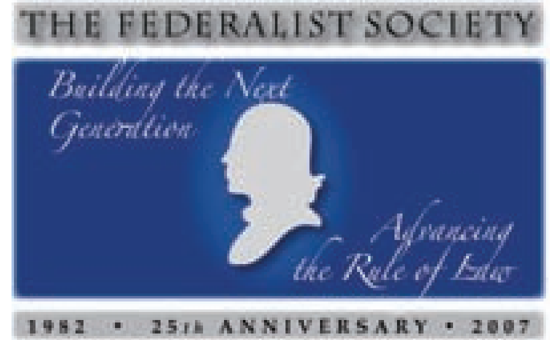
ENGAGE

The Journal of the Federalist Society Practice Groups



Project of the
E.L. Wiegand
Practice Groups

Nota Bene

No Big Surprise: A Review of the Seattle Schools Case
by Harry J.F. Korrell

The Roberts Court's Antitrust Jurisprudence:
The Chicago School Marches On
by Joshua D. Wright

Minding Moral Responsibility: The Supreme Court's
Recent Mental Health Rulings
by Steven K. Erickson

Can the Golden State Catch a Greenhouse Waiver?
by Jonathan H. Adler

How Patents Can Help Those Interested
in the Environment and World Health
by F. Scott Kieff

New Opportunities for Defendants in Securities Class Actions
by Francis J. Menton, Jr.

Book Reviews

David's Hammer, Silence and Freedom,
The Constitution's Text in Foreign Affairs,
Until Proven Innocent,
The Founders on Citizenship & Immigration

Volume 8, Issue 4 October 2007

The Federalist Society for Law and Public Policy Studies is an organization of 40,000 lawyers, law students, scholars and other individuals located in every state and law school in the nation who are interested in the current state of the legal order. The Society takes no position on particular legal or public policy questions, but is founded on the principles that the State exists to preserve freedom, that the separation of governmental powers is central to our Constitution and that it is emphatically the province and duty of the Judiciary to say what the law is, not what it should be.

The Federalist Society takes seriously its responsibility as a non-partisan institution engaged in fostering a serious dialogue about legal issues in the public square. We publish original scholarship on timely and contentious issues in the legal or public policy world, in an effort to widen understanding of the facts and principles involved and to continue that dialogue.

Positions taken on specific issues in publications, however, are those of the author, not reflective of an organization stance. ENGAGE presents articles, white papers, speeches, reprints and panels on a number of important issues, but these are contributions to larger ongoing conversations. We invite readers to submit opposing perspectives or views to be considered for publication, and to share their general responses, thoughts and criticisms by writing to us at info@fed-soc.org. Additionally, we happily consider letters to the editor.

For more information about The Federalist Society, please visit our website: www.fed-soc.org.

The Federalist Society For Law and Public Policy Studies

Directors/Officers

Steven G. Calabresi, *Chairman*
Hon. David M. McIntosh, *Vice Chairman*
Gary Lawson, *Secretary*
Brent O. Hatch, *Treasurer*
Eugene B. Meyer, *President*
. . .
T. Kenneth Cribb, *Counselor*

Board of Visitors

Hon. Robert H. Bork, *Co-Chairman*
Hon. Orrin G. Hatch, *Co-Chairman*
Lillian BeVier
Hon. Lois Haight Herrington
Hon. Donald Paul Hodel
Hon. Frank Keating
Harvey C. Koch
Robert A. Levy
Hon. Edwin Meese, III
Hon. Theodore B. Olson
Andrew J. Redleaf
Hon. Wm. Bradford Reynolds
Nicholas Quinn Rosenkranz
Hon. Gale Norton

Staff

President
Eugene B. Meyer

Executive Vice President
Leonard A. Leo

Senior Vice President & Faculty Director
Lee Liberman Otis

Lawyers Division
Dean Reuter, Practice Groups Director
Lisa Budzynski, Lawyers Chapters Director
Mia Reynolds, State Courts Project Director
David C.F. Ray, Associate Director
Juli Nix, Deputy Director
Alyssa Luttjohann, Assistant Director
Deborah O'Malley, Assistant Director
John Paul Fox, Assistant Director
Marisa Maleck, Assistant Director

Student Division
Peter Redpath, Director
Sarah Roderick, Associate Director
Kyle Reini, Assistant Director
Elizabeth LeRoy, Assistant Director

Pro Bono Center Director
Peggy Little

International Affairs
James P. Kelly, III, Director
Flavius Mihaies, Research Associate

Publications Director
P.C. Aigner

Finance Director
Douglas C. Ubben

Development
Patty Price, Director
Cindy Searcy, Associate Director
Ellen Fuller, Assistant Director

Information Technology Director
C. David Smith

Membership Director
Terry Archambeault

Office Management
Rhonda Moaland, Director
Matt Nix, Assistant Director

Engage

Volume 8, Issue 4

Letter from the Editor . . .

Engage, the journal of the Federalist Society Practice Groups, provides original scholarship on current, important legal and policy issues. The journal is a collaborative effort, involving the hard work and voluntary dedication of each of the organization's fifteen Practice Groups. Through its publication, the Groups aim to contribute to the marketplace of ideas in a way that is collegial, measured, and insightful—and hope to spark a higher level of debate and discussion than is all too often found in today's legal community.

This is the first time that we have been able to offer a fourth issue in a single volume. We are delighted to do so, in this, our twenty-fifth anniversary year. To celebrate this milestone, we have put on a number of special events this year: a televised, moderated discussion of various approaches to constitutional interpretation with U.S. Supreme Court Justices Antonin Scalia and Stephen Breyer, two day-long conferences on the legacy of former Attorney General Edwin Meese, III, and Judge Robert H. Bork, and, finally, *Originalism*, a newly released volume of speeches, essays, and debates, collected by our co-founder, chronicling the discussion over proper constitutional theory this last quarter-century.

Audio and video for all of these events can be found at our Multimedia Archive—a new feature of our recently redesigned website (www.fed-soc.org), intended to be a resource for our student, lawyer, and faculty members. Further information about the book can also be found there, and the latest preview of our 25th annual National Lawyers Convention in November. If you have not signed up already, please do so. This year's convention is sure to gather an extraordinary cast of panelists and speakers, including four U.S. Supreme Court Justices. Also new to the website are two projects of our Faculty Division, SCOTUScast and Originally Speaking—providing commentary on Supreme Court decisions, and cases before the High Court, as they occur. Members can podcast the former through iTunes or the RSS feed in the SCOTUScast section of our website. We hope all these new features provide our members with the high standard of scholarship and balance that they have come to expect from the Society.

As always, readers can expect upcoming issues of Engage to feature articles on matters of importance to them. We hope that you find the work in these pages well-crafted and informative, and strongly encourage you to send us your feedback at info@fed-soc.org.

Volume 8, Issue 4 October 2007

Engage

Administrative Law & Regulation

The Roberts Court Wades into Products Liability Preemption Waters: *Reigel v. Medtronic, Inc.*
by Catherine Sharkey............ 4

Civil Rights

No Big Surprise: A Review of the Seattle Schools Case
by Harry J.F. Korrell............ 11

Can the Federal Transportation DBE Program Be Narrowly Tailored to Remedy Discrimination?
by George R. La Noue............ 18

Point-Counterpoint: Affirmative Action Hiring Practices in the New Court Era
Someone Should Sue *by Roger Clegg*............ 25
Affirmative Action: Legally Sound and Good for American Business
by Wade Henderson............ 26

Corporations, Securities & Antitrust

The Roberts Court's Antitrust Jurisprudence: The Chicago School Marches On
by Joshua D. Wright............ 29

The End of the Road for *Dr. Miles*: *Leegin Creative Leather Products, Inc. v. PSKS, Inc.*
by Robert T. Miller............ 40

Majority Voting: Acting in the Best Interests of the Stockholders, Or Following the Herd?
by Donald R. Moody & J. Ammon Smartt............ 53

Criminal Law & Procedure

Minding Moral Responsibility: The Supreme Court's Recent Mental Health Rulings
by Steven K. Erickson............ 59

Genetic Information and Privacy Interests: The DNA Fingerprinting Act
by Ronald J. Rychlak............ 64

Sanctuary Cities and the Second Circuit's Challenge in *New York v. U.S.*
by Marc Harrold............ 68

Making Prisons Compete: Private Prisons and the Demands on the Public System
by Geoffrey F. Segal & Alexander McCobin............ 71

Environmental Law & Property Rights

Can the Golden State Catch a Greenhouse Waiver?
by Jonathan H. Adler............ 75

Do Inclusionary Zoning Laws Violate *Nollan*, *Dolan*,
And the Doctrine of Unconstitutional Conditions?
by James S. Burling............ *83*

Federalism & Separation of Powers

Federal Aid to the States *by Chris Edwards*............ 91

Financial Services & E-Commerce

Securities and Exchange Commission Issues Guidance on Management's
Evaluation of Internal Control Over Financial Reporting
by Daniel Fisher............ 100

Free Speech & Election Law

Violence ≠ Obsenity

by Geoffrey Hymans 103

Intellectual Property

How Patents Can Help Those Interested in the Environment and World Health

by F. Scott Kieff 106

International & National Security Law

Public Committee Against Torture in Israel v. Government of Israel:
U.S. Military Commissions Through the Lens of the Supreme Court of Israel

by Michael Newton 112

Labor & Employment Law

Sarbanes-Oxley Whistleblower Provisions: A Study in Statutory Construction

by Jay P. Lechner 121

Litigation

New Opportunities for Defendants in Securities Class Actions

by Francis J. Menton, Jr. 130

West Virginia Supreme Court Strikes Down Learned Intermediary Rule

by James M. Beck & Theodore H. Frank 134

Professional Responsibility & Legal Education

Attorneys' Fees in Class Actions: The Problem Remains

by Jack Park 138

Religious Liberties

"Equal Access"? *Faith Center Church Evangelistic Ministries v. Glover*
And the Use of Public Facilities for Religious Worship

by Douglas G. Smith 142

Telecommunications & Electronic Media

Will the Federal Communications Commission Broadcast Flag Order Be Resurrected?

by Stephen T. Yelverton 145

Book Reviews

David's Hammer: The Case for an Activist Judiciary by Clint Bolick

First Review by Timothy Sandefur 147

Second Review by Craig S. Lerner 148

Silence and Freedom by Louis Michael Seidman

Reviewed by Paul Horwitz 150

The Constitution's Text in Foreign Affairs by Michael D. Ramsey

Reviewed by Jeremy Rabkin 152

Until Proven Innocent: Political Correctness and the Shameful Injustices of the Duke Lacrosse Rape Case by Stuart Taylor, Jr. & K.C. Johnson

Reviewed by Michael Madigan 154

The Founders on Citizenship and Immigration: Principles and Challenges for America by Edward J. Erler, Thomas G. West & John Marini

Reviewed by Margaret Stock 155

Administrative Law and Regulation

The Roberts Court Wades into Products Liability Preemption Waters: *Reigel v. Medtronic, Inc.*

*By Catherine M. Sharkey**

With *Riegel v. Medtronic, Inc.* (06-179), the Roberts Court makes its inaugural foray this term into the realm of federal preemption of state-law products liability claims.[1] The Supreme Court's products liability preemption jurisprudence is a small but expanding area that can trace its beginnings to the early 1990s with *Cipollone v. Liggett Group, Inc.*,[2] and continues, most recently, through the 2005 decision of *Bates v. Dow Agrosciences, LLC*.[3] The regulation of public health and safety via common law tort actions falls within the traditional purview of the states. In recent decades, however, the federal government has played an increasingly significant role in the regulation of products. In 1976, Congress enacted the Medical Device Amendments (MDA) "to provide for the safety and effectiveness of medical devices intended for human use."[4]

The Court granted certiorari in *Riegel* to decide:

> [w]hether the express preemption provision of the Medical Device Amendments to the Food, Drug, and Cosmetic Act preempts state-law claims seeking damages for injuries caused by medical devices that received premarket approval from the Food and Drug Administration.

Returning to the field of medical devices, the Court will answer a question left open by its decade-old opinion in *Medtronic v. Lohr*,[5] which held that state-law tort claims as to medical devices subject to a less rigorous pre-market notification process (as opposed to pre-market approval) were *not* preempted.

The fractious opinion in *Medtronic* sets the scene for the issues pending before the Court in *Riegel*. The Court will decide whether the FDA's pre-market approval process for medical devices creates federal preemptive "requirements" sufficient to preempt state common-law tort actions. Tasked with interpreting the language and scope of the express preemption provision of the MDA, the Roberts Court returns to a contentious area of jurisprudence. The Court is presented with an opportunity to resolve looming tensions between competing canons of statutory interpretation: the presumption against preemption and *Chevron* deference to federal agency interpretations of ambiguous statutes.

I. Facts and Procedural Posture

Charles Riegel, a cardiac patient, sued Medtronic, Inc., the manufacturer of a balloon catheter used during his angioplasty. The balloon catheter, a "Class III" medical device, received premarket approval from the FDA in 1994.[6] The catheter ruptured after being over-inflated, causing Riegel extensive injuries and permanent disabilities. Riegel, joined by his wife,[7] brought a number of state-law claims against the medical device manufacturer, including negligent design, testing, manufacture, distribution, labeling, marketing, and sale of the catheter; strict liability; breach of express warranty; breach of implied warranty; and loss of consortium.

The Second Circuit affirmed (2-1) the district court's grant of summary judgment in favor of Medtronic, holding that the majority of Riegels' claims were preempted.[8] According to the court, because the FDA had expressly found the design of the device safe and effective, and approved the precise wording on the label, the plaintiffs' claims were preempted, except insofar as they alleged manufacturer did not adhere to specs submitted to the FDA in manufacturing the specific device used in his operation.[9]

II. Split in the Circuits

With its pro-preemption holding, the Second Circuit joined the large majority of federal circuits to have decided the issue. The Third, Fifth, Sixth, Seventh, and Eighth Circuits have all held that pre-market approval of a medical device preempts state tort claims that challenge the safety or efficacy of a product that was designed, manufactured, and labeled in conformity with the approval process.[10] The Eleventh Circuit is the outlier, having found that comparable state law claims were not preempted.[11]

The Solicitor General counseled the Court against granting review on the grounds that the Second Circuit's decision was correct and that the only cases on the short end of the lopsided split "predate most of the other cases addressing the question, and they were issued without the benefit of the FDA's current judgment that premarket approval of a Class III device imposes federal "requirements" that should be given preemptive effect."[12] Indeed, in *Goodlin*, the Eleventh Circuit took note of a 1997 FDA proposed rule (later withdrawn)[13] that would have enshrined the FDA's earlier anti-preemption view, and found it "unsettling that the agency charged with conducting PMA review has doubts regarding whether an approval pursuant to that process should preclude subsequent state tort liability."[14] Moreover, the Solicitor General reminded the Court that it had "repeatedly denied certiorari petitions that presented questions concerning the preemptive effect of the FDA's issuance of premarket approval for Class III medical devices."[15]

III. MDA Express Preemption and *Medtronic v. Lohr*

The starting point to understanding the issues at stake in *Riegel v. Medtronic, Inc.*, as with its forbear, *Medtronic v. Lohr*, is the express preemption provision set forth in Section 360k(a) of the 1976 Medical Device Amendments (MDA) to the Federal Food, Drug, and Cosmetic Act (FDCA). Section 360k(a) directs preemption of "any [state] *requirement*" "which is different from, or in addition to, any *requirement* applicable under [the FDCA] to the device."[16]

** Catherine M. Sharkey is a Professor of Law at New York University.* *This essay draws from Catherine M. Sharkey, Products Liability Preemption: An Institutional Approach, 76 Geo. Wash. L. Rev. (forthcoming 2008).* *Jaime Sneider of Columbia Law School provided helpful assistance.*

The express preemption clause raises two interpretive issues. First, does the domain of state-law "requirements" that are potentially subject to preemption include common-law tort actions as well as positive enactments of statutory or regulatory law? Second, what precisely constitutes a federal "requirement applicable... to the device" sufficient to preempt state law?

In *Medtronic*, five of the Justices (represented by the concurring and dissenting Justices) answered "yes" to the first question, concluding that the MDA will sometimes preempt state-law tort causes of action. Any ambiguity on this question[17] has arguably now been put to rest in *Bates v. Dow Agrosciences, LLC*.,[18] where the Court reiterated that "the term 'requirements' in [the express preemption provision of the Federal Insecticide Fungicide and Rodenticide Act] reaches beyond positive enactments, such as statutes and regulations, to embrace common-law duties."[19] On the second question, the *Medtronic* Court's majority held that the FDA's pre-market notification process—to which the medical device, a pacemaker, had been subjected—did not create preemptive federal "requirements" applicable to the device.

IV. Preemption Analysis

Preemption is the fiercest battle in products liability litigation today. With the stroke of a pen, Congress could definitively determine when its product regulations displace state common law. But instead, time and again, Congress punts, leaving open the key question of the extent to which federal standards and regulations preempt state common-law remedies. A textualist statutory interpretation approach to preemption will likely come up short. Instead, the products liability preemption inquiry is multidimensional, involving layers of legal and policy issues, beginning with interpretation of the statutory language, but reaching beyond to issues of regulatory policy, federalism, and the level of deference accorded federal agency actions and interpretations.

It is difficult to demonstrate that any consistent principle or explanatory variable emerges from the Supreme Court's products liability preemption jurisprudence. As Professor Jack Goldsmith has aptly summed up: "The [statutory interpretation] canons have an uncertain justification… and they probably conceal more than they enlighten about what drives the judicial decision to preempt or not."[20] Even the "presumption against preemption"—perhaps the leading contender for consistency in the traditional state realm of torts—breaks down in the products realm, rearing its head with gusto in some cases, but oddly quiescent in others. Strangely eluding detection to date, the influence of the position of the relevant federal agency may provide a tighter explanatory fit. To be sure, the Supreme Court has been less than forthcoming about its reliance upon the views of the agency, sometimes put forward in official regulations, but more often simply in amicus briefs submitted by the Solicitor General (which may or may not get explicit mention by the Court).

The Roberts Court ducked an opportunity last term to provide guidance on the potential clash that arises in preemption cases between the presumption against preemption and *Chevron* deference to agency interpretation. The Court granted certiorari in *Wachovia Bank, N.A. v. Watters* on whether the interpretation of the Comptroller of the Currency that its regulation preempted state laws regulating mortgage lending as applied to operating subsidiaries of national banks was entitled to *Chevron* deference.[21] The Court dodged the issue, holding that state laws were preempted by the National Banking Act, independent of the OCC's regulation,[22] prompting a vigorous dissent from Justice Stevens (joined by Chief Justice Roberts and Justice Scalia) that "[w]hatever the Court says, this is a case about an administrative agency's power to preempt state laws."[23] And so, the contentious *Chevron* deference issue was put off for another day.

A. The Presumption Against Preemption

The touchstone of conventional preemption analysis is congressional intent. The "presumption against preemption" in areas "traditionally occupied by the States" has acquired preeminent status as an interpretive canon. Given Congress' track record in failing to address squarely the question of preemption in the products realm, interpretive canons such as the presumption against preemption should, at least in theory, take on added significance.

To date, however, the Court's application of the presumption has been haphazard at best. The presumption appears to do the yeoman's work in some products cases, while eluding mention altogether in others.[24] The *Medtronic* Court, for example, began its preemption analysis with an invocation of the presumption: "[B]ecause the States are independent sovereigns in our federal system, we have long presumed that Congress does not cavalierly pre-empt state-law causes of action."[25] Paradoxically, the Court has applied the presumption when interpreting express preemption provisions (as in *Medtronic*), but not when called upon to engage in implied preemption analysis, where it would seem more warranted given the absence of express statutory language.[26] And it is striking that in the single implied preemption case in which is invoked (*Buckman Co. v. Plaintiffs' Legal Committee*) it is the presumption for the purpose of disavowing it, given the primacy of the federal interest at stake.[27]

Moreover, the subdued role played by the presumption against preemption fits a wider empirical pattern, whereby, for decades, roughly fifty-fifty odds have prevailed in Supreme Court preemption decisions.[28] Moreover, the preemption rate actually *increases* (to greater than 60% odds) when considering preemption of state common law tort claims—a realm in which the putative anti-preemption presumption should be at its zenith.[29]

B. Deference to the FDA

With the presumption against preemption playing a mixed interpretive role at best, deference to the relevant federal agency charged with administering a particular statute emerges as a contender. The role played by the FDA might be significant in two different respects. First, there is the level of scrutiny to which it subjects medical devices before manufacturers are allowed to market them, and the relationship between this regulatory action and the creation of federal preemptive "requirements." Second, the FDA plays a distinct interpretive role as administer of the MDA, and has a variety of means at its disposal to express its position on preemption, from formal

notice-and-comment rulemaking to less formal interpretive statements and preambles to litigation briefs.

i. FDA Regulatory Action

Both the pacemaker at issue in *Medtronic* and the balloon catheter at issue in *Riegel* are "Class III" medical devices regulated by the FDA. A crucial distinction, nonetheless, emerges with respect to the stringency of FDA regulatory review of the respective medical devices.

The FDA's review of the pacemaker at issue in *Medtronic* consisted solely in its determination that the device was "substantially equivalent" to a device that was on the market before 1976 (the effective date of the MDA). Known as the "premarket notification" process (or, alternatively, § 510(k) process), the FDA's review focuses narrowly on equivalence as opposed to safety and effectiveness. It is a streamlined process, completed in an average of twenty hours, that allows manufacturers to avoid the more stringent pre-market approval [PMA] process as a kind of accommodation "to prevent manufacturers of grandfathered devices from monopolizing the market while new devices clear the PMA hurdle, and to ensure that improvements to existing devices can be rapidly introduced into the market."[30] Although designed as a limited exception, in practice most new medical devices are approved via the pre-market notification process.[31]

The balloon catheter at issue in *Riegel* was subjected to the full-bodied pre-market approval process mandated for Class III devices that do not fall within the grandfathering exception. In stark contrast to the pre-market notification process, the PMA process is rigorous, requiring manufacturers to submit detailed information regarding the safety and efficacy of the medical device and demanding an average of 1200 hours of FDA review time per submission:

> [A] manufacturer must submit a PMA application containing full reports of investigations of the device's safety and effectiveness; a statement of the components and principles of operation of the device; a comprehensive description of the methods of manufacture, processing, packing, and installation of the device; and the proposed labeling for the device. In determining whether to approve a PMA application, the FDA considers the information submitted by the manufacturer as well as other information known to the agency. The FDA may also request additional information from the manufacturer, and it may consult with a scientific advisory committee made up of outside experts.[32]

The PMA process culminates in a finding by the FDA that there is a "reasonable assurance" that the device is both safe and effective, so long as the device is used in accordance with any conditions of use included in the proposed labeling.[33]

ii. FDA Interpretation

But the FDA's role in regulating medical devices goes beyond that of conducting the risk-risk analyses and assuring the safety of medical devices. It has also assumed the mantle of statutory interpreter of the MDA. Such a role was justified, in the eyes of the *Medtronic* Court, "[b]ecause the FDA is the federal agency to which Congress has delegated its authority to implement the provisions of the Act, the agency is uniquely qualified to determine whether a particular form of state law 'stands as an obstacle to the accomplishment and execution of the full purposes and objectives of Congress, and therefore, whether it should be pre-empted."[34]

The FDA issued formal regulations construing the scope of the express preemption provision, which cabin its preemptive force to instances where the FDA has established "specific counterpart regulations or... other specific requirements applicable to a particular device."[35] The regulation further provides that the MDA "does not preempt State or local requirements of general applicability where the purpose of the requirements relates either to other products in addition to devices."[36]

In addition, the FDA often weighs in contemporaneously on factors that arguably determine the preemptive effect of its regulatory actions. In *Medtronic*, for example, at the time the FDA issued its "substantial equivalence" letter to the manufacturer, "[t]he agency emphasized... that this determination should not be construed as an endorsement of the pacemaker lead's safety."[37]

Finally, the FDA has shared its views before courts (including the Supreme Court) tasked with deciding preemption questions. In *Medronic*, the FDA adopted a narrowly constricted view of its preemptive power: "Neither the FDCA nor the FDA's regulations prescribe criteria for the design of devices. The design of a device originates with its manufacturer."[38] In other words, in this instance—where its review consisted solely of a "substantial equivalence" finding—it ceded regulation of the design of medical devices to state common law.

Reliance upon federal agency interpretation at each of these three levels—issuance of regulations regarding preemptive scope; contemporaneous views interpreting regulatory action; and expressions of views in amicus briefs before courts—is contentious (with increasing degrees in the move from formal regulations to less formal interpretive positions). In deciding products liability preemption issues, the Supreme Court has been influenced by agency positions, but has not always been upfront about the degree to which the agency's view is dispositive.

The *Medtronic* Court, for example, resisted the idea that the language of the express preemption provision of the MDA decided the preemption issue, given the inherent ambiguity over what is meant by the statutory term "requirement." In the words of Justice Breyer (in concurrence): "Congress must have intended that courts look elsewhere for help as to just which federal requirements pre-empt just which state requirements, as well as just how they might do so."[39] Here, the Court turned to the FDA—the federal agency charged with administering the MDA[40]—for guidance: "The ambiguity in the statute... provide[s] a sound basis for giving *substantial weight* to the agency's view of the statute."[41]

In reaching its ultimate posture against preemption of the plaintiffs' (the Lohrs') state-law claims, the Court relied upon the fact that "[t]he FDA regulations interpreting the scope of § 360k's pre-emptive effect support the Lohrs' view, and our interpretation of the pre-emption statute is *substantially informed* by those regulations."[42] The *Medtronic* plurality emphasized the "critical importance of device specificity" in its understanding of the MDA preemption scheme.[43]

In dissent, Justice O'Connor (joined by then-Chief Justice Rehnquist and Justices Thomas and Scalia) railed against the Court's analysis, noting that its opinion neither explicitly relied on the FDA regulations nor offered "any sound basis for why deference would be warranted."[44] Sharpening its disagreement, the dissent elaborated:

> Apparently recognizing that *Chevron* deference is unwarranted here, the Court does not admit to deferring to these regulations, but merely permits them to "infor[m]" the Court's interpretation. It is not certain that an agency regulation determining the pre-emptive effect of *any* federal statute is entitled to deference, but one pertaining to the clear statute at issue here is surely not.[45]

The Second Circuit, in *Riegel,* deferred to the FDA in both its regulatory and interpretive capacities. With respect to state law, the court held that the common law tort actions would impose state-law requirements "different from, or in addition to" the federal requirements.[46] With respect to federal law, the court construed the term "requirement" to encompass product specifications set forth in the PMA application that is submitted by the manufacturer and approved by the FDA.[47] The court reasoned that "[o]nce the PMA process is complete, all PMA-approved devices are subject to the same federal device-specific regulation: complying with the standards set forth in their individual approved PMA applications."[48] Moreover, whereas substantial-equivalence "does not reflect the FDA's determination that the device should 'take any particular form for any particular reason,' the PMA process expressly provides the FDA with the power to require the device to take a particular form in order to be approved as safe and effective."[49] Finally, the court noted that its pro-preemption conclusion was "further supported by the FDA's recent determination that preempted is warranted with respect to this universe of cases, as indicated by the content of the May 14, 2004 amicus brief that the FDA submitted upon request to the Third Circuit in connection with the *Horn* case."[50]

V. Questions for the Roberts Court

The issue before the Roberts Court is whether the PMA process gives rise to specific federal manufacturing and labeling requirements in conflict with state common law actions. In *Medtronic*, the presumption against preemption and deference to the FDA pointed in the same anti-preemption direction. In *Riegel*, the potential clash rises to the fore, given the FDA's pro-preemption position.

One threshold question—as yet unanswered by the Supreme Court—is whether the FDA's position on preemption matters, and, if so, whether the formality with which it conveys its view affects the level of deference courts should accord. It is certainly an understatement to suggest that the degree to which courts regard the FDA's position as dispositive or persuasive remains unclear. In *Horn* and *Riegel*, the Third and Second Circuits, respectively, deferred to the FDA's position,[51] while in *McMullen* the Seventh Circuit disclaimed reliance on the FDA's position in support of its pro-preemption holding.[52] Moreover, the dissents in *Horn* and *Riegel* sharply criticized judicial reliance on informal agency views.[53]

To the extent the issue surfaced in *Medtronic*, the Court was deeply divided. The plurality accorded some degree of deference to the FDA's formal regulation; Justice Breyer would apparently give wide berth to federal agencies to communicate their views on preemption formally via regulations or more informally via preambles and the like:

> [T]his Court has previously suggested that, in the absence of clear congressional command as to pre-emption, courts may infer that the relevant administrative agency possesses a degree of leeway to determine which rules, regulations, or other administrative actions will have pre-emptive effect…. It can communicate those intentions, for example, through statements in "regulations, preambles, interpretive statements, and responses to comments[.]"[54]

While agency briefs and preambles arguably lack "the force of law" necessary to warrant *Chevron* mandatory deference,[55] the doctrine on deference to agency preambles and amicus briefs—particularly in the realm of preemption—is far from pellucid.

A second question is to what extent is an agency's change in position relevant to the deference issue. At the time of its intervention in *Medtronic* (1996), the FDA recommended that the Court leave open the issue whether the PMA process (as opposed to the pre-market notification process) engendered preemptive federal requirements.[56] In fact, the FDA had already taken the position that the PMA process did *not l*ead to preemption of state common law claims.[57] And a year after *Medtronic*, in an amicus brief urging the Court to grant certiorari in another medical devices case (where the catheter device at issue had gone through the full PMA process), the FDA maintained its view that "Section 360k does not... preempt respondents' common law tort claims."[58] The FDA's position at that time (as defended by then-FDA Chief Counsel) was driven by a view that "FDA product approval and state tort liability usually operate independently, each providing a significant, yet distinct, layer of consumer protection…. Even the most thorough regulation of a product such as a critical medical device may fail to identify potential problems presented by the product."[59]

The FDA subsequently changed its view and first articulated its new pro-preemption position for PMA devices in 2004 in an amicus brief before the Third Circuit Court of Appeals in *Horn v. Thoratec Corp*. To justify its change in position, the FDA explained:

> [B]ased on further analysis of the relevant legal and policy issues by FDA, the agency charged with administering the MDA and implementing its preemption provisions—as well as the recent rulings by several courts of appeals and state courts—the Government has instead determined that state tort claims such as those raised here are indeed preempted with respect to FDA-approved devices.[60]

According to the FDA, the PMA process creates specific federal requirements because, following approval, the device's "attributes are fixed in place, as they can be materially changed only with FDA approval."[61] The Third Circuit relied heavily on the FDA's views in holding plaintiff's state law claims preempted.[62] Since that time, while the FDA has steadfastly maintained its pro-preemption position, it has not attempted to embody its view in a new regulation or else amend the prior

regulation (§ 808.1(d)) that provided justification for the *Medtronic* Court to construe the MDA's express preemption clause narrowly to imbue only "device specific" requirements with preemptive effect.

This inconsistency appeared to be relevant to the Court's anti-preemption position in *Bates v. Dow Agrosciences, LLC*.[63] The Court reasoned that "[t]he notion that FIFRA contains a nonambiguous command to pre-empt the types of tort claims that parallel FIFRA's misbranding requirements is particularly dubious given that just five years ago the United States advocated the interpretation that we adopt today."[64] However, in *National Cable & Telecommunication Association v. Brand X Internet South*, a fairly recent decision (post-dating the Court's line of products liability preemption cases), the Court held that agency inconsistency is not relevant to a court's decision whether to accord *Chevron* deference to an agency interpretation.[65] Instead, the Court continued, "[u]nexplained inconsistency is, at most, a reason for holding an interpretation to be an arbitrary and capricious change from agency practice under the Administrative Procedure Act."[66] Three years prior, in *Barnhart v. Walton*, however, the Court stated: "The Agency's regulations also reflect the Agency's own longstanding interpretation... [and thus] should be accorded particular deference."[67] While the relevance of agency consistency as a factor in applying *Chevron* deference is somewhat ambiguous, there is no doubt as to its relevance in applying the weaker "power to persuade" *Skidmore* deference.[68]

In *Riegel*, the Second Circuit was not troubled by the FDA's change of heart; as it explained: "It is certainly true that the FDA previously took a different view, but as the Third Circuit noted in *Horn*, 'an agency may change its course so long as it can justify its change with a "reasoned analysis,"' a standard satisfied here."[69] In its amicus brief filed at the petition stage, the Solicitor General explained: "The FDA has since reexamined the issue and determined that the position it announced at the time of the filing in *Kernats* was erroneous."[70] In addition to embracing the reasons set forth in the *Horn* amicus brief, the Solicitor General claimed that "[t]he government's position in *Kernats* is also inconsistent with the risk-management principles that the FDA currently follows."[71]

CONCLUSION

Riegel presents an opportunity for the Court not only to revisit *Medtronic* and the issue of express preemption under the MDA but also to begin to fashion a framework for preemption jurisprudence that reconciles the often competing demands of the presumption against preemption and deference to agency interpretations. Significantly, the Court will also hear a second products liability preemption case, *Warner-Lambert Co. LLC v. Kent*, a pharmaceutical drug case, where the argument for preemption lies in implied (as opposed to express) grounds, and which calls for the Court to interpret the scope of its previous holding in *Buckman* in preempting claims of fraud on the agency.[72] Finally, the Court has called for the views of the Solicitor General in a pharmaceutical drug preemption case, *Levine v. Wyeth*.[73] Preemption in the pharmaceutical drug context is even more fraught than that of medical devices. Unlike the MDA, the FDCA contains no express preemption provision that pertains to drugs. The FDA has taken a similarly aggressive pro-preemption stance, expressing its views in a preamble to a rule on the form and content of drug labels as well as in amicus briefs before courts. The case presents a clash (analogous to that in *Riegel*) between the presumption against preemption and deference to agency views, but in a context that has not yet ripened in the lower courts.[74]

Endnotes

1 The Court has granted certiorari in a second products liability case, Warner-Lambert Co. LLC v. Kent (No. 06-1498), 2007 WL 1420397 (Sept. 25, 2007).

2 505 U.S. 504 (1992).

3 544 U.S. 431 (2005).

4 90 Stat. 539.

5 518 U.S. 470 (1996).

6 Class III medical devices "presen[t] a potential unreasonable risk of illness or injury" or are "purported or represented to be for a use in supporting or sustaining human life or for a use which is of substantial importance in preventing impairment of human health." 21 U.S.C. § 360c(a)(1)(C).

7 Mr. Riegel in fact died in 2004. The Supreme Court granted the untimely motion of Donna Riegel (which was opposed by Medtronic) to be substituted in place of Charles Riegel. Chief Justice Roberts and Justice Scalia dissented on the ground that the motion was filed more than six months after petitioner's death, in violation of Supreme Court Rule 35.1.

8 Riegel v. Medtronic, Inc., 451 F.3d 104, 119 (2d Cir. 2006), *cert. granted*, 127 S. Ct. 3000 (2007).

9 *Id.* at 106.

10 Horn v. Thoratec Corp., 376 F.3d 163, 169 (3d Cir. 2004); Martin v. Medtronic, Inc., 254 F.3d 573, 585 (5th Cir. 2001), *cert. denied*, 534 U.S. 1078 (2002); Kemp v. Medtronic, Inc., 231 F.3d 216, 227-28 (6th Cir. 2000), *cert. denied*, 534 U.S. 818 (2001); McMullen v. Medtronic, Inc. 421 F.3d 482, 487-88 (7th Cir. 2005), *cert. denied*, 126 S. Ct. 1464 (2006); Brooks v. Howmedica, Inc., 273 F.3d 785, 798 (8th Cir. 2001), *cert. denied*, 535 U.S. 1056 (2002).

11 Goodlin v. Medtronic, Inc. 167 F.3d 1367, 1382 (11th Cir. 1999).

12 Brief for the United States as Amicus Curiae at 8, Riegel v. Medtronic, Inc. (No. 06-179), 2007 WL 1511526 (May 23, 2007) [hereinafter U.S. *Riegel* Brief].

13 62 Fed. Reg. 65,384 (1997), *withdrawn*, 63 Fed. Reg. 39,789 (1998).

14 *Goodlin*, 167 F.3d at 1375 n.15.

15 U.S. *Riegel* Brief at 20.

16 The section provides in full:

§ 360k. State and local requirements respecting devices

(a) General rule

Except as provided in subsection (b) of this section, no State or political subdivision of a State may establish or continue in effect with respect to a device intended for human use any requirement—

(1) which is different from, or in addition to, any requirement applicable under this chapter to the device, and

(2) which relates to the safety or effectiveness of the device or to any other matter included in a requirement applicable to the device under this chapter.

21 U.S.C. § 360k(a).

17 There is some ambiguity with respect to the *Medtronic* plurality's take on this point. The plurality (consisting of Justices Stevens, Kennedy, Souter, and Ginsburg) seems to accept, albeit grudgingly, that the MDA may preempt some common-law tort actions, when it states its view that "§ 360(k) simply was not intended to pre-empt most, let alone all, general common-law duties enforced by damages actions." *Medtronic*, 518 U.S. at 491. But the plurality ultimately reserves final judgment on this issue, reasoning:

> [I]t is apparent that few, if any, common-law duties have been pre-empted by this statute. It will be rare indeed for a court hearing a common-law cause of action to issue a decree that has "the effect of establishing a substantive requirement for a specific device." Until such a case arises, we see no need to determine whether the statute explicitly pre-empts such a claim.

Id. at 502-03.

Justice Breyer, in a separate concurrence, provides a clear "yes" response:

> I believe that ordinarily, insofar as the MDA pre-empts a state requirement embodied in a state statute, rule, regulation, or other administrative action, it would also pre-empt a similar requirement that takes the form of a standard of care or behavior imposed by a state-law tort action. It is possible that the plurality also agrees on this point, although it does not say so explicitly.

Id. at 504-05 (Breyer, J., concurring). The dissent (Justice O'Connor, joined by then-Chief Justice Rehnquist, and Justices Scalia and Thomas) likewise have no difficulty reading "requirement" to include state-law causes of action. *Id*. at 509 (O'Connor, J., dissenting) ("I conclude that state common-law damages actions do impose 'requirements' . . . .").

18 544 U.S. 431 (2005).

19 *Id.* at 443 ("The phrase '[n]o requirement or prohibition' sweeps broadly and suggests no distinction between positive enactments and common law; to the contrary, those words easily encompass obligations that take the form of common-law rules") (quoting *Cipollone*, 505 U.S. at 521 (plurality opinion)). Clouding the issue, however, the *Bates* Court retreated from its statement by concluding that "an event, such as a jury verdict, that merely motivates an optional decision is not a requirement." *Id.* at 432.

20 Jack Goldsmith, *Statutory Foreign Affairs Preemption*, 2000 Sup. Ct. Rev. 175, 200.

21 126 S. Ct. 2900 (No. 05-1342) (June 19, 2006).

22 127 S. Ct. 1559, 1572 (2007) (opining that the deference issue was "beside the point, for under our interpretation of the statute, the level of deference owed to the regulation is an academic question").

23 *Id.* at 1585 (Stevens, J., dissenting); *id.* at 1584 ("No case from this Court has ever applied such a deferential standard to an agency decision that could so easily disrupt the federal-state balance.")

24 *See, e.g.*, Freightliner Corp. v. Myrick, 514 U.S. 280 (1995) (no mention of presumption); Geier v. American Honda Motor Co., 529 U.S. 861 (2000) (same); Sprietsma v. Mercury Marine, 537 U.S. 51 (2002) (same).

25 *Medtronic*, 518 U.S. at 485. Most recently, the presumption reared its head again in *Bates*, 544 U.S. at 449 (quoting *Medtronic*); *id*. at 499 ("Even if Dow had offered us a plausible alternative reading of [the express preemption provision]—indeed, even if its alternative were just as plausible as our reading of that text—we would nevertheless have a duty to accept the reading that disfavors pre-emption.").

26 Thus, the presumption prevails in *Cipollone*, *Medtronic*, and *Bates*—all decided on express preemption grounds; but is not invoked in *Geier* or *Sprietsma*—the seminal implied preemption cases.

27 531 U.S. 341, 347-48 (2001) (refusing to apply presumption given that "[p]olicing fraud against federal agencies is hardly 'a field which the States have traditionally occupied,'" and "the relationship between a federal agency and the entity it regulates is inherently federal in character because the relationship originates from, is governed by, and terminates according to federal law").

28 *See* Michael S. Greve & Jonathan Klick, *Preemption in the Rehnquist Court: A Preliminary Empirical Assessment*, 14 Sup. Ct. Econ. Rev. 43, 57 (2006) (in an empirical analysis of universe of 105 preemption decisions from the Rehnquist Court era, finding preemption in 52%); *accord* Note, *New Evidence on the Presumption Against Preemption: An Empirical Study of Congressional Responses to Supreme Court Decisions*, 120 Harv. L. Rev. 1604, 1605 (2007) ("Between the 1983 and 2003 Terms the Supreme Court decided 127 cases involving federal preemption of state law, finding state law preempted approximately half of the time.") (footnotes omitted).

29 Greve & Klick, *supra* note 29, at 52 (finding 62.5% preemption rate in 32 cases involving preemption of state common law claims; figure increases to 67.6% when cases are restricted to "Second Rehnquist Court," beginning in 1994).

30 *Medtronic*, 518 U.S. at 478 (citing 21 U.S.C. § 360e(b)(1)(B)).

31 *Id.* at 479 ("[T]he House reported in 1990 that 80% of new Class III devices were being introduced to the market through the § 510(k) process and without PMA review."); *see also* Richard C. Ausness, *After You, My Dear Alphonse!: Should the Courts Defer to the FDA's New Interpretation of § 360K(A) of the Medical Device Amendments?*, 80 Tul. L. Rev. 727, 733 (2006) ("Most medical devices do not undergo the FDA PMA process.").

32 U.S. *Riegel* Brief at 2.

33 21 U.S.C. § 360e(d)(2).

34 *Medtronic*, 518 U.S. at 496 (quoting Hines v. Davidowitz, 312 U.S. 52, 67 (1941)); *id.* at 506 (Breyer, J., concurring) (FDA has a "special understanding of the likely impact of both state and federal requirements, as well as an understanding of whether... state requirements may interfere with federal objectives").

35 21 C.F.R. § 808.1(d) (1995).

36 *Id.* § 808.1(d)(1).

37 *Medtronic*, 518 U.S. at 480.

38 Brief for the United States as Amicus Curiae Supporting Respondents at 20, Medtronic v. Lohr, 518 U.S. 470 (1996) (Nos. 95-754, 95-886), 1996 WL 118035 [hereinafter U.S. *Medtronic* Brief].

39 *Medtronic*, 518 U.S. at 505 (Breyer, J., concurring).

40 The FDCA vests authority in the Secretary of Health and Human Services, who subsequently delegated authority to the FDA, "to promulgate regulations for the efficient enforcement of" the Act. 21 U.S.C. § 371(a).

41 *Medtronic*, 518 U.S. at 496 (citing Chevron U.S.A. Inc. v. Natural Resources Defense Council, Inc., 467 U.S. 837 (1984); Hillsborough County v. Automated Medical Laboratories, Inc., 471 U.S. 707, 714 (1985)).

42 *Id.* at 495 (emphasis added).

43 *Id*. at 503.

44 *Id.* at 509 (O'Connor, J., concurring in part and dissenting in part).

45 *Id.* at 512 (O'Connor, J., concurring in part and dissenting in part).

46 Riegel v. Medtronic, Inc., 451 F.3d 104, 119 (2d Cir. 2006).

47 *Id.* at 119.

48 *Id.* at 118.

49 *Id.* at 121.

50 *Id*. at 124-25 (citing *Horn*, 376 F.3d at 177-79; *Medtronic*, 518 U.S. at 496).

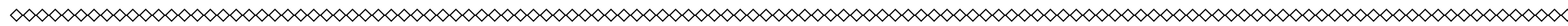
51 *Horn*, 376 F.3d at 170-171; *Riegel*, 451 F.3d at 124-25 .

52 McMullen v. Medtronic, Inc., 421 F.3d 482, 489 n.3 (7th Cir. 2005) ("An agency's urging, however, does not change a permissive provision into a mandatory one.").

53 *Horn*, 376 F.3d at 182 (Fuentes, J., dissenting) ("[A]rguments advanced by the United States in a litigation brief are entitled to 'near indifference,' and are only as persuasive as their own merits dictate."); *Riegel*, 451 F.3d at 129 (Pooler, J. dissenting) ("[T]he idea that all state tort claims are unambiguously preempted is "particularly dubious" considering it appears that until relatively recently neither the industry nor the FDA thought such claims were preempted.").

54 518 U.S. 470, 505–06 (1996) (Breyer, J., concurring) (citations omitted).

55 *See* United States v. Mead Corp., 533 U.S. 218, 228 (2001)



("[A]dministrative implementation of a particular statutory provision qualifies for *Chevron* deference when it appears that Congress delegated authority to the agency generally to make rules carrying the *force of law*, and that the agency interpretation claiming deference was promulgated in the exercise of that authority.") (emphasis added).

56 U.S. *Medtronic* Brief at 20 n.14 ("There is no occasion in this case to consider the extent to which the PMA process may result in requirements applicable to a device under the FDCA that would trigger preemption under Section 521(a).").

57 *See, e.g.*, Brief for the United States as Amicus Curiae at 14 n.5, Talbott v. Bard, Inc., 63 F.3d 25 (1st Cir. 1995) (No. 94-1951) ("This Court previously indicated that premarket approval procedures that apply to Class III devices constitute 'specific requirements' within the meaning of 21 C.F.R. § 808.1(d). Although the FDA holds a contrary view of the effect of its regulations, these issues need not be addressed to resolve this appeal.").

58 Brief for the United States as Amicus Curiae at 13, Smith Indus. Med. Sys., Inc. v. Kernats, 522 U.S. 1044 (1997) (No. 96-1405), 1997 WL 33561767. According to the Solicitor General, "We have been informed by the FDA that it imposes such specific [federal] requirements on Class III devices only in extraordinary situations." *Id*. at 15.

59 Margaret Jane Porter, *The* Lohr *Decision: FDA Perspective and Position*, 52 Food & Drug L.J. 7, 11 (1997).

60 Brief for the United States as Amicus Curiae at 3, Horn v. Thoratec Corp., 376 F.3d 163 (3d Cir. 2004) (No. 02-4597), 2004 WL 1143720.

61 *Id*. at 16-17.

62 *Horn*, 376 F.3d at 170-71 ("[T]he Supreme Court has instructed us that the FDA's preemption determinations are significant and should inform our interpretation of § 360k(a).").

63 544 U.S. 431 (2005).

64 *Id.* at 450.

65 545 U.S. 967, 981 (2005) ("Agency inconsistency is not a basis for declining to analyze the agency's interpretation under the *Chevon* framework.").

66 *Id*.; *see also* 5 U.S.C. § 706(2)(A).

67 535 U.S. 212, 213 (2002).

68 *See, e.g*., Gonzales v. Oregon, 546 U.S. 243 (2006) ("The weight of such a judgment in a particular case will depend upon the thoroughness evidence in its consideration, the validity of its reasoning, its consistency with earlier and later pronouncements, and all those factors which give it power to persuade, if lacking the power to control.") (citing Skidmore v. Swift & Co., 323 U.S. 134, 140 (1944).

69 *Riegel*, 451 F.3d at 125 (citing *Horn*, 376 F.3d at 179; Motor Vehicle Mfrs. Ass'n v. State Farm Mut. Auto Ins. Co., 463 U.S. 29, 42 (1983)).

70 U.S. *Riegel* Amicus Brief at 17.

71 *Id.* The SG has offered somewhat more in the way of explanation than it did in *Bates*, where it disavowed its previous position, but provided no explanation. *See* Brief for the United States as Amicus Curiae Supporting Respondents at 20, Bates v. Dow Agrosciences LLC, 544 U.S. 431 (2005) (No. 03-388), 2004 WL 1205202 ("The United States has properly considered and disavowed its prior position that Section 136v(b) does not preempt state common-law duties.").

72 No. 06-1498, 2007 WL 1420397 (Sept. 25, 2007).

73 2006 WL 3041078 (Vt. Oct. 27, 2006), *petition for cert. filed,* No. 06-1249 (May 21, 2007) (inviting the Solicitor General to file a brief in this case expressing the views of the United States).

74 The Third Circuit is poised to be the first court of appeals to decide the preemption issue. Colacicco c. Apotex, Inc., 432 F. Supp. 2d 514 (E. D. Pa. 2006), *appeal docketed,* No. 06-3107 (3d. Cir. 2006); McNellis v. Pfizer, 2005 WL 3752269 (D. N.J. Dec. 29, 2005), *motion to vacate denied, interlocutory appeal granted,* 2006 WL 2819046 (D. N.J. Sept. 29, 2006), *appeal docketed,* No. 06-5148 (3d Cir. 2006).

Civil Rights

No Big Surprise: A Review of the Seattle Schools Case

*By Harry J.F. Korrell**

On June 28, 2006, the Supreme Court issued its decision (along with several concurring and dissenting opinions) in *Parents Involved in Community Schools v. Seattle School District No. 1*.[1] This case has received widespread attention, and has been called one of the most significant equal protection cases in decades. This article provides a summary of the factual background and the complex procedural history of the Seattle litigation that led to the decision. It also examines some of the criticism of the Court's decision to review the case and the result. Contrary to the hyperbole from some quarters, a close review of the underlying Ninth Circuit opinion, the developing rift among courts (and among the judges of those courts) about the legality of racial balancing programs, and the holding and reasoning of the *Parents* decision, reveals (1) that review by the Supreme Court was necessary to resolve a split among lower courts on an issue of national importance; and (2) that the decision is a straight-forward application of long-established Equal Protection Clause jurisprudence. The decision to grant certiorari and the outcome are not surprising in light of the Court's earlier Equal Protection Clause cases.

I. Factual Background and Procedural History

Prior to the Supreme Court's decision in *Parents*, the Supreme Court had never decided whether a public school district could make admission decisions based on race, absent a need to remedy de jure segregation; that is, prior discrimination by that district. *Parents* presented an opportunity for the Court to answer this question and to clarify how the equal protection rights of public high school students were affected by the landmark decisions in *Grutter v. Bollinger*[2] and *Gratz v. Bollinger*.[3]

A. Seattle's Race-Based Admissions Plan and Parents' Suit

The racial composition of the students attending Seattle public schools is about 60% non-white and 40% white.[4] It was undisputed in the litigation that Seattle's public high schools were never intentionally racially segregated,[5] though the racial composition of individual schools varied across the district, with two schools enrolling student bodies that were 10% and 8% white, and three enrolling student bodies that were between 55% and 63% white (thus the "whitest" school in the district enrolled a student body that was about 37% non-white).

The schools also varied widely in the quality of education provided, measured by objective and subjective criteria, and in popularity. At the time suit was filed, the district operated ten regular high schools. Five were regarded by most families as providing significantly better educational opportunities than the others. At three of these popular schools, without use of the race preference, enrollment would have been predominantly white (between 55% and 63%) and predominantly non-white (53% and 80%), at two others.

Rather than assigning students to specific schools, as many districts do, the Seattle School District allowed students to select whatever school they desired to attend, and students completed forms ranking their preferred schools. This "open choice" assignment plan allowed families to vote with their feet. Due to the differing quality of the schools, 82% of all students selected one of the five better schools as their first choice, with the result that more students wanted to attend the popular schools than the schools were willing to enroll. In district parlance, these schools were "oversubscribed."

To allocate admissions to these popular schools, and in an effort to achieve a racial balance in these schools that approximated the district's 60% non-white to 40% white ratio, the district employed a series of preferences to determine admission. When a school was oversubscribed, the district first admitted siblings of enrolled students. The district next looked at a school's racial composition and used race to determine who would be admitted. If the ratio of non-white to white pupils in an oversubscribed school deviated by more than a set number of percentage points from the desired 60/40 balance, then a student whose race would have moved the school closer to the desired racial balance would have been admitted, and a student whose race would have moved the school away from the desired balance would have been denied.[6] In effect, seats at such a school were reserved for preferred-race students, and only after all preferred-race students were admitted would others be admitted. There was no individual consideration of applicants, and whenever race was considered it was the sole deciding factor.

In 2000-2001, the trigger for the operation of the race preference was a school's deviation from the preferred 60/40 balance by ten percentage points. That year, the district denied about 300 students admission to their first-choice schools solely because of race. About 210 students were denied their first choice (and many were denied their second and third choice) because they were white; about ninety were denied their first choice because they were non-white.

These race-based assignments imposed significant burdens on affected families, among them (1) denial of admission to a chosen school (in an otherwise open choice system), (2) imposition of cross-town commutes, and (3) the concomitant difficulty of parental involvement in the schools. While these assignments denied hundreds of students admission to chosen schools solely because of skin color, they had only a marginal effect on the racial balance of the schools: the district's data show that without the use of race, all the oversubscribed schools would enroll substantial numbers of white and non-white students. For example, *without using the race preference*, in 2000-01 Roosevelt High School would have enrolled a population that was 54.8% white and 45.2%

** Harry Korrell is a Partner in the Seattle office of the national law firm Davis, Wright, Tremaine, LLP. Along with DWT partner Daniel Ritter, he represented the Parents association in the seven-year litigation, and argued the case to the U.S. Supreme Court.*

non-white. The district's race-based assignments changed the racial balance at Roosevelt by less than four percentage points, increasing the minority enrollment from 45.2% to 48.9%. Similarly, using race changed the white/non-white percentages at other oversubscribed schools by only about two and a half to six percentage points.

The petitioner at the Supreme Court, and the plaintiff below, was an association of families who were either affected or likely to be affected in the future by the district's race-based admissions plan. Parents Involved in Community Schools ("Parents"), formed as a Washington nonprofit corporation. Contrary to the impression left by some of the reports in the popular press, Parents included both white and non-white families. Parents filed suit in federal district court asserting claims under the Washington Civil Rights Act,[7] the Equal Protection Clause of the Fourteenth Amendment, and Title VI of the federal Civil Rights Act of 1964.

After suit was filed, the district modified its admissions plan by changing the trigger for the race preference from a ten point deviation to a fifteen point deviation from the desired racial balance, limiting the use of race to ninth grade assignments (previously the race tie-breaker also applied to new assignments to upper grades), and installing a "thermostat," so that when a school reached the desired balance the use of race as a factor was stopped for that year (previously the preference applied to all assignments in a given year once it was triggered). The district rejected further narrowing proposals advocated by the superintendent of schools, such as changing the trigger to a twenty-point deviation and granting a preference for students who identified a school as a first choice on their rankings.

The district offered several justifications for seeking its preferred racial balance. These included the educational benefits argued to flow from racial diversity, increased racial and cultural understanding, and the desire to avoid racially "isolated" schools, which the district argued would result in the absence of race-based student assignments because of Seattle's housing patterns.

B. The District Court Granted Summary Judgment to the School District

On cross motions for summary judgment, the district court granted judgment in favor of the district.[8] The court found no violation of state law, the Equal Protection Clause, or the federal Civil Rights Act, holding that "achieving racial diversity and mitigating the effects of *de facto* residential segregation... are compelling government interests as a matter of law." Explicitly deferring to the district's judgment, the court concluded that the district had a "sufficient basis" for implementing the race preference, and that the race preference was narrowly tailored to achieve those objectives.

C. The Court of Appeals for the Ninth Circuit Reversed, Holding that the Plan Violated the Equal Protection Clause

On appeal to the Ninth Circuit, a three-judge panel unanimously found for Parents on the state law claim.[9] On Parents' motion (because a new round of assignments was to be made before the court would issue its mandate), the court enjoined the use of the race preference.[10] When the district sought rehearing, the panel withdrew its decision, vacated the injunction, and certified the state law issues to the Washington Supreme Court,[11] which decided those issues in favor of the district.[12]

While the federal claims were still pending in the Ninth Circuit, the U.S. Supreme Court decided *Grutter* and *Gratz*. Parents then rebriefed and reargued their Equal Protection claim in light of those decisions. The panel decided in favor of Parents, holding that the district's plan was not narrowly tailored because it "is virtually indistinguishable from a pure racial quota;"[13] it "fails virtually every one of the narrow tailoring requirements;"[14] and the record revealed "an unadulterated pursuit of racial proportionality that cannot possibly be squared with the demands of the Equal Protection Clause."[15] One judge dissented.[16]

D. A Sharply Divided En Banc Panel Affirmed the District Court, Holding that the Plan Was Constitutional

A rehearing en banc resulted in a decision in favor of the District by a vote of seven (including one concurrence) to four.[17] The en banc majority, relying on the observation in *Grutter* that "context matters," extended the reasoning in that decision in several ways. The majority held racial diversity, pursued for its "educational and social benefits," and to avoid "racially concentrated or isolated schools," can be a compelling governmental interest for high schools.[18] The majority also held that much of the rigorous narrow tailoring analysis of *Grutter* and *Gratz* does not apply in the high school context,[19] so that, inter alia, a mechanical race-based admissions scheme can satisfy the narrow tailoring prong of strict scrutiny when implemented to achieve a pre-determined racial balance. In reaching this conclusion, the majority deferred to the judgment of the local school board regarding the need for a race-based admissions plan.[20] It also adopted a theory of equal protection rights as group rights, holding that a racial classification scheme does not "unduly harm any students," so long as it does not "uniformly benefit any race or group of individuals to the detriment of another."[21]

Judge Kozinski concurred in the judgment.[22] He urged the Supreme Court to abandon strict scrutiny and adopt a "rational basis" standard for evaluating the constitutionality of race-based school assignment plans of the kind at issue.[23]

Judge Bea, joined by three others, dissented.[24] They rejected, as inconsistent with strict scrutiny, the majority's "relaxed," "deferential" standard of review;[25] its deference to the local school board;[26] and its group rights theory of the Equal Protection Clause.[27] The dissent concluded that when strict scrutiny is applied, the district's race preference violated the Equal Protection Clause because it sought to accomplish only a predetermined white/non-white racial balance (not "genuine" diversity);[28] because the plan operates as a quota system;[29] and because it does not satisfy the other narrow tailoring requirements set out in *Grutter* and *Gratz*.[30]

Parents filed a petition for a writ of certiorari, granted on June 5, 2006. As noted above, the Court also agreed to review the *McFarland* case out of the Sixth Circuit, which raised similar issues in the context of elementary school assignments in a school district that had a history of de jure segregation and that had recently achieved unitary status.

II. The Decision to Grant the Petitions

The decision in *Parents* has been the subject of a great deal of discussion and criticism,[31] and that criticism has extended even to the decision to accept review. However, the Supreme Court does not appear to have jumped at the chance to review the Seattle case once Justice O'Connor was replaced by Justice Alito.[32] Such speculation overlooks the fact that after the Ninth Circuit decision in *Parents* there was a significant split among lower courts and judges of individual circuit courts about whether, absent the need to remedy past discrimination, the Equal Protection Clause allowed government schools below the university level to admit or deny students on the basis of race. The lower courts and school officials nationwide needed guidance from the Court.

Prior to the decision in *Parents,* the Supreme Court had not decided whether a school district may use race-based pupil assignments for any purpose other than remediation of the effects of past de jure segregation. Before the 2003 decisions in *Grutter* and *Gratz,* lower courts reviewing racial classifications by government applied the reasoning of Justice Powell's opinion in *Regents of Univ. of Calif. v. Bakke,*[33] and of subsequent Equal Protection decisions such as *Wygant v. Jackson Bd. of Educ.*,[34] *City of Richmond v. J .A. Croson Co.,*[35] *Freeman v. Pitts,*[36] and *Adarand Constructors Inc., v. Pena*.[37] Accordingly, the federal courts of appeal consistently struck down racial balancing schemes by government, including race-based admission and assignment plans of secondary and primary schools.[38]

In those cases, the First and Fourth Circuits, without deciding that diversity can be a compelling interest for secondary and primary schools, held that in the absence of de jure segregation plans designed to achieve a particular racial balance are unconstitutional, citing Justice's Powell's opinion in *Bakke* and subsequent equal protection cases applying strict scrutiny.[39] Similarly, the Ninth Circuit in *Ho* allowed racial quotas only to remove "vestiges of segregation."[40] Except for the Second Circuit in *Brewer*, prior to *Grutter* and *Gratz* there had been no federal court of appeals decision authorizing a plan of racial balancing even to remedy de facto segregation.

In 2003, the Supreme Court addressed equal protection challenges to the race-conscious admissions plans at the University of Michigan's law school,[41] and its undergraduate school.[42] In those cases, the Court explicitly endorsed Justice Powell's *Bakke* opinion and adopted its reasoning.[43] In *Grutter,* the Court affirmed that equal protection rights are "personal" rights, not group rights, and that strict scrutiny applies to all government racial classifications: the government must prove that the racial classification scheme is justified by a compelling interest and is narrowly tailored to achieve that goal.[44] The Court agreed with Justice Powell's *Bakke* opinion, and held that "genuine diversity" (distinguished from mere racial or ethnic diversity) in the student body could be a compelling interest for institutions of higher education.[45] The Court also expressly endorsed Justice Powell's view that an interest in assuring that a student body contained "some specified percentage of a particular group merely because of its race... would amount to outright racial balancing, which is patently unconstitutional."[46]

Grutter also set out the elements of the narrow tailoring prong of strict scrutiny: to pass muster, any race conscious plan must (1) provide for individualized consideration of applicants, (2) not operate as a quota system by imposing a fixed percentage that cannot be exceeded, (3) provide serious, good faith consideration of race-neutral alternatives, (4) not impose undue harm, and (5) have a logical end point.[47] Elaborating on these elements of the analysis, the Court stated that race must "be used in a flexible, nonmechanical way."[48] The plan cannot "make[] an applicant's race or ethnicity the defining feature of his or her application." It must "consider race or ethnicity only as a 'plus' in a particular applicant's file."[49] Applying those factors, the *Grutter* Court held that the law school plan was narrowly tailored, noting inter alia that it was flexible, provided serious individualized consideration to applicants, weighed many other diversity factors besides race, and did not operate mechanically such that race was always a determining factor when it was considered.[50]

In *Gratz*, the Court reiterated its endorsement of Justice Powell's view that "[p]referring members of any one group for no reason other than race or ethnic origin is discrimination for its own sake."[51] Applying the standards articulated in *Grutter*, the Court held that the University of Michigan's undergraduate admissions plan was unconstitutional because it was not narrowly tailored: the plan did not provide for individualized consideration of an applicant's potential contributions to diversity (apart from his or her race), it was mechanical, and race was a decisive factor for virtually every minimally qualified minority applicant.[52]

Despite the Court's express adoption of Justice Powell's *Bakke* rationale, and the *condemnation* of racial balancing in *Grutter* and *Gratz*, the court of appeals in *Parents* (in a sharply divided en banc decision) read *Grutter* and *Gratz* as an invitation to *approve* of racial balancing as a means for government to accomplish mere racial diversity.[53] The court of appeals also rejected most of the rigorous narrow tailoring requirements of *Grutter* and *Gratz* as inapplicable to the "context" of high school assignment plans and held that racial balancing could be a permissible means to accomplish the district's goals.[54]

Soon after the court of appeals' decision in *Parents*, the Sixth Circuit in *McFarland v. Jefferson Cty. Pub. Schs.*[55] affirmed per curium and adopted the opinion of the district court reported at 330 F.Supp.2d 834 (W.D. Ky. 2004). In that case, parents challenged racial guidelines that affected some admissions to some schools and that sought to avoid in any school a black population of less than 15% or greater than 50% in a system whose overall student population was 34% black, and which had operated under a desegregation decree until 2000.[56] Applying *Grutter* and *Gratz*, the lower court found a compelling interest in "maintaining integrated schools,"[57] and determined that the guidelines were narrowly tailored to achieve that objective,[58] except at one group of schools where white and black applicants were put on separate assignment tracks—there the guidelines were held to constitute an "illegal quota."[59]

A few days after the appellate decision in *McFarland,* another Sixth Circuit district court granted a temporary

restraining order prohibiting denial of a student request for hardship transfer where the denial was based solely on race pursuant to a racial balancing plan.[60] Some Sixth Circuit courts thus appeared to be following the pre-*Grutter* line of cases condemning racial balancing (listed in note 31 above).

Likewise in the Fifth Circuit, where *Cavalier v. Caddo Parish School Board* struck down a race-based magnet school admissions plan aimed at achieving a pre-determined racial balance.[61] The court held that the school board had no compelling interest for its use of race because, although the district had operated under a desegregation decree between 1981 and 1990, there was no evidence of either current segregation or vestiges of past segregation,[62] and that under the narrow tailoring analysis of *Grutter* the school board "cannot justify its outright racial balancing absent a showing of current effects of prior segregation, which it has not done."[63]

In summary, while *Grutter* and *Gratz* continued and developed the equal protection doctrine enunciated in *Bakke* and *Croson*, including the prohibition of racial balancing, the Ninth Circuit's decision in *Parents* moved the law in the opposite direction. By taking *Grutter* and *Gratz* as a license to approve racial balancing, the Ninth Circuit in *Parents* (joining the First Circuit in *Comfort*) opened wide the door to race-based school assignments (and by logical extension to other racial classifications claimed to promote racial diversity in other areas of government).

Not only was that decision a deviation from established equal protection jurisprudence (*see* note 31), it confirmed a new uncertainty in the courts about the legality of such plans. For example, counting the initial panel opinion and the en banc opinion in *Parents,* six judges of the Ninth Circuit concluded that Seattle's racial preference violated the Constitution, while eight judges concluded otherwise. The three-judge panel in the First Circuit struck down the race-balancing plan in *Comfort*, but the en banc panel divided three to two and upheld it. Among the circuits, racial balancing to increase diversity in public schools was condoned by post-*Grutter* decisions in the First and the Ninth Circuits and condemned by post-*Grutter* decisions in the Fifth and (apparently) the Sixth Circuits. Pre-*Grutter* decisions that condemn such racial balancing remained precedent in the Fourth Circuit. The situation in the Second Circuit was unclear, as the *Brewer* court held, pre-*Grutter,* only that racial balancing may be used to remedy de facto as well as de jure segregation.[64] If the Supreme Court had not accepted the Seattle and Louisville cases, substantial uncertainty would have plagued school systems and parents nationwide.

III. Did *Parents* "Roll Back" Civil Rights Protections?

The pundits who posit that the Court agreed to review the Ninth Circuit decision in *Parents* so as to roll back advances in civil rights protection also argue that this is exactly what has happened. An examination of the Supreme Court majority opinion reveals, however, a straight-forward application of well-established equal protection jurisprudence. What is most significant about the decision in *Parents* is not any new doctrine, but the fact that the Court once again rejected a call for a less strict scrutiny of government racial classifications.

A. The Supreme Court Majority in Parents Relied on Long-Established Equal Protection Clause Analysis

Justice Roberts wrote an opinion that was for the Court in most respects and for a plurality in others. In the parts of that opinion joined by Justice Kennedy (parts I, II, III-A and III-C), the Court held inter alia as follows:

(1) that strict scrutiny applies to the school districts' racial classification schemes (which denied students admission to chosen schools based solely on race);[65]

(2) that in the education context the Court has only acknowledged two compelling interests: remediation of past de jure segregation and—in the context of higher education—achievement of a broad notion of diversity of which race could be only one of many facets (the plans at issue pursue neither);[66]

(3) that, unlike the genuine or holistic diversity pursued in *Grutter,* racial balancing (except to remedy past de jure segregation) is not a compelling interest and is (still) patently unconstitutional;[67]

(4) that the plans at issue in *Parents* also fail the narrow tailoring prong of strict scrutiny because (a) they do not provide for any individualized review but instead rely on race in a mechanical way;[68] (b) the marginal alleged benefit of these plans did not outweigh the cost of subjecting hundreds of students to disparate treatment based solely on skin color;[69] and (c) the districts did not show that they had earnestly considered race-neutral alternatives to their racial classification schemes.[70]

Justice Kennedy joined these aspects of the Roberts opinion.[71] Writing separately, he made clear his view that the Constitution does not require complete color blindness: for example, in his view, the Constitution does not prohibit school officials' consideration of the effects on the racial composition of schools when making general administrative decisions, such as where to build a school or what magnet programs to fund.[72] But he reserved some of his most impassioned language for his repeated condemnation of government's classifying people by race and making decisions, such as school admissions decisions, based on a person's race.[73]

To anyone familiar with the Court's modern Equal Protection Clause jurisprudence, none of these holdings is remarkable. In *Parents,* the Court simply applied long-standing Equal Protection Clause analysis to yet another government racial classification scheme.

So what, then, is the doctrinal significance of the *Parents* decision? The case is most important not for any new analysis but for the Court's rejection of a new "diversity" jurisprudence being pressed by school districts and others committed to racial balancing efforts (*see, e.g.*, many of the more than fifty amicus briefs filed in support of the school districts in the two cases) and recently adopted by some lower courts: a less exacting, more deferential standard of review for so-called "benign" racial classifications. In the past, a similar analysis has been endorsed by some of the justices on the Court, including the dissenters in *Parents* and *Gratz,* but it has been unable to command a majority.[74]

B. The Ninth Circuit's Decision Was a Radical Departure from Established Equal Protection Jurisprudence

Although the court of appeals majority purported to apply *Grutter* and *Gratz* to the differing "context" of public high schools, as recognized by the en banc dissent and explained below, the Ninth Circuit deviated radically from established equal protection jurisprudence (1) by allowing public schools that were never segregated to engage in racial balancing, (2) by deferring to school officials on racial matters instead of conducting a genuinely strict scrutiny, and (3) by adopting the theory that Equal Protection rights belong to racial groups, not individuals.

1. The Ninth Circuit Found a Compelling Interest in Racial Diversity Defined as a Pre-determined Racial Balance Between White and Non-White Students And Thereby Eliminated Much of the Established Narrow Tailoring Inquiry

Seattle school officials defined racial diversity in a given high school as a ratio of white to non-white students that deviates by no more than a set percentage from its preferred 40/60 ratio of white to non-white students. Accepting this definition, the court of appeals allowed the use of a racial classification—the race preference—by which children were granted or denied admission to their preferred high schools according to a mechanical, arithmetic formula implemented by computer algorithm and based solely on race. The immediate purpose of the preference and its only effect was to move the racial composition of a few schools closer to a pre-determined racial balance. Yet it is just this process of racial balancing—the mechanical use of a racial classification to accomplish a pre-determined racial balance, whether of students, employees, or government contractors—that the Supreme Court had repeatedly condemned, beginning with *Bakke* and continuing through *Grutter* and *Gratz*. Therefore, by accepting diversity as a compelling interest, where diversity means achievement of a pre-determined racial balance, the en banc majority found a compelling governmental interest in doing exactly what the Supreme Court has repeatedly said the Constitution forbids.

Moreover, when racial balancing becomes a permissible government objective, few of the narrow tailoring requirements of strict scrutiny apply in any meaningful way. If racial balance is a permissible goal, there is no need for individualized consideration of applicants or consideration of other ways in which a student could contribute to diversity. The requirement that race not be a determining factor has no place. Likewise, the requirement that the plan not be a quota has no application if a particular racial balance is a permissible goal. Indeed, as the dissent from the en banc opinion pointed out, the court of appeals dispensed with much of the narrow tailoring inquiry as inapplicable in the high school "context."[75]

2. Instead of Applying "Strict Scrutiny," the Ninth Circuit Held that Courts should Defer to Local School Officials on Matters of Race

In *Grutter*, the Court accorded deference to the judgment of officials at the University of Michigan Law School that genuine diversity in the classroom was essential to its educational mission and thus was a compelling interest.[76] This unusual deference was justified because of a university's academic freedom and "a constitutional dimension, grounded in the First Amendment, of educational autonomy," that includes the freedom of a university to select its student body.[77] The Court expressly limited this deference (to the school's determination of what interests are compelling) to the university context and the "expansive freedoms of speech and thought associated with the university environment... a special niche in our constitutional tradition."[78]

In *Parents*, the court of appeals improperly deferred to local school officials. For example, rather than scrutinize strictly whether a race-based plan was necessary and whether the district seriously considered and properly rejected race neutral alternatives to its race-based scheme, the en banc majority cited the deference accorded university officials in *Grutter* and deferred to the district's judgment that race-neutral alternatives would have been inadequate to obtain the benefits of a racially diverse student body.[79]

In *Parents*, the testimony of the Superintendent of Schools on these issues was plain. When asked whether the district "g[a]ve any serious consideration to the adoption of a plan... that did not use racial balancing as a factor or goal," he testified:

> *I think the general answer to that question is no...* I don't remember a significant body of work being done. I mean it's possible informally ideas were floated here or there, but I don't remember any significant staff work being done.[80]

Testimony by other school officials confirmed this admission.[81]

In light of this testimony, the court of appeals could find an earnest consideration by the district of race-neutral alternatives—an essential requirement of narrow tailoring—only by adopting a less exacting "rational basis" standard, which is exactly what the court did: according to the majority, the plan satisfied these aspects of narrow tailoring because "the record reflects that the district *reasonably concluded* that a race-neutral alternative would not meet its goals."[82]

The en banc majority's deference to the government and the resulting rational basis scrutiny was inconsistent with the rigorous analysis required by the Constitution.[83]

3. The Ninth Circuit Treated Equal Protection Rights as Belonging to Racial Groups Rather than to Individual Students

The race preference operated in some instances to deny white children their preferred assignments solely because they were white. In other instances it denied non-white children their preferred assignments solely because they were not white. Because equal protection rights are individual, personal rights,[84] each student affected by the operation of the preference suffered injury under traditional equal protection analysis: an infringement of her personal right to be free from race-based decision-making by government, and the denial of an otherwise generally available benefit (the opportunity to choose her high school) solely because of her race.

The en banc majority, however, abandoned this bedrock principle of constitutional law and treated equal protection rights as group rights. The majority held the district's plan

was narrowly tailored in part because the tie-breaker "does not uniformly benefit any race or group of individuals to the detriment of another," and thus does not "unduly harm any students in the District."[85]

This group rights analysis was contrary to the established understanding of the right to equal protection as a *personal* right.[86]

C. The Supreme Court Merely Rejected the Invitation to Change its Equal Protection Jurisprudence

Prior to *Parents*, the Court had never expressly held that racial balancing plans by local school boards were unconstitutional. It seems that many (relying on dicta in cases decided decades before the Court expressly adopted strict scrutiny) hoped the Court would recognize an exception to strict scrutiny for government "diversity" initiatives and other "benign" racial classifications. The Seattle School District, many of the amicus briefs filed in support of it, and the en banc majority at the Ninth Circuit advocated a less exacting standard by arguing that the court should *defer* to the judgment of school authorities and that racial classifications should be constitutional if they were *reasonable* efforts by government to address an honestly-perceived problem of racial imbalance in some schools. They sought implicitly what Judge Kozinski advocated expressly: adoption of a rational basis standard of review for racial classifications that appear to be benign and that do not uniformly work to the detriment of any particular racial group.[87]

The Supreme Court rejected this call for a new Equal Protection Clause jurisprudence. As noted above, the reasoning in the Court's opinion was straight-forward: the Court held (again) that racial balancing, except to remedy past discrimination, is unconstitutional, and that the appropriate standard of review for all racial classifications is strict scrutiny, even if the program under review is defended as a "diversity" measure. The Court rejected a plea for unprecedented deference to local school boards on racial matters, and it rejected the Ninth Circuit's group rights theory of the Equal Protection Clause. It also reiterated that, even when pursuing a compelling interest, school districts must prove that any racial classification scheme is really necessary and that school officials seriously considered and rejected alternatives to using race—something that was impossible for the Seattle School District to do in light of the testimony of school officials. None of this is remarkable or new, but it is very significant that the Court reiterated, in yet another context, that it will not retreat from its application of strict scrutiny to all governmental racial classifications.

CONCLUSION

The Supreme Court's decision in *Parents* was not surprising, though not because, as some complain, President Bush appointed Justices Roberts and Alito to the Court. The result was unsurprising because the admissions plans employed by the school districts were, in the words of the Seattle School District's superintendent, "blunt" instruments. They were employed to accomplish only the crudest kind of "diversity": a pre-determined white/non-white ratio (in Seattle) and a black/"other" ratio (in Louisville). The school districts did not seriously consider any alternatives to their racial classifications because they were committed to accomplishing racial balance, not the kind of "genuine" or "holistic" diversity sought by law schools such as the University of Michigan in *Grutter*.

It was risky for the school districts and their allies to rely on decades-old dicta and dissenting opinions. Nothing in the Court's modern equal protection decisions suggested that these plans could have survived strict scrutiny as that test has been repeatedly articulated and applied in numerous other contexts. What is significant about the decision is not so much the doctrine announced—the analysis was straight forward and predictable—but the rejection yet again of the call for a relaxed scrutiny of supposedly benign racial classification schemes. This should give pause to those who might be inclined to look for ways to work around the Constitution's prohibitions, even for what they are convinced are good reasons.

Endnotes

1 ___ U.S. ___ 127 S. Ct. 2378 (2007) ("*Parents*").

2 539 U.S. 306 (2003).

3 539 U.S. 244 (2003). By accepting review of the Sixth Circuit's per curiam decision in *McFarland v. Jefferson Cty. Bd. Of Educ.*, 416 F.2d 513, the Court was also able to address this issue for lower schools as well.

4 The city population is about 70% white and 30% non-white; the school age population—including children who attend private schools—is about 50% white and 50% non-white.

5 In the Sixth Circuit case under review at the same time, the school district had been segregated de jure, but it had recently achieved unitary status.

6 A student was deemed to be of the race specified in her registration materials (and if a parent declined to identify a child's race, the district assigned a race to the child based on a visual inspection of the student or parent).

7 Wash. Rev. Code § 49.60.400 (1999).

8 *Parents,* 137 F. Supp. 2d 1224 (2001).

9 *Parents,* 285 F.3d 1236 (2002)

10 *Parents,* 2002 WL 841345 (9th Cir. Apr. 26, 2002).

11 *Parents,* 294 F.3d 1085 (2002).

12 *See Parents,* 149 Wash. 2d 660, 72 P.3d 151 (2003).

13 *Parents,* 337 F.3d 949, 969 (2004).

14 *Id.*

15 *Id.* at 976.

16 *Id. a*t 989.

17 *Parents,* 426 F.3d 1162 (2005).

18 *See, e.g., id.* at 1179.

19 *See, e.g., id.* at 1184, 1186.

20 *See, e.g., id.* at 1188, 1190-91.

21 *Id.* at 1191-92.

22 *Id.* at 1193.

23 *Id.* at 1195.

24 *Id.* at 1196.

25 *Id.* at 1197, 1199.

26 *Id.* at 1207–09, 1214.

27 *Id.* at 1216-17.

28 *See, e.g., id.* at 1202-03, 1209, 1221.

29 *Id.* at 1212-14.

30 *Id.* at 1210, 1214-16, 1218-20.

31 That criticism seems not to be joined by most of the American public. *See* Aug. 16, 2007, Quinnipiac Univ. Poll, *available at* www.quinnipiac.edu/x1295.xml? Release ID=1093 ("By a 71-24 percent margin, American voters agree with a recent U.S. Supreme Court decision that public schools may not consider an individual's race when deciding which students are assigned to specific schools, according to a Quinnipiac University National Poll released today.").

32 Those making this argument point out that, shortly before Justice O'Connor retired, the Court denied review of a First Circuit case raising similar issues, *Comfort v. Lynn Sch. Comm.*, 418 F.3d 1 (1st Cir. 2005), *cert. denied,* 126 S. Ct. 798 (2005). *See, e.g.,* Comments of Prof. Pamela Karlan at Am. Const. Law Society's Annual S. Ct. Review, 6/28/2004, *available at* www.acslaw.org/node/5088 (calling the decision to review these cases the "clearest example of an agenda—driven supreme court" after Justice O'Connor's retirement.) But as explained below, until the Ninth Circuit's en banc decision, *Comfort* was an outlier, reaching a decision contrary to the overwhelming weight of circuit court authority (including the decision of the Ninth Circuit's three-judge panel that struck down the district's plan).

33 438 U.S. 265 (1978).

34 476 U.S. 267 (1986).

35 488 U.S. 469 (1989).

36 503 U.S. 467 (1992).

37 515 U.S. 200 (1995).

38 *See, e.g.,* Johnson v. Board of Regents of Univ. of Ga., 263 F.3d 1234 (11th Cir. 2001) (university admissions); Smith v. Univ.of Wash., 233 F.3d 1188 (9th Cir. 2000) (applying *Bakke* to law school admissions); Eisenberg v. Montgomery Cty. Pub. Schs, 197 F.3d 123 (4th Cir. 1999) (transfers to magnet school); Tuttle v. Arlington Cty. Sch. Bd., 195 F.3d 698 (4th Cir. 1999) (admissions to oversubscribed school); Wessmann v. Gittens, 160 F.3d 790 (1st Cir. 1998) (admission to Boston Latin School); Ho v. San Fran. Unified Sch. Dist., 147 F.3d 854 (9th Cir. 1998) (racial quotas for schools); Lutheran Church-Missouri Synod v. FCC, 141 F.3d 344 (D.C. Cir. 1998) (radio station hiring); Monterey Mech. Co. v. Wilson, 125 F.3d 702 (9th Cir. 1997) (public contracting); Podberesky v. Kirwan, 38 F.3d 147 (4th Cir. 1994) (college scholarships); *In re Birmingham Reverse Discrimination Employment Litig.*, 20 F.3d 1525 (11th Cir. 1994) (hiring quotas); Cunico v. Pueblo Sch. Dist. No. 60, 917 F.2d 431 (10th Cir. 1990) (employment). Brewer v. West Irondequoit Cent. Sch. Dist., 212 F.3d 738 (2d Cir. 2000), *superseded on other grounds as stated in Zervos v. Verizon N.Y., Inc.,* 252 F.3d 163, 171 n.7 (2d Cir. 2001), held that racial balancing may be used to remedy *de facto* as well as de jure segregation of schools and remanded for trial. *Cf.* Hunter v. Regents of Univ. of Calif., 190 F.3d 1061 (9th Cir. 1999) (allowing racial balancing for research purposes in university laboratory school).

39 *See Wessman; Eisenberg; Tuttle, id.*

40 147 F.3d at 865.

41 *Grutter,* 539 U.S. 306.

42 *Gratz,* 539 U.S. 244.

43 *See, e.g., Grutter*, 539 U.S. at 325; *Gratz,* 539 U.S. at 270-71.

44 539 U.S. at 326-27.

45 *Id.* at 328-30.

46 *Id.* at 329-30 (internal quotation marks and citations omitted).

47 *Id.* at 334-42.

48 *Id.* at 334.

49 *Id.*

50 *Id.* at 336-38.

51 539 U.S. at 270.

52 *Id.* at 271-76.

53 *Parents,* 426 F.3d at 1174-79.

54 *Id.* at 1180-93.

55 416 F.3d 513 (6th Cir. 2005).

56 *Id.* at 840-42.

57 *Id.* at 849-55.

58 *Id.* at 855-62.

59 *Id.* at 862-4.

60 Tharp v. Board of Educ. of N.W. Local Sch. Dist.*,* 2005 WL 2086022 (S.D. Ohio 2005) (citing *Grutter* and the Fourth Circuit's decision in *Eisenberg*, 197 F.3d 123).

61 403 F.3d 246 (5th Cir. 2005).

62 *Id.* at 285-60.

63 *Id.* at 260-61.

64 212 F.3d 738.

65 *Parents,* 127 S. Ct. at 2751-52.

66 *Id.* at 2752-53

67 *Id.* at 2753-54.

68 *Id.*

69 *Id.* at 2759-60.

70 *Id.* at 2760.

71 *Id.* at 2788.

72 *Id.* at 2792-93. Of course, the sorts of general administrative decisions Justice Kennedy was concerned about appear not to involve racial classifications of the kind at issue in the litigation, and no such decisions were before the Court in *Parents*. Any comment by the Court on the legality of such facially race-neutral (but race-conscious) decisions would have been dicta, as *Parents* involved explicit racial classifications by which individual students were admitted or denied admission to a preferred school solely because of skin color, not decisions about where to place a school or what programs to fund.

73 *See, e.g., id.* at 2796-98.

74 *See, Parents,* 127 S. Ct. at 2764 (collecting cases).

75 *Parents,* 426 F.3d at 1216, n. 25.

76 *Grutter,* 539 U.S. at 328-29.

77 *Id.* at 329.

78 *Id.*

79 *See Parents,* 426 F.3d at 1188 ("Implicit in the [*Grutter*] Court's analysis was a measure of deference..."); *id.* at n.33 ("The Supreme Court repeatedly has shown deference to school officials at the intersection between constitutional protections and educational policy.")

80 *See id.* at 1214 (emphasis added).

81 *Id.; Parents,* 377 F.3d at 970 n.23.

82 *Parents,* 426 F.3d at 1188 (emphasis added).

83 *See* Johnson v. California, 543 U.S. 499, 125 S. Ct. 1141, 1146 n.1 (2005) ("deference [by the courts in applying strict scrutiny] is fundamentally at odds with our equal protection jurisprudence"); *id.* at 1150 (the Supreme Court "has refused to defer to state officials' judgments on race... where those officials traditionally exercise substantial discretion"); Goss v. Lopez, 419 U.S. 565, 574 (1975) (the "Fourteenth Amendment... protects citizens against the State itself and all of its creatures—Boards of Education not excepted").

84 *See, e.g., Grutter,* 539 U.S. at 326 ("the Fourteenth Amendment 'protect[s] *persons,* not *groups*'") (emphasis in original, quoting *Adarand*, 515 U.S. at 227).

85 *Parents,* 426 F.3d at 1192.

86 *See, e.g., Grutter,* 539 U.S. at 326; Miller v. Johnson, 515 U.S. 900, 911 (1995) ("At the heart of the Constitution's guarantee of equal protection lies the simple command that the Government must treat citizens as individuals, not as simply components of a racial, religious, sexual, or national class." (internal quotation marks omitted)); *Adarand*, 515 U.S. at 230 ("any individual suffers an injury when he or she is disadvantaged by the government because of his or her race, whatever that race may be."); Loving v. Virginia, 388 U.S. 1 (1967) ("the fact of equal application [of a miscegenation statute] does not immunize the statute from the very heavy burden of justification which the Fourteenth Amendment has traditionally required of state statutes drawn according to race.")

87 *See* 426 F.3d at 1195 (Kozinzki, J., concurring).

Can the Federal Transportation Disadvantaged Business Enterprise Program Be Narrowly Tailored to Remedy Discrimination?

*By George R. La Noue**

The federal transportation Disadvantaged Business Enterprise (DBE) program is a relic from another era that no longer serves a civil rights purpose. DBE is the progeny of the 10% minority set-aside provision in the 1977 Public Works Employment Act (PWEA).[1] That Congress was responding to a temporary economic downturn and the 1968 Kerner Commission, which highlighted the special plight of African-Americans. Creating a share for them and other minority communities seemed like an important part of overall economic pump-priming.

In 1982, Congress adopted that same 10% set-aside for transportation spending and included the same minority groups as beneficiaries, while adding women. The legislative history of both Acts makes references to remedying various forms discrimination, and courts have used those references to justify the Acts' racial provisions. But both these set-asides were also clearly redistributive, making certain that firms owned by members of various groups (particularly African-Americans) received a share of federal contracting dollars.

Since 1989, however, distributing shares of public contracts by race has been regarded by the Supreme Court as a violation of the individual right to equal protection.[2] The use of race, if permissible at all, is limited to the "extreme case where some form of narrowly tailored racial preference is necessary to break down problems of deliberate exclusion."[3] In short, any use of race must be part of a civil rights remedy. Political patronage, or even economic development, is not legitimate reasons for using racial classifications.[4]

The modern DBE program, however, while no longer employing set-asides, does not serve a remedial purpose. It does not identify or sanction the state and local recipients of those funds that discriminate; indeed, it does not affect their award of prime contracts at all. It does not identify contracts where discrimination occurred. It does not identify firms subjected to discrimination and provide remedies to them. It does not identify prime contractors that discriminate in the selection of subcontractors and sanction them. What it does do is redistribute subcontracting dollars to firms owned at least 51% by women or minorities. Examining a small, federally funded paving contract, the Ninth Circuit found that the "prime contractor did not select Western States [a non -DBE], even though its bid was $100,000 less than that of the minority firm that was selected. The prime contractor explicitly identified the contract's minority [DBE] utilization requirement as the reason that it rejected Western States' bid."[5]

Deciding the case on the facts before it, the Ninth Circuit had only a micro view of the racial and gender redistribution involved in DBE programs. No macro view of the extent of the problem existed. A new study of the results of the DBE programs in 413 airports, however, has found that DBE subcontractors were massively over-utilized in every one of the seven regions across the county.[6] In the two regions covered by the Ninth Circuit, DBE subcontractors received 468% of their expected number of contracts and 328% of their expected amount of dollars in the Western Pacific region and 559% of their expected contracts and 580% of their expected dollars in the Northwest region.[7] If this is justice at all, it is very rough justice. So the question this article addresses is: Could the DBE program be modified so that it prevented, sanctioned, and remedied discrimination instead of causing it?

I. The Application of the Compelling Interest Prong

In 1989, the Supreme Court in *City of Richmond v. Croson* applied the strict scrutiny test to state and local contracting programs that provided racial and ethnic preferences.[8] Recognizing that Richmond's program may have been politically inspired,[9] the Court insisted that the first task of a government wishing to create race conscious contracting programs was to identify the discrimination it sought to remedy. Justice O'Connor declared:

> Proper findings are necessary in this regard to define the scope of the injury and the extent of the remedy necessary to cure its effects. Such findings also serve to assure all citizens that the deviation from the norm of equal treatment of all racial and ethnic groups is a temporary matter, a measure taken in the service of the goal of equality itself.[10]

Justice O'Connor went on to criticize the various forms of evidence Richmond offered in defense of its program. Evidence of societal discrimination was not sufficient.[11] Nor could a city depend on generalized assertions of discrimination in entire industry.[12] Finding discrimination in one market did not permit an assumption that discrimination exists in all markets.[13] Nor did a finding of discrimination against one ethnic group lead to a conclusion that discrimination affects other groups.[14] O'Connor did suggest that a properly done disparity study might lead to an inference of discrimination.

> Where there is a *significant* statistical disparity between the number of *qualified* minority contractors *willing and able to perform a particular service* and the number of such contractors actually engaged by the locality or the locality's prime contractors, an inference of discriminatory exclusion *could arise*. (Emphasis added.)[15]

The *Croson* case proved to be a formidable hurdle for local preferential programs. In the first three cases invalidating preferential programs in Philadelphia, Columbus, and Miami-Dade County courts found that the local governments lacked sufficient evidence that they were remedying discrimination, and therefore lacked a compelling interest.[16] Later, other governments also failed to establish a compelling interest to support their preferential contracting programs, though courts often considered narrowly tailoring issues as well.

After the Supreme Court's 1995 decision in *Adarand Constructors, Inc. v. Pena* establishing that strict scrutiny was

* *George R. La Noue is a Professor of Political Science and Public Policy at the University of Maryland, where he directs the Project on Civil Rights and Public Contracts. He wrote on this subject in the March 2006 issue of Engage.*

also the proper standard of review for federal contracting programs,[17] the Clinton administration adopted a "mend, don't end" strategy toward affirmative action. Three new studies were created to show the DBE program was still necessary[18] and the program was amended to make it more narrowly tailored. The 10% national DBE goal was replaced by a requirement that each recipient of federal funds set its own goal, after determining the availability of DBEs to do its projected work. That goal was also to reflect a recipient's determination of the level of DBE participation expected "absent the effects of discrimination."[19]

From the outset of the program, DBEs had to be socially and economically disadvantaged. The definition of social disadvantage successfully kept firms owned by white males out of the program and was essentially left unchanged in the amended program.[20] After Republicans complained that the fabulously wealthy Sultan of Brunei could buy a business and become a DBE, a new definition capping the net worth of firm owners and the total revenues of firms was added to create some semblance of economic disadvantage, though most Americans would consider the caps a very generous definition of disadvantage.

Region	Capacity	Utilization	Disparity %
E. North Central	10.5	16.5	154
E. South Central	11.8	11.5	97
Middle Atlantic	9.1	11.0	121
Mountain	13.3	27.0	203
New England	9.6	26.0	271
Pacific	14.3	16.9	112
South Atlantic	7.0	16.1	230
W. North Central	8.0	9.8	123
W. South Central	13.3	11.1	83

Despite *Adarand*'s strict scrutiny admonition and that Court's willingness to closely scrutinize Congressional findings in other areas,[21] federal courts reviewing challenges to the DBE program have proved to be quite deferential to whatever evidence the Justice Department has told them was "before" Congress.[22] There has been no judicial requirement that Congress actually adopt any set of findings in committee reports or legislative preambles.[23] That action might have created the virtue of some debate about the merits of the asserted evidence of discrimination. Instead, it has been sufficient for judicial reviewers that individual Congressional DBE program advocates have alluded to anecdotes or studies in floor speeches or in hearings that support existence of various inequalities that may be caused by discrimination.[24] No further Congressional identification of discrimination related to federal transportation expenditures has been thought necessary.

Indeed, when the Clinton administration sought to do a *Croson*-like disparity study, the results did not support a national finding of discrimination in federal transportation construction procurement. In 1998, guided by the Justice Department, the Commerce Department conducted the so-called benchmark limits study.[25] The theory behind this study was that *Adarand* required contracting preferences to be used only in industries where evidence of discrimination in federal contracting existed. Federal procurement, therefore, was classified by the Census Bureau's two digit Standard Industrial Codes (SIC) and the dollar shares Small Disadvantaged Businesses (SDB)[26] received in each of 68 industries was compared to estimates of SDB availability and capacity.

The Justice Department also decided that, while all other industries could be analyzed on a national basis, the three construction industry specialties were uniquely local, and so the benchmark limits study divided them into nine geographical regions. As the table below shows, the patterns of availability and utilization differed in each region, but overall SDBS were overutilized, not underutilized. [27]

There were a number of methodological problems in this study that inflate the availability of minority firms,[28] but at least it was an objective effort to narrow tailor the use of racial preferences in federal procurement prime contracting. Further, it shows that a similar study could have been done on the use of federal transportation funds by state and local recipients, but such was rarely attempted. The benchmark study was released by the Clinton administration two weeks after the 1998 reenactment of the transportation DBE program, and never affected that consideration.

As a part of reenactment, however, Congress asked the General Accountability Office (GAO) to investigate the operation and results of the DBE program. Their 2001 report came back titled, "Disadvantaged Business Enterprises: Critical Information Is Needed to Understand Program Impact."[29], GAO concluded that the majority of recipients surveyed had made no analysis of barriers to DBE participation and of those recipients who had done so there was "little agreement among the officials we contacted on whether [barriers] were attributable to discrimination."[30] Further, it found that the fourteen transportation disparity studies they examined were all methodologically deficient.[31]

Nevertheless, every court reviewing the issue has accepted without reservation assertions by the Clinton and Bush Justice Departments that Congress had an abundance of evidence available that created a compelling interest for the DBE program.[32] When Republicans controlled Congress they sought to deal with this problem by not holding hearings on the subject in the hope that the Clinton administration evidence would become obsolete.[33] But now that the political pendulum has swung, it is possible that a Democratic Congress will create a

new predicate supporting the DBE program. Since the federal courts do not seem to require any particular statistical evidence of discrimination which might have to meet some objective standards, the past record suggests little likelihood, short of a new Supreme Court ruling, that the DBE program will be affected by the compelling interest prong.

II. Application of the Narrow Tailoring Prong

The current version of the federal transportation program (Safetea-Lu) provides for the grants of $ 244 billion federal dollars over a five year period to local recipients; mainly to build highways, expand airports, and maintain mass transit systems. Despite the limitations of the compelling interest prong of strict scrutiny as currently interpreted, there is more possibility that the narrow tailoring prong might serve to tie the DBE program to remedying actual discrimination in the expenditures of this massive program. The major issues are discussed below.

A. Market Area Tailoring

According to *Croson*, "Congress explicitly recognized that the scope of the [discrimination] problem would vary from market area to market area,"[34] but lower courts did not require any evidence of local discrimination in the administration of the DBE program. Then in 2005, the Bush Justice Department told the Ninth Circuit in *Western States Paving Co. v. Washington State Department of Transportation* that local findings of discrimination were necessary before local recipients could make the decision to use race-conscious rather than race-neutral measures to meet DBE goals. The circuit panel responded, "As the United States correctly observed in its brief and in oral argument, it can not be said that TEA-21 is a narrowly tailored remedial measure unless its application is limited to those States in which the effects of discrimination are actually present."[35] Otherwise, the court noted:

> Whether Washington's DBE program is narrowly tailored to further Congress's remedial objective depends upon the presence or absence of discrimination in the State's transportation industry. If no such discrimination is present in Washington, then the State's DBE program does not serve a remedial purpose; it instead provides an unconstitutional windfall to minority contractors based solely on race.[36]

The consequence of this decision is that every Ninth Circuit state (Alaska, Arizona, California, Hawaii, Idaho, Montana, Nevada, and Oregon) has or will produce a new disparity study.[37] Furthermore, some local airports and transit districts in that region are doing their own studies.

So, the question is raised: What is a market area? It does not seem larger than the boundaries of a state, but what about federal funding recipients that operate only within a local area in a state? Utilization data will be different for each of these local recipients and availability may be too. It is unlikely that many small DBEs would be equally willing to work in San Diego and Humboldt Counties, which are about 750 miles apart. Moreover, airports and transit districts may attract very different competitors for their large contracts.

In the two post-*Western States* highway transportation studies completed thus far, the Nevada study divided its analysis into three regions, while the California analyzed twelve districts.[38] Not surprisingly, the regional results produced a patchwork of under and over-utilization of DBEs. Will courts permit the imposition of DBE race preferences in local market areas where there is no evidence of discrimination? The only federal court to have considered that question so far, albeit about a state highway program, criticized a statewide conclusion about discrimination, when the relevant disparity study failed to analyze data by specific districts which varied substantially in their demographic composition.[39]

B. Group Specific Tailoring

Firms wishing to be DBEs need formal certification by local recipients. Federal regulations require that a DBE has to 51% owned by an economically and socially disadvantaged person. Economic disadvantage is defined objectively. If the person has a net worth of less than $750,000 (excluding the value of the principal residence and the business) and a construction business (depending on specialty) has revenues of less than $6.5 million to $31 million, the economic test is met. Persons are presumptively socially disadvantaged if they are women or self-identified members of a minority group. If challenged, a person might have to show he or she was accepted by the minority community as a member, but in this era of mixed marriages who has the authority to say that person is or is not a member of a particular minority community is a bit fuzzy.

In modern, multi-cultural America, what is a minority group? The short answer is a group thought to be composed of persons of color and long answer is codified in the following list:

> "Black Americans," which includes persons having origins in any of the Black racial groups of Africa); "Hispanic Americans," which includes person of Mexican, Puerto Rican, Cuban, Dominican, Central or South American, or other Spanish or Portuguese culture or origin, regardless of race); "Native Americans," which includes persons who are American Indians, Eskimo, Aleut or Native Hawaiian); "Asian- Pacific American," which includes persons are from Japan, China, Taiwan, Korea, Burma, (Myanmar), Vietnam, Laos, Cambodia (Kampuchea), Thailand, Malaysia, Indonesia, The Philippines, Brunei, Samoa, Guam, U.S. Trust Territory of the Pacific Islands (Republic of Palau), Federated States of Micronesia, the Commonwealth of the Northern Mariana Islands, Macao, Fiji, Tonga, Kiribati, Juvalu, [sic] Nauru, Federated States of Micronesia, or Hong Kong; "Subcontinent Asian-Americans," which includes persons whose origins are India, Pakistan, Bangladesh, Bhutan, the Maldives Islands, Nepal or Sri Lanka.[40]

In *Croson*, the Supreme Court established the principle that separate findings of discrimination had to be made for each minority group eligible for preferences. Richmond's uncritical adoption of an earlier version of the above list drew Justice O'Connor's retort that:

> The random inclusion of racial groups, that, as a practical matter, may never have suffered from discrimination in the construction industry in Richmond suggests that perhaps the city's purpose was not in fact to remedy past discrimination.... The gross

> overinclusiveness of Richmond's racial preferences strongly impugns the city's claim of remedial motivation. [41]

The Court noted that Richmond's adoption of the federal group categories created a situation in which:

> There is *absolutely no evidence* of past discrimination against Spanish-speaking, Orientals, Indians, Eskimos, or Aleut persons in any aspect of the Richmond construction industry.... It may well be that Richmond has never had an Aleut or Eskimo citizen. (emphasis in the original)[42]

Since that time, almost every state and local disparity study has examined evidence of discrimination for each separate major racial and ethnic group. When they have failed to do so, courts have been harsh in their criticism. [43]

Nevertheless, the Department of Transportation (DOT) has contended that DBE is a category that does not require separate analysis of the groups composing it.[44] DBE annual and contract goals are set for DBEs as a category and not for individual groups. Thus, a prime contractor could consistently discriminate against all subcontractors that were not Hispanic, for instance, but still meet the DBE goals and be free from sanction or even scrutiny. DOT does require recipients to report annually their dollars and contract awards for six major groups, but the federal agency does not appear to do anything with that data.

DOT's DBE single category policy, however, was challenged in *Western States,* and the Ninth Circuit ruled that

> each of the principal minority groups benefited by Washington's DBE program—Black Americans, Hispanic Americans, Native Americans, Asian-Pacific Americans, Subcontinent Asian Americans and women—must have suffered discrimination in the State. If that is not the case, then the DBE program provides minorities who have not encountered discriminatory barriers with an unconstitutional competitive advantage at the expense of both non-minorities and any minority groups that have actually been targeted for discrimination.[45]

The consequence of the group specific narrowly tailoring requirement may be as dramatic for DBE programs as it has been for local MWBE programs. If the data show that there has been no under-utilization of one of the large group components of the DBE category and no other evidence of discrimination against that group in a particular marketplace exists, firms owned by that group would no longer be eligible for preferences and would be treated like non-DBEs in making subcontract choices. That result would create a considerable administrative and political shift. First, it should reduce the size of race and gender conscious contract goals, since there would be fewer firms eligible to fill them. Second, exclusion of some major groups from being beneficiaries of preferential programs may reduce the political support for such policies. For example, legislative members of an African-American caucus or trade association may be less inclined to support contracting preferences, if only Hispanics or white women are the beneficiaries.

While it is too soon to measure the political consequences of the *Western States* group specific rule, the results of the first two post-*Western States* disparity studies and the state DOT responses are now public. Both the Nevada and California DBE goals submission proposals illustrate the enormous impact of the requirement of treating each group separately in the analysis of who is eligible for preferences. The Nevada study found mwbes overall were under-utilized at a 50% ratio, but women-owned firms were over-utilized and African-American firms were in the parity range at 85%.[46] Since after *Western States* firms owned by members of those groups could not be used to fulfill race or gender conscious goals and still have a narrowly tailored program, Nevada will seek to meet its 5.7% DBE goal solely through race-neutral means. In California, the overall mwbe utilization on race-neutral state contracts (the data used by Caltrans to set its proposed 2008 federal DBE goal) was 79.5 %, but Hispanics were at parity and Subcontinent Asians were over-utilized. Caltrans has proposed a 13.5% DBE goal split evenly between race-conscious and race-neutral means. Race-conscious contract goals will exclude firms owned by Hispanics and Subcontinent Asians.[47]

C. Industry Specific Tailoring

Croson bars the use of generalized assertions about discrimination in an entire industry as a basis for race-conscious remedies.[48] Within an industry there may be important specialties with different proportions of DBEs and non-DBEs. Discrimination against DBEs in one construction specialty might have no effect on DBEs in another. Consequently some disparity studies, including the federal benchmark study divide construction into its three major components (building, heavy and highway, and special trades).[49] When that has occurred, the disparities often vary by specialty. For example, according to the benchmark study in the mid-Atlantic region, minority building contractors were underutilized, while minority highway contractors and specialty contractors were over-utilized.

Transportation-related contracts can be even more complex, involving many specialties that have separate professional identities and are not viewed by the Census as part of the "construction industries;" for example, architects, engineers, aerial photographers, and truckers. A discriminatory problem affecting truckers may have no impact on architects. Consequently, the post-*Western States* Nevada and California disparity studies each made about 120 separate disparity analysis. The outcomes were a patchwork of under- and over-utilization.

When a local government conducts a disparity study and finds no statistical support for a finding of discrimination against firms in some industries where it makes purchases, narrow tailoring requires that it no longer employ race-conscious purchasing in those categories.[50] But very few DOT recipients have previously conducted disparity studies, so DBE goals were applied regardless of whether there was any evidence of local discrimination in the various industries involved in their contract. These goals were based solely on DBE availability and neither identified nor remedied any particular discrimination.

D. Goal Setting Tailoring

The key to any recipient's administration of its DBE program is the annual exercise it must go through which results in a goals submission proposal that DOT reviews.[51] DOT

permits a wide variation in the data sources and methods used in this exercise. Sometimes simple headcounts of minority and women owned firms (mwbes) and those owned by white males and stockholders are used for goal setting. When that occurs, it almost always inflates the availability of the newer smaller mwbes which do not have the same capacity as other firms. Other goal setting efforts are sophisticated measures of availability based on a weighted mix of contracts the recipient will award and the variety of specialties that will be required. The difference in approaches can be seen in the post-*Western States* disparity studies. In Nevada, the percentage of mwbes compared to non-mwbes based on their headcount was 24.4%. But when weighted by size, specialty, and location, mwbes availability was reduced to 14.6%, and when the mwbes too large to be certified DBEs were eliminated the mwbes availability figure dropped to 5.7%. It was this percentage that Nevada adopted for its proposed FY 2008 goals [52]

But Nevada's availability estimate is still too large. In a DBE program, only hiring certified DBE firms can be counted by a prime contractor in fulfilling contract goals. But Nevada and many other recipients set goals based on the total of minority- and women-owned firms, not just on the fraction of certified DBEs. In Nevada three of four mwbe firms were not certified.[53] Counting uncertified firms for goal-setting, but not for utilization on specific contracts, can create a windfall for the fraction of firms that are DBE-certified.

Inflated DBE annual goals mean that race-neutral measures will almost never fulfill them, increase the number of contracts to which race-conscious measures must be applied, and accelerate pressure on non-DBE primes to over-select DBE subcontractors. Nevertheless, DOT apparently approves almost all goals submissions, unless it suspects the goals are set too low.

There is another problem. While inflated annual DBE goals create the problems described above, it is the goals set on particular contracts that drive contractor behavior. These goals are always set on total dollar amount of the contracts. But the prime contract is awarded through a race-neutral low bid process. So if a non-DBE firm wins the prime contract, he must seek to meet the DBE goal on the work subcontracted. Frequently, the recipient will require a prime contractor to do certain proportion of the work himself, so subcontracting opportunities are decreased. Further, efficiency and profitability may limit the amount of work prime contractors subcontract. Nevertheless, if the DBE goal is 20% of the total contract dollars and 40% of the work is subcontracted, then the prime contractor must award half of the subcontracting work to DBEs to meet the 20% overall goal. This policy is the principal cause of DBE subcontractor over-utilization. If goals were set only on the subcontracted work, the DBE program would be much more narrowly tailored.

E. Race Neutral Tailoring

As a part of the post-*Adarand* narrow tailoring initiative, the DBE regulations were amended to require recipients to maximize the use of race neutral alternatives to achieve annual DBE goals.[54] Serious consideration of race neutral alternatives is now a standard requirement of equal protection law.

The key to any successful use of race neutral alternatives is problem identification. You usually can not solve a problem you do not know exists. Recipients are supposed to set goals to overcome barriers caused by discrimination. But what barriers actually exist in a particular locality?

As Justice O'Connor articulated in *Croson*, there are two parts to the race-neutral contracting process: (1) enforcing anti-discrimination provisions,[55] and (2) reducing barriers that might discourage DBEs and other small firms from public contracting.[56] DOT does not require that either of those steps take place. Thus, recipients need take no steps to identify any discrimination that might affect their contracting process. DOT does not systematically track discrimination complaints. When GAO surveyed state and transit authorities, 81% said they had not received any discrimination complaints in a two-year period.[57]

Nor do recipients have much statistical evidence of discrimination. Until the *Western States'* mandate, only three state DOTs had completed disparity studies and they were almost unknown for airports or transit districts. So the DBE process is not being used to identify or remedy specific discrimination. Nor is it necessarily being used to reduce barriers. While some recipients have identified barriers and put in responsive programs, it is not required.

What DOT does require to maximize race neutral alternatives is a completely passive accounting process. If a DBE wins a low bid prime contract, or if DBE subcontractor utilization exceeds the goal set on a contract, either outcome is considered a race-neutral attainment that can be used to lower next year's race-conscious goal.[58] But neither accounting step identifies or remedies any act of discrimination, though they do reflect the underlying numbers oriented redistributive purpose of the DBE program.

CONCLUSION

All of the narrow tailoring steps described above would be helpful in confining preferences to the market areas, groups, and industries where there was at least a statistical inference that discrimination was taking place. Improving the goal-setting process would limit the windfalls now accruing to some DBE subcontractors. But none of these actions, though useful, would really solve the civil rights issue of using preferences only as remedies for those instances where identified discrimination had occurred. To tackle that problem, the concept of group presumptions of social disadvantage has to be reconsidered.[59]

In 1976, criticizing the PWEA set-asides, Justices Stewart commented:

> In today's society, it constitutes far too gross an oversimplification to assume that every single Negro, Spanish-speaking citizen, Oriental, Indian, Eskimo, and Aleut potentially interested in construction contracting currently suffers from the effects of past or present racial discrimination. Since the MBE set-aside must be viewed as resting upon such an assumption, it necessarily paints with too broad a brush.[60]

What was true about these groups then is even truer today. The addition of all white women, whose firms usually are the largest beneficiaries of DBE contracting preferences, to the presumption of social disadvantage only makes the problem

worse. In the current DBE program, Hillary, Condoleezza, and Alberto, to take a few names not at random, would all be presumed to be socially disadvantaged despite their personal elite educations, social status, and political power. If Barak Hussein Obama were from the Middle East, instead of having a Kenyan father and white American mother, he would have to prove personally by a preponderance of evidence that he is socially disadvantaged. The Senator need merely check off a box on the DBE certification form. In *Croson*, Justice O'Connor criticized Richmond for not inquiring into "whether or not a particular MBE seeking a racial preference has suffered from the effects of discrimination by the city or prime contractors" and declared that "the interest in avoiding the bureaucratic effort necessary to tailor remedial relief to those who truly suffered the effects of prior discrimination can not justify a rigid line drawn on the basis of a suspect classification."[61]

To be a certified DBE, firms need to supply a considerable amount of financial and managerial information, but the one question those considered presumptively socially disadvantaged never have to answer is whether the owners personally or their firms specifically suffered from discrimination. The presumption preempts the question.[62] This presumption of social disadvantage adheres for life as though social mobility were uncommon in America. Nor is the extent of the preferences stemming from the presumption of social disadvantage limited in time or number. Only the economic presumption is measured objectively, and serves to limit benefits. Discrimination against persons based on their immutable characteristics damages the economy and democracy itself. In the long run, the only way to narrow-tailor the DBE program is to restrict it to people and firms that have actually suffered discrimination, or to open it up to all small disadvantaged businesses. The bureaucracy and rules are already in place to do that, if skin color or genitalia are not the defining criteria.

Endnotes

1 In 1980, the Supreme Court upheld the PWEA set-asides by deferring to Congressional spending powers *(*Fullilove v. Klutznick, 448 U.S. 448), "without using the strict scrutiny test or any other traditional standard of equal protection review." (City of Richmond v. Croson, 488 U.S. 469,487 (1989).

2 Justice O'Connor wrote "To whatever racial group these citizens belong, their "personal rights" to be treated with equal dignity and respect are implicated by a rigid rule erecting race as the sole criterion in an aspect of public decision making."(*Croson* at 493).

3 *Id.* at 509.

4 In *Croson*, Justices O'Connor at 493, Stevens at 516 and Scalia at 524 expressed concern about the role of racial politics in creating race conscious contracting programs.

5 Western States Paving, Inc. v. Washington State Department of Transportation (WSDOT) 407 F.3d 983, 988 (9th Cir. 2005).

6 George La Noue, *Follow the Money: Who Benefits From The Federal Aviation Authority's DBE Program?*, The American Review of Public Administration (forthcoming).

7 These percentages are based on comparison of the DBE goals airports set based on their own estimates of DBE and non-DBE availability for their work and the actual distribution of subcontracting dollars and awards to DBEs and non-DBEs. The data set was obtained through a Freedom of Information Act request and encompassed the "Uniform Report of DBE Awards or Commitments and Payments" for the fiscal year 2004.

8 488 U.S. 469 (1989).

9 Seven years after *Croson*, two Richmond political scientists, one black and one white, found "The attempt to shape the agenda for downtown revitalization in a manner that would benefit black Richmonders through set-aside policies was part of a nationwide strategy for using urban political power to the advantage of African American constituents." W. Avon Drake & Robert D. Holsworth, Affirmative Action and the Stalled Quest for Black Progress, 161-162 (1996).

10 *Croson*, at 510.

11 *Id.* at 499.

12 *Id.* at 498.

13 *Id.* at 504.

14 *Id.* at 506.

15 *Id.* at 509.

16 Contractors Association of Eastern Pennsylvania v. City of Philadelphia, 893 F. Supp. 438 (E.D. Pa 1995), affirmed 91 F.3d 586 (3rd Cir. 1996); AGC v. Columbus, 936 F. Supp. 1363 (S.D. Ohio 1996) and Engineering Contractors of South Florida. v. Metropolitan Dade County, 943 F. Supp. 895, (S. D. Fla. 1996), affirmed 122 F. 3rd 895 (11th Cir. 1997).

17 "All racial classifications, imposed by whatever federal, state, or local governmental actor, must be analyzed by a reviewing court under strict scrutiny." 515 U.S.200, 227(1995).

18 For a analysis of the Clinton Administration's three compelling interest documents (the benchmark limits study, the Urban Institute report and Appendix A), *see* "Disparity Studies as Evidence of Discrimination in Federal Contracting" A BRIEFING HELD BEFORE THE UNITED STATES COMMISSION OF CIVIL RIGHTS, Washington D.C. December 16, 2005. Edited with findings in May 2006. (*hereinafter* USCCR report), pp.34-36,58-60, 79-80.

19 49 C.F.R 26.45 (b).

20 A small fig leaf was added to try to cover the overt racial and gender classification by permitting those not presumptively socially disadvantaged to qualify by showing they were personally and economically disadvantaged by a preponderance of evidence. This provision has turned out to be a very narrow gateway and less than 2% of DBEs are not members of presumptively eligible groups. Mostly, they are white males with physical handicaps.

21 In legislation, even where the standard of strict scrutiny was not involved, the Supreme Court has been critical of relying on limited legislative history. For example, *in Kimel v. Florida Board of Regents*, the Court held that:

> Congress never identified any pattern of age discrimination by the States, much less any discrimination that rose to the level of constitutional violation. The evidence supplied by petitioners to demonstrate such attention by Congress to age discrimination by the States falls well short of the mark. The evidence consists almost entirely of isolated sentences clipped from floor debates and legislative reports. 582 U.S. 62, 89 (2000).

22 *Adarand VII,* 228 F.3rd 1147 (10th Cir.2000), *Sherbrooke Turf, Inc.,* 345 F.3d 964 (8th Cir.2003, Western States Paving Co., Inc. v. Washington State Department of Transportation 407 F.3d 983 (9th Cir. 2005) and Northern Contracting, Inc. v. Illinois Department of Transportation, 473 F.3d 715 (7th Cir. 2007).

23 W. H. Scott Construction C. v. City of Jackson, 199 F. 3rd 206, 218-9 (5th Cir. 1999) (City can not rely on a disparity study it has not formally adopted).

24 *Western States,* 407 at 991. (Racial minorities constitute more than 20% of the population of the United States, but own only 9% of all businesses).

25 61 Fed. Reg. 26042, 28045, 1996.

26 SDBs are based on the same presumptions of minority social disadvantage as the DBE program.

27 There are several ways to calculate disparity ratios, but the most common is to divide capacity or availability into utilization. Thus in the East North Central region there was 11.8% SDB capacity divided into 11.5% SDB utilization which yields 97% or near parity.

28 National Research Council, "Analyzing Information on Women-Owned Small Businesses in Federal Contracting, The National Academy Press, Washington, 60. *See also,* George La Noue, *To the 'Disadvantaged' Go the Spoils?,* The Public Interest, No. 138, Winter 2000, 91-98.

29 GAO Report GAO-01-586, June 8, 2001. (*hereinafter* GAO report)

30 *Id.* at 7.

31 *Id.* at 29.

32 See cases cited in footnote 22.

33 The Bush administration took a similar stance and never updated the benchmark study.

34 *Croson* at 506.

35 *Western States* at 998.

36 Ibid., at pp. 997-998.

37 USDOT is providing a large part of the funding for these studies. Washington State had a study in progress before *Western States* was decided and is basing its new race conscious program on that study.

38 "Availability and Disparity Study, Nevada Department of Transportation" was completed by BBC Research and Consulting on June 15, 2007. (*hereinafter* Nevada study); "Availability and Disparity Study: California Department of Transportation" June 29, 2007 (*hereinafter* Caltrans study.)

39 Phillips and Jordan v. Watts, 13 F. Supp. 2d 1308, 1315 (N. D. Fla. 1999).

40 49 CFR 26.5.

41 *Croson* at 506.

42 *Id.*

43 O'Donnell Construction Co v. District of Columbia, 963, F.2d 420, 427(D.C.1992), Engineering Contractors Association of South Florida v. Metropolitan Dade County, 122 F. 3d 895, 928 (11th Cir. 1997), Monterey Mechanical Co. v. Wilson, 125 F.3d at 704, (9th Cir. 1997), Associated General Contractors of Ohio, Inc. v. Drabik, (6th Cir. 2000), and Builders Association of Greater Chicago v. County of Cook, 256 F. 3d 642, 767 (7th Cir. 2001).

44 26.45 (h).

45 *Western States* at 999.

46 Nevada study, Figure IV-8. These disparity ratios refer to all construction and engineering contracts awarded in the pre-*Western States* era. Different configurations of over and underutilization emerge in the post-*Western States* data (IV-10), the Nevada state highway race neutral program (IV-9), and the subcontractor awards data (V-4 showing mwbes at 115% of availability). Prior to *Western States*, Nevada had a 6% DBE goal equally divided between race conscious and race neutral means. The lack of any pattern in the group specific findings, however, did not provide a basis for Nevada to reestablish its race conscious program.

47 In FY 2006, Caltrans DBE goal was 12% all race-conscious.

48 *Croson* at 498.

49 Criticizing the results of a disparity ratio that aggregated the three construction SIC codes, one federal court noted "firms that build hospital (SIC 15) do not compete for County contracts with firms that lay asphalt (SIC 16) or firms that install plumbing. (SIC 15)." *Engineering Contractors*, 943 F. Supp. at 1560.

50 Hershall Gill Consulting Engineers, Inc. v. Miami-Dade County, 333 F. Supp. 2nd 1305 (S. D. Fla. 2004). (making separate findings for construction and architects and engineers).

51 For a comprehensive review of goal setting issues, *see,* George R. La Noue, *Setting Goals in Federal DBE Programs*, George Mason Civil Rights Law Review, Spring 2007.

52 In California, raw figures showed mwbes were 32% of the firms in highway -related specialties, but when that number was adjusted for the size, specialties and location of the firms and those mwbes too large to be certified DBEs were removed, the mwbe availability percentage fell to 13.5%. (California study ES, p2)

53 Nevada study, III-1. Minority and women owned firms may not seek certification because they can not meet the regulatory definition of "economic disadvantage" or because they may not be interested in government contracts. In California, 4of 5 mwbe firms were not certified DBEs either and Caltrans is proposing to count these uncertified mwbes firms for annual goal setting purposes.

54 49 C.F.R. Sec. 26.51 (a)

55 "... a city is justified in penalizing the discriminator and providing appropriate relief t the victim of such discrimination." *Croson* at 509.

56 "Even in the absence of evidence of discrimination, the city has at its disposal a whole array of race-neutral devises to increase the accessibility of city contracting opportunities to small entrepreneurs of all races." *Croson* at 509-10.

57 GAO report at 5-7. Of the 31 discrimination complaints across the nation, 29 had been investigated and discrimination was found in only four of these investigations. In *Croson*, that the lack of evidence that Richmond had tried to enforce its contracting anti-discrimination ordinance cast doubt in the Court's mind about whether any exclusionary practices actually existed. (469 U.S. at 503).

58 The regulations also permit counting as a race neutral outcome utilization of a DBE subcontractor selected for reasons other than DBE status, but few recipients actually collect that data. Nevertheless, about six states and 9% of all airports meet all their DBE goals through race neutral means.

59 George R. La Noue & John C. Sullivan, *Deconstructing the Affirmative Action Categories*, American Behavioral Scientist, Vol. 41, No. 7 April 1998.

60 *Id.* at 530.

61 *Croson* at 508.

62 The presumption is technically rebuttable, but no criteria have been established for challenging the social presumption and it not clear that such a challenge has ever been attempted or been successful.

Point-Counterpoint: Affirmative Action Hiring Practices in the New Court Era

Someone Should Sue

*By Roger Clegg**

Two things that the past few months have proved about the Supreme Court, when it comes to civil rights: five justices insist on interpreting statutory language to mean what it says, and are very, very skeptical about racial and ethnic preferences.

The thesis of this article follows rather directly from this observation: since American companies frequently use employment preferences based on race, ethnicity, or sex, and since these preferences are inconsistent with the text of the Civil Rights statutes, they will likely be struck down. Companies need to rethink them.

First, though, we need to define one term, namely "affirmative action." Its original meaning was taking positive, proactive steps—affirmative action, get it?—to get rid of discrimination. Another meaning is casting a wide net—recruiting far and wide for the best candidates, not just using an old-boy network. Neither of these kinds of affirmative action is controversial today or raises any legal issues. But that is not true of the use of preferences based on race, ethnicity, or sex—"affirmative discrimination," as Nathan Glazer aptly termed it. This means that the best qualified people are not being hired and promoted because of their skin color, the country their ancestors came from, or their gender, and this is both unfair and presumptively illegal.

Eight out of ten business executives said that affirmative-action programs had resulted in them giving jobs and promotions to applicants who were less qualified than others, according to a survey conducted ten years ago by Yankelovich Partners, and commissioned by the PBS show "Nightly Business Report." Things have only gotten worse since then.

It is not difficult to find evidence of corporate preferential treatment. Just visit some corporate websites, or look at their own brochures. The trumpeting of minority numbers is deafening, and it is implausible that this bean-counting does not reflect and encourage the use of quotas and preferences.

Consider Wal-Mart. The company has told its managers that they have "diversity goals," and that they should avoid discrepancies between the percentage of qualified minorities/females who apply and the percentage actually chosen—or risk losing at least part of their annual bonuses. Thus, if a manager is faced with hiring the most qualified candidate or meeting the diversity goal, she or he knows what to do.

Here is another example. Recently the Center for Equal Opportunity received an e-mail, apparently from one of Intel Corporation's employees, forwarding a description by Intel of its "Diversity Employee Referral Program." The gist is that Intel will pay a $6,000 bonus to employees who make successful hiring referrals of "women, African Americans, Hispanics and Native Americans," but only $2,000 for successful hiring referrals of anyone else, *i.e.*, men who are of European, Asian, or Middle Eastern background.

One employee of a large Fortune 500 company contacted the Center when the company announced (internally only, of course) that when managers were hiring interns, if they hired three, one had to be female, one a minority, and one a "top performer." (Note the soft bigotry of low expectations.)

According to an October 17, 2005 article in *Newsweek* about Xerox, "Managers are judged—and compensated —on meeting diversity goals." The article indicates that the company's CEO, Anne Mulcahy, is dismissive of affirmative discrimination concerns: "Tales of preferential treatment —along with numerical targets for women—might raise the ire of affirmative-action opponents. So be it. 'If [somebody] wanted to write an editorial in *The Wall Street Journal*, I don't particularly care,' Mulcahy says."

Companies may be assuming that some "diversity" justification in hiring and promotions will shield them from legal challenge, since the Supreme Court has accepted it for university admissions. This is not true. Statutory language makes the legal justifications for employment discrimination much weaker.

Student admission decisions are, for the most part, governed by Title VI of the Civil Rights Act of 1964, while hiring and promotion decisions are addressed by Title VII of that act. The courts have interpreted the two statutes differently, so that what is permissible under Title VI is not necessarily permissible under Title VII.

Title VII contains a categorical ban, forbidding any employer to "discriminate" on the basis of "race, color, religion, sex, or national origin" in hiring, firing or "otherwise… with respect to [an employee's] compensation, terms, conditions, or privileges of employment." And, unlike the situation with Title VI, the Court has not conflated Title VII with the Equal Protection Clause. Accordingly, the Court's recent ruling in the University of Michigan cases that the latter permits discrimination in the name of "diversity" is inapplicable.

Will the courts nonetheless create a "diversity" exception to Title VII's prohibition of racial and ethnic discrimination, as they have for Title VI? Even before a case reaches the Supreme Court, that is very unlikely.

The statute, again, admits to no exceptions. To be sure, the Court did allow racial preferences in *United Steelworkers v. Weber*, handed down in 1979, and preferences on the basis of sex in *Johnson v. Santa Clara Transportation Agency*, a 1987 decision. But the rationale the Court approved in these two cases was not based on "diversity" but on "remedying" or "redressing" past employment practices that resulted in a "manifest imbalance" of the discriminated against groups "in traditionally segregated job categories." In 2007, and with every tick of the clock, it is becoming less and less likely that a company can plausibly assert that any imbalance, manifest or not, is traceable to "traditional[] segregat[ion]."

** Roger Clegg is President and General Counsel of the Center for Equal Opportunity in Falls Church, Virginia.*

It is one thing to say that an anti-discrimination statute allows preferences in order to remedy discrimination; it is very different to say that such a statute allows discrimination so long as the employer and the courts think there is a good reason for it. There is simply no way to reconcile the latter "interpretation" with the words of the statute. (The distinction between remedial and non-remedial preferences is one that proved critical in the Court's decisions in the Seattle and Louisville school cases, by the way.)

Note also that the Court in *Johnson* stressed that preferences could be used only to "attain" and not to "maintain" greater balance; this would make no sense if the justification is diversity. What is more, the diversity rationale is premised on a belief in racial, ethnic, and gender differences that is quite at odds with the insistence in Title VII—and by five justices of the Court—that people be judged individually, without regard to stereotypes.

If a "diversity" exception is created, it is hard to see why other exceptions might not also be put forward. Yet, in the case of Title VII, and not for Title VI, Congress explicitly declined to create even a "bona fide occupational qualification" exception to the statute for race, even as it did so for sex, religion, and national origin.

Furthermore, the diversity rationale could be—and frequently is—used to support discrimination *against* members of racial, religious, and ethnic minority groups and women. If a company's aim is greater "diversity" and less "underrepresentation" in its workforce, this means that any group that is "overrepresented" will be on the short end of any preferential hiring or promotion. That means, depending on the company, racial and ethnic minorities and women could all lose out. It seems very unlikely that Title VII was written to allow such anti-minority and anti-female discrimination so long as an employer could adduce a business reason for it.

Does anti-minority policy in the name of diversity sound far-fetched? Xerox recently lost an employment discrimination case before the U.S. Court of Appeals for the Fifth Circuit. At issue was the company's "Balanced Workforce Initiative," begun "in the 1990s for the stated purpose of insuring that all racial and gender groups were proportionately balanced at all levels of the company." The Houston office detected a racial imbalance, and so its general manager took steps "to remedy the disproportionate racial representation" there, "set[ting] specific racial goals for each job and grade level…." The Fifth Circuit found that "the existence of the [Balanced Workforce Initiative] is sufficient to constitute direct evidence of a form or practice of discrimination." After all, "Xerox candidly identified explicit racial goals for each job and grade level," and the evidence "indicate[d] that managers were evaluated on how well they complied with" the initiative's objectives—an appalling company policy, and an excellent judicial decision. And here is the kicker: The plaintiffs were African-Americans, and the company had concluded that "blacks were over-represented and whites were under-represented."

So, it is not surprising that the two federal courts of appeals—one in the Third Circuit, and one in the Fifth—that have been presented with the diversity rationale in Title VII cases have refused to accept it.

Even before the ascendancy of Chief Justice Roberts and Associate Justice Alito, it was unlikely that there would have been five votes for the diversity rationale. In 1997, when the Court had granted review in the Third Circuit case just mentioned, the civil rights establishment was so afraid of losing on this issue that it raised enough money to pay off the claims of the plaintiff and the fees of her lawyer.

It is fine for companies to celebrate their diversity—and use "affirmative action"—if that means making their workplaces attractive and friendly to as many people as possible. But it is wrong for them to aim for a predetermined racial, ethnic, and gender mix, and use preferences in order to achieve it—wrong, and illegal.

Affirmative Action: Legally Sound And Good for American Business

*By Wade Henderson**

Corporate affirmative action makes good business sense and remains lawful under Supreme Court precedent which has been on the books for decades. American companies can rest assured: Employers have substantial latitude to use affirmative action to hire and promote a diverse workforce under Title VII of the Civil Rights Act of 1964, the law addressing race- and gender-based employment discrimination in the private sector.

Despite the recent Supreme Court ruling about school districts attempting to achieve classroom diversity through race-conscious policies, a wide range of corporate affirmative action programs remain on firm legal footing, because Congress has the final say on what private companies can do. And Congress, as the Supreme Court has recognized for thirty years, views such programs as tools to achieve Title VII's goal of eradicating the vestiges of discrimination in the private sector workforce.

Even a Supreme Court aggressively opposed to race-conscious policies would be loath to assail the lawfulness of affirmative action under Title VII. The current Court's most conservative members have recognized that stare decisis is at its high-water mark on issues of statutory interpretation, because if Congress disagreed with the Court's interpretation of what the legislature meant it could change the statute. Congress left the door open for corporate affirmative action policies to play a role in reaching the goal of workplace fairness. Why? Because centuries of discrimination made a simple ban on conscious discrimination against women and minorities inadequate to the task of restoring racial and gender fairness in the job market. The Supreme Court observed that Congress intended Title VII as a "catalyst to cause employers and unions to self-examine and to self-evaluate their employment practices," in order to root out the vestiges of discrimination. To this day, the residue of a long history of discrimination continues to manifest itself in the form of insidious bias in the American workplace, even though conscious discrimination is banned.

Seton Hall law professor Tristin Green observed that subtle forms of discrimination, not easily addressed by

** Wade Henderson is President and CEO of the Leadership Conference on Civil Rights.*

anti-discrimination enforcement, are an impediment to the advancement of minorities in the workplace.[1] African Americans and other minorities continue to face subconscious bias concerning their job qualifications, and they are often excluded from business circles that facilitate opportunities for white men. Meanwhile, the white men who have long held the great majority of top positions simply do not have to worry about unspoken devaluing of one's skills, or not being connected to the right social cliques at the office.

And studies show that a majority of both men and women in corporate offices agree that a "glass ceiling" exists for women. In other words, subtle bias—unspoken discomfort with women in supervisory roles, the lack of women in the clubby circles of management—makes it harder for women to break into various types of corporate jobs.

In the first few years following the enactment of Title VII, the prohibition of outright discrimination against women and minorities made only modest inroads into the vast gender and racial disparities in the job market. These subtle obstacles to the advancement of women and minorities were not readily overcome solely by anti-discrimination enforcement. The advent of corporate affirmative action programs beginning in the late 1970s brought about slow but sure progress.

Moreover, overcoming the continuing effects of discrimination is a decidedly good business practice. The list of companies that have implemented affirmative action programs is not limited to corporate do-gooders. As a Goldman Sachs adviser stated, "diversity is a good business practice;" "there is a connection between diversity and financial success," though not always readily quantifiable.

Affirmative action programs are good for business because they offset subconscious factors affecting a company's recruiters and interviewers, often rooted in negative stereotypes or "comfort levels," rather than explicit bigotry, which result in the exclusion of highly qualified minority and female applicants. And affirmative action counteracts the exclusion of talented women and minorities from informal good-old-boys networks.

Finally, by making the effort to promote and train qualified women and minorities so that one's workforce better reflects the diversity of the labor pool, companies can foster an image and environment that appeals to the many women and minorities entering the workforce from which they hire—thus achieving a wider pool of applicants, including talented applicants of all races and genders.

Although some ideological opponents of any type of affirmative action would claim that the recent Supreme Court cases involving the use of race by government entities, including the *Parents Involved* school integration case, raise questions about corporate affirmative action, the truth is that these cases shed little light on the issue of private sector affirmative action.

Under current Supreme Court precedent, the Equal Protection Clause of the Constitution limits public sector affirmative action to cases where the entity in question has a history of discriminatory conduct—unless there is another compelling justification for the race-conscious policy. In *Parents Involved*, the Court was split, with a narrow 5-4 majority concluding that diversity is a compelling justification for race-conscious school assignment policies where the schools did not themselves have an unremedied history of discrimination. It remains unclear whether a majority on the Court would support a diversity rationale for private actors governed by Title VII.

Ultimately, however, private employers do not need to point to diversity as a justification for race- and gender-conscious policies. The critical difference between Title VII and the Equal Protection Clause lies in the fact that under Title VII a private company can pursue affirmative action policies to correct an imbalance between its workforce and the labor pool at large, *even absent any demonstrable history of discrimination within that company*.

Given the continuing racial and gender discrepancies in many of America's business sectors, the goal of remedying such an imbalance continues to provide ample support for many corporate affirmative action policies. Whether the current Supreme Court likes such private sector policies is irrelevant; by declining to proscribe their use, Congress has tied the Court's hands.

The Supreme Court laid out the specific criteria for corporate affirmative action plans in two major Title VII affirmative action cases, *United Steelworkers v. Weber* (1979) and *Johnson v. Transportation Agency* (1987). The Court unequivocally held in those cases that Congress intended Title VII to allow voluntary affirmative action programs by private employers if they are designed to address "a manifest imbalance" in the representation of women or minorities in traditionally segregated job categories, as determined by comparing the percentage of minorities in an employer's workforce and the percentage in the qualified labor pool.[2] Unlike in the public sector, the company implementing the practice need not itself have engaged in any discriminatory practices which led to the imbalance.

Thus, for example, where an employer recruits nationally among college graduates and has shown that the rate of participation for a minority group or women in its entry-level workforce is conspicuously smaller than the percentage of recent college graduates from one of the respective groups, an affirmative action plan should be lawful under Title VII. Such plans are most likely to be viewed favorably by the courts when, rather than using set-asides or quotas, race and gender are used only as factors considered in a more broad-based evaluation of the individual applicant.

The Supreme Court has also held that companies may facilitate the selection of qualified minority or women employees for executive or other high-ranking positions, or for training programs for these positions, if the percentage of minorities and women in these upper-level positions is conspicuously out of balance with the percentage in the labor pool.

The contours of corporate affirmative action programs the law permits vary depending on the industry and labor pool. However, studies suggest that race and gender imbalances persist in many sectors. Across sectors, those imbalances tend to be especially pronounced in upper-level positions, even as the representation of women and minorities slowly improves at the entry level.

Using law firms as an example, recent data show that approximately 50 % of law school graduates are women and nearly 20 % are minorities. Law firm employment of women and minorities at all levels still lags behind their numbers in the qualified labor pool. At the entry level, there has been significant improvement in recent years, particularly for women. According to the EEOC, as of 2003, the number of women associates was approximately 40 %. African-American and Latino representation among associates at firms is much further behind, at approximately half the rate of their representation among law school graduates.

The contrast is even starker at higher levels. According to a 2005 National Association for Legal Career Professionals survey, only about 17 % of law firm partners are women, and less than 5 % belong to any minority group. Similar patterns exist in other industries, like finance, where EEOC data suggest that progress in participation for women and minorities has also been slow.

As the Supreme Court has recognized, the goal of corporate affirmative action programs should be to move the private sector to a place where such programs are no longer necessary. Well-designed programs are moving us in that direction but it is clear from extensive employment and education data that both the disparities and their underlying causes persist.

Corporate efforts to improve the representation of women and minorities among their employees remain legal. They are also sound business policies that offset the stubborn barriers to the participation of women and minorities in our economy, making American companies stronger and more competitive in the process. According to Jeffrey Norris, President of the Equal Employment Advisory Council (EEAC), "Affirmative action continues to be needed in employment to address the inequalities that still exist in some workplaces for women and minorities."

Endnotes

1 Tristin K. Green, *Discrimination in Workplace Dynamics: Toward a Structural Account of Disparate Treatment Theory*, 38 Harv. C.R.-C.L. L. Rev. 91 (2003).

2 The Supreme Court offered little guidance to elucidate what kinds of job categories can be considered "traditionally segregated." The Court appears to consider the manifest imbalance in gender or minority representation to be itself evidence of traditional segregation, which makes sense given the body of evidence demonstrating that these imbalances continue to emanate from both explicit and structural discrimination.

Corporations, Securities & Antitrust

The Roberts Court's Antitrust Jurisprudence: The Chicago School Marches On

*By Joshua D. Wright**

The Supreme Court issued four antitrust decisions this term—(the most since the 1989-90 term)—and seven cases over the past two years. The antitrust activity level of the Roberts Court thus far has exceeded the single case average of the Court prior to the 2003-04 term by a significant margin.[1] In addition to these decisions, the Roberts Court requested input from the government in six antitrust cases over the past three years. This flurry of antitrust activity, combined with an apparent willingness to reconsider long established precedents that conflict with modern antitrust theory, suggest that the Roberts Court will play a relatively significant role in shaping antitrust doctrine for years to come.

This article examines three of the Supreme Court's 2006-2007 decisions—*Leegin*, *Twombly*, and *Weyerhaeuser*—with the goal of characterizing the antitrust philosophy of the Roberts Court.[2] To preview my conclusion: I argue that the Roberts Court's jurisprudence is heavily influenced by the Chicago School of antitrust analysis. This is not a function of the Court's composition, but rather the inevitable result of what has been a largely uninterrupted march by the Chicago School on antitrust analysis. Chief Justice Roberts and Justice Alito were presumed to be conservative antitrust thinkers, but there was little evidence from their prior judicial output or litigation experience that either would offer any distinctively Chicagoan influence on the Court's jurisprudence.

"Chicago School," is a term that means many different things to different people in the antitrust community. It has been used to describe the contributions to economic thought from the University of Chicago in the 1930s and 40s, the school of antitrust analysis that derived from Aaron Director's teachings at the University of Chicago. The term also, unfortunately, has been used pejoratively to describe reflexively naïve non-interventionist antitrust policy. However, in this article, I employ the term to describe the three pillars of antitrust analysis derived from the Chicago Law and Economics movement led by Aaron Director: (1) rigorous application of price theory; (2) commitment to empiricism; and (3) appreciation of the role of error costs on the optimal design of legal rules.[3]

I. The Chicago School of Antitrust Analysis: A Brief History and Some Defining Characteristics

The history of the Chicago School's influence on antitrust analysis has been well documented.[4] Professors Jonathan Baker and Timothy Bresnahan usefully break the Chicago School's influence on antitrust into two separate analytical components.[5] The first component, "the Chicago School of industrial organization economics," consists of the work in industrial organization economics which aimed, and succeeded, at debunking the structure-performance-conduct paradigm and its hypothesized relationship between market concentration and price or profitability.[6] Especially influential in the dismantling of the structure-conduct-performance hypotheses was UCLA economist Harold Demsetz,[7] whose work was central to exposing the misspecification of this relationship in previous work by Joe Bain and followers, as well as offering efficiency justifications for the observed correlation: firms with large market shares could earn high profits as a result of obtaining efficiencies, exploiting economies of scale, or creating a superior product.[8]

The second component, "the Chicago School of antitrust analysis," primarily (but not exclusively) contributed empirical work in the form of case studies demonstrating that various business practices previously considered manifestly anticompetitive could be explained as efficient and pro-competitive. Perhaps the most well-known contribution of the Chicago School of antitrust was the "single monopoly profit theorem," which posits that only a single monopoly profit is to be had in any vertical chain of distribution. The logic of the theorem is that a firm with monopoly power at one level of distribution would prefer competition at every other level of the supply chain because that will reduce the price of the product to consumers, increase sales, and maximize total profits. The theorem has been applied to monopoly leveraging theories, as well as tying, essential facilities, vertical integration, and vertical restraints.

The basic features of this second component are generally attributable to the work of Aaron Director[9] and others from 1950 to the mid-1970s.[10] A group of eminent antitrust scholars such as Richard Posner, Robert Bork, and Frank Easterbrook followed in Director's footsteps, building on these studies and economic analysis, and advocating bright line presumptions, including per se legality, which reflected the growing consensus that most conduct is efficient most of the time.

This is not to say that the Chicago School's contributions to antitrust economics were completed by the 1970s, nor that they were limited to the ultimate rejection of the structure-conduct-performance paradigm. For example, Chicagoans have continued to contribute to our economic understanding of various business practices, despite the fact that developments in industrial organization economics for the past twenty years have relied primarily upon game theoretic modeling techniques. Recent "Chicagoan" contributions to antitrust economics include work on exclusive dealing,[11] slotting contracts,[12] and vertical restraints theory.[13]

** Joshua D. Wright is a Scholar in Residence in the Bureau of Competition at the Federal Trade Commission, and an Assistant Professor at George Mason University School of Law.* *He thanks Brandy Wagstaff for research assistance and Jon Baker, Thom Lambert, Geoff Manne, and Timothy Bresnahan for helpful comments and suggestions.* *This article relies on analysis that appears in Joshua D. Wright, The Roberts Court and the Chicago School of Antitrust: The 2006 Term and Beyond, __ Competition Policy International (2007).*

There is little doubt that the influence of the Chicago School on antitrust law and policy has been substantial, particularly in the Supreme Court. Supreme Court decisions such as *Sylvania*,[14] *Khan*,[15] *Trinko*,[16] and *Brooke Group*[17] were influenced by Chicago School thinking, not to mention the development of the 1982 Horizontal Merger Guidelines by Assistant Attorney General William Baxter. Indeed, the 1970s and '80s were marked by a dramatic shift in antitrust policies, a significant reduction in enforcement agency activity, and calls from Chicago School commentators for the use of bright line presumptions,[18] per se legality of vertical restraints,[19] and even repeal of the antitrust laws altogether.[20] Perhaps the Chicago School's most important and visible victory has been the continual narrowing of the per se rule, which, after *Leegin* lifted the prohibition on minimum resale price maintenance, exists only in naked price-fixing cases and in a weakened form in tying cases.

There are undoubtedly many themes to the Chicago School movement in antitrust. While some of these themes are shared universally by Chicago School proponents, it would be a mistake to contend that the movement was monolithic. To the contrary, the Chicago School exhibited substantial variance in interests, beliefs, and methodologies employed. However, I contend that the following three "methodological commitments" are distinctively, while perhaps not exclusively, Chicagoan in nature: (1) rigorous application of price theory; (2) the centrality of empiricism; and (3) emphasis on the social cost of legal errors in the design of antitrust rules. While the first claim probably will not generate any significant dispute, the second and to a lesser extent the third will attract some dissent and warrant greater discussion. Consequently, I spend the bulk of this section arguing that both (2) and (3) are indeed distinctively Chicagoan, while conceding that the Post-Chicago and Harvard Schools shared some of these commitments some of the time.

A. Rigorous Application of Price Theory

The first defining characteristic is the rigorous application of economic theory, especially, but not exclusively, neoclassical price theory, to problems of antitrust analysis. Richard Posner described the key distinguishing attribute of the Chicago School of antitrust was that it "view[ed] antitrust policy through the lens of price theory."[21] Because I suspect that most commentators will agree that application of price theory is indeed a distinctive characteristic of the Chicago School of antitrust, I will not expand on this point other than to offer two caveats.

The first caveat is that Chicago's application of price theory does not imply that both the Harvard School and post-Chicago applications of economic theory to antitrust lacked rigor. Although this criticism has been leveled at the contributions of the Harvard School to industrial organization in the 1950s and '60s,[22] most criticisms of the post-Chicago movement have focused on its excessive mathematical complexity and highly stylized models rather than lack of theoretical rigor. The primary difference between the post-Chicago and Chicago Schools with respect to economic theory is likely that the latter rejects game theory as a useful tool for policy analysis, while the former embraces it as its primary weapon. Importantly, one reason that the Chicago School favored price theory is its ability to generate testable implications for the purpose of empirical testing, while game theory has been criticized on the grounds that it produces too many equilibria to be useful.[23]

The second caveat is to recognize that many of the Chicagoan's contributions, especially in the area of vertical restraints, do not rely solely upon neoclassical price theory and the model of perfect competition. Several of the key contributions by Chicagoans shed the confines of the neoclassical price theory model of perfect competition in favor of reliance on the New Institutional Economics and its focus on institutional details and transaction costs. In a series of articles, Professor Alan Meese has correctly noted that strict adherence to the perfect competition model envisioned in neoclassical economics was not consistent with the Chicago explanations of vertical restraints, which depend on the presence of downward sloping demand curves.[24] While noting that this objection is not without some force, I adopt a "big tent" view of the philosophical underpinnings of the Chicago School here, which is inclusive of these contributions.

Adherence to neoclassical price theory was no doubt a hallmark characteristic of Chicago analysis—and much progress was made in advancing antitrust analysis with simple application of price theory. However, embracing a one-to-one correspondence between perfect competition and Chicago would be overly narrow and not capture the contributions of many members of the Chicago movement. Chicago School economists frequently deviated from the confines of the model of perfect competition where such deviation was useful to generate helpful insights about various business practices.[25] In fact, Chicagoans themselves were among the first to criticize reliance on the model of perfect competition as a useful benchmark for antitrust analysis.[26]

B. The Centrality of Empiricism

The second defining feature is the centrality of empiricism to the Chicago antitrust analysis research agenda. This, I realize, is a somewhat more controversial claim. Post-Chicago scholars have frequently argued that it is the Chicagoan views that are without empirical support.[27] This argument is in some tension with recent empirical surveys of vertical restraints which appear to support the view that these practices are not likely to produce anticompetitive effects and favor a presumption of legality.[28] The question I address here, however, is not whether the predictions of Chicago School models have generated superior predictive power relative to their Post-Chicago counterparts. Rather, my claim is merely that empirical testing is a central feature of the Chicago School analysis.

There is at least one set of generally undisputed empirical contributions from Chicago School economists: the debunking of the purported relationship between concentration and price asserted by proponents of the structure-conduct-performance paradigm.[29] However, even holding aside the contributions of these "early" Chicagoans, it is clear that the relative weight attached to empirical evidence by later Chicago antitrust scholars was also relatively high.

Perhaps the most striking example of a Chicago School scholar who offered substantial empirical contributions to the antitrust literature was George Stigler. Seminal Chicago School figures Ronald Coase and Demsetz have both noted Stigler's

dedication to empiricism with a note of admiration. Coase describes Stigler as moving effortlessly "from the marshaling of high theory to aphorism to detailed statistical analysis, a mingling of treatments which resembles, in this respect, the subtle and colourful Edgeworth. It is by a magic of his own that Stigler arrives at conclusions which are both unexpected and important."[30] Demsetz eloquently elaborates on this theme:

Housed in Stigler's mind, neoclassical theory had more than the usual quality of material with which to work. It was coupled with a joy in verification and with a strong work ethic and sense of duty to his profession. Intelligence, insight, wit, and style were evident in his writings. His articles and essays could not be ignored. They provoked readers to think and often to follow his lead. For some readers, they simply provoked. Stigler's passion for evidence gathering is also evident in his work, and he made no secret of it.[31]

Stigler's work lived up to the billing described by these prominent Chicagoan colleagues and displayed an unmistakable passion for empirics. It was the empirical flavor of his economic analysis that landed Stigler a Nobel Prize in 1982 for his "seminal studies of industrial structures, functioning of markets, and causes and effects of public regulation." Ironically, Stigler was initially rejected by the University of Chicago Economics Department for being "too empirical." In his 1964 presidential address to the American Economic Association, Stigler announced that the "age of quantification is now full upon us," and that this age would be characterized by policy analysis informed by empirical evidence.[32]

Stigler's body of work in industrial organization, which he referred to often as "microeconomics with evidence," is powerful proof of the centrality of empiricism to his own approach. For example, Stigler offered an early study of the effects of the antitrust laws,[33] an empirical assessment of block booking practices,[34] and a study of the economies of scale[35] introducing the survivorship principle. Perhaps the strongest support for Stigler's dedication to empirical evidence in the development of antitrust policy was his change in position in favor of de-concentration policy in the early 1950s. This change was in response to the state of empirical evidence debunking the consensus views concerning the relationship between concentration and profitability.[36]

The uniquely Stiglerian commitment to empiricism is a noteworthy feature of the Chicago School's contribution to antitrust analysis in its own right, but there are others who demonstrate a similar commitment. For example, the case studies offered by many Chicagoans have played an important role in antitrust policy. Former FTC Chairman Tim Muris has made special note of Benjamin Klein's case studies emphasizing the role of vertical restraints in facilitating dealer supply of promotional services when performance is difficult to measure.[37]

In sum, the Chicago School of antitrust analysis places a strong emphasis on empiricism both in the form of statistical analysis and case studies of specific restraints. One might view the Chicago commitment to price theory, and even measured deviations from price theory where useful to explain economic phenomenon, as an extension of the emphasis on empiricism because of the testable implications that follow from its application.

C. Adoption of the Error-Cost Framework

A third defining feature of the Chicago School of antitrust analysis is the emphasis on the relationship between antitrust liability rules, judicial error, and the social costs of those errors. From an economics perspective, it is socially optimal to adopt the rule that minimizes the expected cost of false acquittals, false convictions, and administrative costs. Not surprisingly, the error-cost approach is distinctively Chicagoan because it was pioneered by Judge Frank Easterbrook, a prominent Chicagoan.[38] Subsequently, several commentators have adopted this framework as a useful tool for understanding the design of antitrust rules.[39]

The error-cost framework begins with the presumption that the costs of false convictions in the antitrust context are likely to be significantly larger than the costs of false acquittals since judicial errors that wrongly excuse an anticompetitive practice will eventually be undone by competitive forces. On the other hand, judicial errors which wrongly condemn a pro-competitive practice are likely to have significant social costs, as such practices are condemned and not offset by market forces.

The insights of Judge Easterbrook's error-cost framework, combined with the application of price theory and sensitivity to the state of empirical evidence, could be powerful tools for improving antitrust policy. For example, David Evans and Jorge Padilla demonstrate that such an approach to tying favors a modified per se legality standard in which tying is deemed pro-competitive unless the plaintiff presents strong evidence that the tie was anti-competitive.[40] Their conclusion is based upon the formulation of prior beliefs concerning the likely competitive effects of tying grounded in an assessment of the empirical evidence evaluating both Chicago and post-Chicago economic theories. Evans and Padilla label their approach "Neo-Chicago" because it purportedly adds to the conventional Chicago approach to the error-cost framework. To the extent that this label helps to distinguish calls for presumptions of legality informed by decision-theoretic analysis from those who would argue for per se legality based solely upon the Chicago School "impossibility theorems," it may be a useful addition to the antitrust nomenclature. However, largely for expositional convenience, and also because it is quite fair to credit Judge Easterbrook's contribution of the error-cost framework to the Chicago School, I will use "Chicago" as synonymous with Evans and Padilla's "Neo-Chicago."

This is not to say that the Chicago School possesses an exclusive claim to placing significant weight on error and administrative costs in the design of antitrust standards. Indeed, FTC Commissioner William Kovacic has persuasively demonstrated that the Harvard School has played an integral role in promoting the administrability of antitrust rules, which is a predecessor of the error-cost framework discussed above.[41] Perhaps the most well-known proponents of this position are Professors Phillip Areeda and Donald Turner, who have consistently argued that antitrust rules should be administrable.[42] Harvard School's then-Judge Stephen Breyer incorporated the insights of the Harvard approach into antitrust doctrine in *Barry Wright*, noting that "antitrust laws very rarely reject... 'beneficial birds in hand' for the sake of more

speculative... 'birds in the bush.'"[43] Again, the Harvard School's sensitivity to the possibility of deterring pro-competitive conduct as a result of judicial error is related to the Chicago School's error-cost framework.

II. The Roberts Court's 2006-07 Antitrust Output

The Supreme Court heard four antitrust cases this term. In relative and historical terms, this is an astonishing level of activity. The Roberts Court's production over the past two terms, and its apparent comfort with complex antitrust issues, suggests that this Court is likely to remain interested and engaged in antitrust, even if not at its current rate of output. In this section, I summarize three of these decisions before turning to my central claim.

A. Leegin Creative Leather Products, Inc. v. PSKS, Inc.[44]

Leegin is a straightforward resale price maintenance (RPM) case involving a terminated dealer. The plaintiff, PSKS, operated a women's apparel store in Texas. The defendant, Leegin, manufactures and distributes a number of leather goods and accessories including handbags, shoes, and jewelry under the "Brighton" brand name. In 1997, Leegin introduced its RPM program, the "Brighton Retail Pricing and Promotion Policy," a marketing initiative under which it would sell its products exclusively to those retailers who complied with the suggested retail prices. When Leegin learned that PSKS was discounting the Brighton product line below the suggested retail prices, Leegin terminated PSKS and PSKS in turn filed suit alleging that Leegin's new marketing and promotion program violated the Sherman Act. The trial court found Leegin's policy per se illegal under *Dr. Miles*,[45] and the jury awarded a $1.2 million verdict which was upheld by the Fifth Circuit.[46]

Justice Kennedy authored the Supreme Court's majority opinion, reversing the Fifth Circuit, and was joined by Justices Scalia, Thomas, Roberts and Alito. Justice Kennedy's analysis largely adopted the structure of the argument offered by both the antitrust agencies and a group of economists in amicus briefs filed in support of Leegin and in favor of overturning *Dr. Miles*, and offered four central points: (1) per se analysis is reserved for restraints that, echoing the language of *Sylvania*,[47] "always, or almost always, reduce consumer welfare by limiting competition and output;" (2) economic theory strongly suggests that RPM does not meet that stringent standard; (3) empirical evidence comports with economic theory on RPM; and (4) stare decisis rationales for continuation of a per se rule and adhering to *Dr. Miles* are unpersuasive.

The majority launched their attack on *Dr. Miles* with a reminder that the rule of reason, and not per se analysis, is the default rule for antitrust analysis of any economic restraint, and deviation from this default is warranted only when the restraint is known to be "manifestly anticompetitive"[48] and "would always or almost always tend to restrict competition and decrease output."[49] Measured against this standard, Justice Kennedy finds the case for continued application of the per se rule profoundly lacking after a review of the theoretical justifications for RPM and the empirical evidence concerning its competitive effects.[50] While recognizing the potential for RPM to produce anticompetitive effects by facilitating collusion, the majority finds that the empirical literature suggests that efficient uses of RPM are not "infrequent or hypothetical," and that the standard for applying the per se rule has not been satisfied.

Justice Breyer's dissent offers an enthusiastic defense of *Dr. Miles*. Unfortunately, as discussed in greater detail by Professor Miller, the enthusiasm is not supported by evidence or economic theory.[51] While Justice Breyer begins his dissent by recognizing the "*always or almost always*" standard that must be satisfied in order to apply the per se rule (in the absence of overriding stare decisis concerns), his failure to understand the economics of vertical restraints and to recognize the state of empirical evidence are fatal to his argument. A brief summary of the flaws is instructive. First, with respect to the empirical evidence, Justice Breyer relies heavily on studies that cannot possibly show that RPM meets the relevant standard.[52] Second, Justice Breyer displays surprising unfamiliarity with the economics of vertical restraints, failing to recognize the point emphasized in both the majority opinion (and by extension, the FTC/ DOJ Brief and the Economists' Brief), that the key explanation for the use of RPM is Klein and Murphy's demonstration that RPM may be used to enforce efficient contracts involving promotional services or other non-contractible elements of performance. Breyer's contention that he does not "understand how, in the absence of free-riding, an established producer would need RPM" is also puzzling. The argument that vertical restraints can facilitate retailer supply of promotion even in the absence of dealer free-riding is cited in the majority opinion and explained in the Economists' Brief in a fairly accessible manner. This argument has been well accepted in the economics literature for over twenty years.

Of course, the antitrust enterprise does not turn solely on the view of economists and economic theory.[53] The dissent offers two further defenses of the *Dr. Miles* rule that turn upon principles of stare decisis and identifying Congressional intent in 1975. The stare decisis defense depends critically on Justice Breyer's assessment that the economic arguments in favor of overturning *Dr. Miles* have not changed "for close to half a century." This is not so. As discussed above, this characterization is undermined by the dissent's erroneous interpretation of the empirical evidence concerning RPM and a failure to understand the role of RPM in facilitating the increased supply of promotional services even without inter-dealer free-riding.

The dissent next argued that overruling *Dr. Miles* would effectively repeal the Consumer Goods Pricing Act of 1975 that repealed the 1937 Miller-Tydings Act, which allowed states to authorize RPM. The dissent argues that the repeal of the 1937 Act should be interpreted as a statement of Congressional intent to endorse application of the per se rule against RPM. The majority rejects this argument, noting that "the text of the Consumer Goods Pricing Act did not codify the rule of *per se* illegality for vertical price restraints. It rescinded statutory provisions that made them *per se* legal" and, therefore, merely placed RPM once again within the ambit of the Sherman Act.[54]

It remains to be seen what impact *Leegin* will have on antitrust jurisprudence more generally. In many ways, the decision's impact is likely to be limited for several reasons. First, manufacturers and retailers had adapted to *Dr. Miles* by creating innovative arrangements that avoided the application

of the per se rule and accomplished the functional equivalent of RPM. In this sense, *Leegin*'s marginal impact will be to allow transactors to accomplish these goals directly rather than circuitously, and presumably at a lower cost. Congress, presumably along with state legislatures, might also reduce *Leegin*'s impact by reviving *Dr. Miles*. One possible result will be a patchwork of laws on RPM, which are likely to impose significant costs on manufacturers attempting to navigate these standards across state lines.[55] Nonetheless, *Leegin* is a significant improvement in antitrust jurisprudence on a much broader level because it reconciles previously incoherent antitrust doctrine with modern economic thought. It is also a symbolic victory for the Chicago School in persuading the Court to abandon one of the last vestiges of the "pre-Chicago" era's hostility to vertical restraints.

B. Bell Atlantic Corp. v. Twombly[56]

While *Twombly* offered the Court an opportunity to clarify the pleading requirements under Section 1 of the Sherman Act, it has also been viewed as having greater procedural implications outside of the antitrust context for its apparent rejection of notice pleading in favor of a "plausibility pleading."[57] While some commentators have argued that *Twombly* is not likely to become very significant,[58] it undoubtedly alters the Section 1 landscape considerably by increasing the pleading burden imposed on plaintiffs alleging horizontal conspiracies. Some factual and procedural background is necessary to place the decision in context.

The plaintiff class alleged that four major local exchange carriers—Bell Atlantic, Bell South, Qwest Communications International, and SBC (ILECs)—colluded to block competitive entry by Competitive Local Exchange Carriers (CLECs), pursuant to the framework established by the 1996 Telecommunications Act, which required the incumbent carriers to sell local telephone services at wholesale rates, lease unbundled network services, and permit interconnection. The allegations themselves consisted of claims that the defendants agreed not to enter each other's territories as CLECs and to jointly prevent CLEC entry altogether.

The district court found that these allegations amounted simply to assertions of parallel conduct, and as such were vulnerable to dismissal, pursuant to defendants' Rule 12(b)(6) motions without allegations of additional "plus factors," such as those required at the summary judgment stage. The Second Circuit reversed unanimously, despite some hesitation and concern regarding the "sometimes colossal expense" of discovery in complex antitrust cases, and held that Rule 8(a) did not require allegations of the "plus factors" required to survive summary judgment.

Justice Souter authored the majority opinion in a 7-2 decision holding that "stating [a Section 1 claim] requires a complaint with enough factual matter (taken as true) to suggest that an agreement was made.... [This requirement] simply calls for enough fact to raise a reasonable expectation that discovery will reveal evidence of illegal agreement."[59] The majority makes clear that allegations of parallel conduct alone are not sufficient to survive the pleading stage, "retiring" and rejecting the "no set of facts" formulation favored by *Conley v. Gibson*,[60] despite the conventional rule disfavoring motions to dismiss in antitrust cases. The Court's rationale for increasing the pleading burden faced by plaintiffs in antitrust conspiracy cases is explicitly motivated by the desire to avoid the extraordinary costs of discovery in such cases unless there is good reason to believe that an agreement will be unearthed.

One lesson from Twombly is entirely clear: a conclusory "allegation of parallel conduct [with] a bare assertion of conspiracy" is not sufficient to plead a conspiracy without "a context that raises a suggestion of a preceding agreement, not merely parallel conduct that could just as well be independent action."[61] The application of the new "plausibility" standard to plaintiffs' claims was relatively straightforward as the allegations consisted of parallel conduct alone and no independent allegation of actual agreement among the ILECs. But it remains to be seen precisely what sort of allegations will be sufficient to survive a motion to dismiss. In one recent case, *In re OSB Litigation*,[62] plaintiffs' Section 1 allegations survived a post-*Twombly* motion to dismiss largely because the complaint described alleged repeated communications between rivals announcing an intention to shutdown plants and reduce output, and detailed the mechanism by which the collusive agreement was formed (involving use of published prices in a trade publication), monitored, and enforced.

The full implications of *Twombly* are yet to be seen. Concerns with false positives in Section 1 cases and the massive social costs of discovery clearly motivated the Court's push towards an increased pleading burden. An open question remains as to precisely what "plus factor" allegations will be sufficient, when added to parallel conduct, to survive *Twombly*'s more rigorous standard. One result of *Twombly*, which appears unavoidable, is that the plausibility standard may operate as "Full Employment Act" for economists who will now be called in at the pleading stages to declare that market conditions are conducive to coordination or tend to exclude the possibility of independent action.

C. Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber Co.[63]

Weyerhaeuser raised the issue of identifying the appropriate standard for "predatory buying" claims under Section 2 of the Sherman Act. Ross-Simmons, a saw mill in the Pacific Northwest, alleged that Weyerhaeuser overpaid for alder sawlogs in a scheme designed to drive its rivals out of business. The district court instructed the jury that Ross-Simmons was required to prove that Weyerhaeuser engaged in "conduct that has the effect of wrongly preventing or excluding competition or frustrating or impairing the efforts of the firms to compete for customers within the relevant market." With respect to the "predatory buying" allegation specifically, the district court instructed the jury that:

> One of [respondents'] contentions in this case is that the [petitioner] purchased more logs than it needed or paid a higher price for logs than necessary, in order to prevent [respondent] from obtaining the logs [it] needed at a fair price. If you find this to be true, you may regard it as an anti-competitive act.

The jury found in favor of Ross-Simmons and awarded $78.7 million. The Ninth Circuit affirmed the judgment, despite Weyerhaeuser's contention that the district court erred by not including both prongs of the *Brooke Group*[64] standard in the jury instruction. The Department of Justice and FTC

petitioned the Supreme Court for certiorari and submitted joint amicus briefs recommending that the Court apply the *Brooke Group* standard to predatory buying.

Justice Thomas authored the unanimous decision on behalf of the Court, which agreed with the position advocated by the enforcement agencies. In predatory buying cases, plaintiffs must demonstrate both that the buyer's conduct led to below-cost pricing of the buyer's outputs and that the buyer "has a dangerous probability of recouping the losses incurred in bidding up input prices through the exercise of monopsony power."[65] Because Ross-Simmons conceded that it had not satisfied the *Brooke Group* standard, the Court vacated the Ninth Circuit's judgment and remanded the case.

The Supreme Court's endorsement of the *Brooke Group* standard appears to rest on three principles. The first is that "predatory-pricing and predatory- bidding claims are analytically similar" as a matter of economic theory, suggesting that similar legal standards are appropriate.[66] The second is that the Court espouses a view that the probability of successful predatory buying, like predatory pricing, is very low,[67] in part because of the myriad of explanations for "bidding up" input prices in an effort to increase market share and output, hedge against price volatility, or as a result of a simple miscalculation.[68] Finally, the Court notes that like low output prices, higher input prices may result in increased consumer welfare as firms increase output.[69]

While the Supreme Court does not take the lower court to task for allowing this jury instruction, there appears to be little, if any, doubt that the Supreme Court was correct to reverse the Ninth Circuit's affirmation of a disastrous jury instruction that would require a determination as to whether a firm purchased more inputs than it "needed" or paid more than "necessary." Rather, the Supreme Court focused almost exclusively on the theoretical similarities between predatory pricing and buying, the attributes of the *Brooke Group* standard, and why the economic similarity should translate into symmetrical legal treatment. Interesting questions remain concerning the implications of *Weyerhaeuser* (does this unanimous opinion suggest that the Supreme Court may be willing to adopt the *Brooke Group* test to bundled discounts, "compensated" exclusive dealing, all-units discounts, or other forms of allegedly exclusionary conduct?) However, there seems to be very little dispute that the decision is correct on the merits.

I claim that these decisions, taken together, suggest an unmistakable connection to the characteristics of the Chicago School of antitrust analysis discussed above. So what is it about these decisions that suggests the Roberts Court has adopted a Chicago School approach to antitrust analysis? And, if I am correct, what does it tell us about where this prolific Court might venture next in the world of antitrust jurisprudence? The remainder of this essay is dedicated to a discussion of these issues.

III. The Roberts Court and The Chicago School

The Roberts Court's productivity in the 2006-07 term alone has supplied sufficient fodder to keep both commentators and practitioners busy analyzing this output for likely trends in future antitrust jurisprudence. There is no doubt that this Court is quite comfortable with antitrust. It has not shied away from complex issues requiring analysis of economic theory or, in the case of *Leegin*, overturning century-old precedent. Perhaps this is because the current justices, led by Justices Breyer and Stevens, have significant antitrust experience.[70] Justice Scalia is considered the Court's only true Chicago School author. Despite the fact that Justice Breyer taught antitrust at the University of Chicago, he is generally acknowledged as a member of the Harvard School with substantial antitrust expertise.[71]

The new Supreme Court justices are also familiar with antitrust issues. Chief Justice Roberts was involved in a significant amount of antitrust litigation, representing both plaintiffs and defendants in a wide variety of cases. Justice Alito's most discussed antitrust moment came in joining an important and vigorous dissent by Judge Greenberg in the controversial and heavily criticized *LePage's* decision.[72]

The antitrust output and experience of these two new Justices certainly would not have allowed one to confidently predict that the Roberts Court's jurisprudence would exhibit a distinctively Chicago flare. For example, consider the following excerpt from an article written by Chief Justice Roberts in 1994 addressing whether the Supreme Court was "conservative":

In the antitrust area, the Court seems to regain its equilibrium after the dizzying *Kodak* decision of two Terms ago. That decision surprised most observers by upholding a predatory pricing verdict based on dubious if not implausible economic theory. In the 1992–93 Term, in three decisions the Court returned to a regime in which the objective economic realities of the marketplace take precedence over fuzzy economic theorizing or the conspiracy theories of plaintiffs' lawyers. This is bad news for professors and lawyers, good news for business.[73]

Admittedly, the implicit critique of *Kodak* appears to be consistent with Chicago School views. But the excerpt also exhibits some aversion to the application of economic theories —at least fuzzy ones—and academic theorizing more generally and especially when it is detached from real world market conditions and empirical realities. While there are kernels in the antitrust history of both judges that might encourage Chicagoans and post-Chicagoans, it is difficult to generalize any antitrust philosophy from these limited sources.[74]

Leegin bears all of the identifying marks of Chicago School influence. Justice Kennedy's analysis applies Chicago economic theory to minimum RPM in order to assess its likely competitive effects. The *Leegin* majority recognizes the several pro-competitive rationales for vertical restraints in the economics literature, many pioneered by Chicagoans, including the use of vertical restraints to facilitate the provision of promotional services in the absence of dealer free-riding. Importantly, *Leegin* at least implicitly broadens the Court's view of the role of vertical restraints outside of the conventional "inter-dealer" or "discount" dealer free-riding rationale, which does not appear to explain many instances of RPM. In summarizing the theoretical literature, the Court notes that the "economics literature is replete with procompetitive justifications for a manufacturer's use of resale price maintenance."[75]

Leegin also displays the two remaining Chicago School characteristics: reliance on empiricism and sensitivity to error costs in designing antitrust rules. Justice Kennedy certainly displays sensitivity to the available empirical evidence concerning the competitive effects of RPM, emphasizing documentation that the practice is infrequently associated with anticompetitive effects. Specifically, the Court notes that "[t]he few recent studies documenting the competitive effects of resale price maintenance also cast doubt on the conclusion that the practice meets the criteria for a *per se* rule."[76]

Finally, the majority also embraces the error-cost framework. This is not surprising since this framework is embodied in *Business Electronics*, limiting the application of per se rules to restraints that are "always or almost always" anticompetitive. But the Court goes further than such an implicit recognition of the error-cost framework when rejecting the argument that per se illegality is the appropriate antitrust default rule on the grounds that per se rules decrease administrative costs. The Court's response clearly reveals that its view of the proper scope of per se rules is illuminated by Judge Easterbrook's error-cost framework:

> *Per se* rules may decrease administrative costs, but that is only part of the equation. Those rules can be counterproductive. They can increase the total cost of the antitrust system by prohibiting procompetitive conduct the antitrust laws should encourage. See Easterbrook, *Vertical Arrangements and the Rule of Reason*, 53 Antitrust L.J. 135, 158 (1984).[77]

Twombly also strongly exhibits two of the three Chicago characteristics set forth above, and arguably the third as well. There is no doubt that the Court's decision to heighten the pleading burden facing plaintiffs alleging conspiracy in violation of Section 1 is influenced by the error-cost analysis. As discussed above, the Court explicitly motivates its reasoning with reference to the massive social costs imposed by allowing discovery in cases that are not likely associated with real collusion. The Court notes that conspiracy allegations are especially ripe for false positives because parallel conduct might well arise from competitive behavior, and that those considerations favor more rigorous pleading standards.

But does *Twombly* have separate antitrust content, or is it an opinion about procedure with some collateral antitrust implications? I would argue that the former interpretation is correct. Justice Souter's opinion extends the logic of *Matsushita* and *Monsanto*, seeking to avoid false inferences of conspiracy at the pleading stage. This extension itself has important antitrust implications. One such implication is that lower courts will be faced with the challenge of assessing whether conditions tending to exclude the possibility of independent action are present *before* discovery has occurred.

Returning to the claim that *Twombly* was influenced by Chicago logic, the majority's analysis also displays commitment to the application of economic theory. *Twombly*'s primary antitrust lesson is that lower courts are to analyze the "plausibility" of the conspiracy allegations in light of "common economic experience." This lesson combines the Chicago School principles of application of economic theory and the centrality of empiricism. What role does evaluation of the "common economic experience" have in determining "plausibility"? *Twombly*'s analysis of market conditions suggests that rational, profit-maximizing independent action is the likely explanation of the ILEC's parallel conduct. Applied outside the case at bar, *Twombly* requires that the market conditions must be conducive to coordination *and* tend to exclude the possibility of independent action.

But where does a court turn to evaluate whether the "common economic experience" and market conditions are conducive to agreement? The answer is economic theory, and an evaluation of empirical realities. Specifically, the modern oligopoly theory built upon the work of Chicago's George Stigler lays the foundation for this analysis in a manner that provides useful guidance to courts by focusing on the conditions that lower the costs of forming, monitoring, and enforcing a collusive agreement.[78] *Twombly* requires lower courts to evaluate market realities to determine whether they are consistent with those conditions that would support an inference of conspiracy.

Weyerhaeuser also fits nicely into the Chicago School framework described above, with respect to its application of economic theory to predatory bidding and its consistency with the error-cost framework. Justice Thomas's opinion, however, demonstrates very little interest in empiricism. As discussed above, Justice Thomas's opinion on behalf of the unanimous Court begins with what reads much like a literature survey, noting the consensus view of economists that predatory buying is analytically identical to predatory pricing. This reliance on economic theory allows the Court to both equate monopsony and monopoly analysis for the purposes of antitrust and set the stage to adopt the *Brooke Group* standard. The reliance on *Brooke Group* makes clear that the error-cost framework plays a central role in Justice Thomas's analysis, relying on both the low probability of competitive harm associated with predatory buying,[79] as well as the economic logic that predatory pricing is likely to benefit consumers, to justify adoption of the *Brooke Group* standard.

The Roberts Court's antitrust output generally appears to embrace the Chicago School principles identified in Part II. I offer this as a descriptive theory of these cases rather than a normative judgment on their merits. Such a description may be useful in its own right in highlighting these aspects of the Roberts Court's antitrust jurisprudence. Nor do I wish to overstate my claim as denying the existence of any distinctively Harvard or Post-Chicago themes in these cases. But, for the most part, I believe these cases largely adopt what can accurately be described as a Chicago School approach.

One can anticipate the objection that the Supreme Court, at least since *Sylvania*, has long been influenced by Chicago School and so the Roberts Court's antitrust output is merely reflective of the status quo that persisted prior to the 2006-07 term. While that argument is not without force, and it is certainly true that Chicago School principles are not new to Supreme Court antitrust jurisprudence, it was unclear that the Roberts Court would adopt a Chicago School approach to antitrust analysis. Even if it were true that the Roberts Court's antitrust jurisprudence represents a mere continuation of a pre-existing trend, that point would not detract from the importance of identifying the distinctive themes displayed by

the Roberts Court, which has proven unique in its productivity, willingness to engage antitrust issues, and its familiarity and expertise with the subject matter. Holding these points aside, another useful application of this descriptive theory is the generation of some predictions concerning the future antitrust output of the Roberts Court.

IV. Where Will The Roberts Court Go Next?

The Roberts Court's interest in and proclivity for antitrust analysis raises the question: Where will they go next? Is the Court going to limit itself to "clean up" decisions such as *Independent Ink* and *Leegin* that correct long standing and broadly criticized precedents? Will the Court intervene only in cases where an economic consensus is apparent in the literature, such as *Weyerhaeuser* and *Leegin,* rather than engaging in its own hands-on economic analysis? An aversion to taking on complex antitrust issues where such a consensus does not exist might explain the Court's unwillingness to grant certiorari in *Tamoxifen*.[80] Or will the Court be willing to engage some of the more difficult and complex issues of the day, such as addressing the correct standard for unilateral "exclusionary pricing" in cases such as *LePage's*? Or perhaps the Roberts Court will tackle a horizontal merger case? I offer some predictions on topics that the Supreme Court may take on in the near future that are consistent with the analysis above.[81]

The first prediction is that the Roberts Court will finally take on a horizontal merger decision. The Supreme Court has not offered any substantive guidance on horizontal mergers in over thirty years,[82] allowing merger analysis to develop amongst the lower courts with substantial influence from the antitrust agencies in the form of the Horizontal Merger Guidelines. There are, of course, significant obstacles to the Supreme Court addressing a merger case in the near future even if it is so inclined, such as the elimination of automatic direct appeal. Nonetheless, a Supreme Court merger opinion may be consistent with the pattern exhibited in the 2006-07 term. Economic theory, and the Merger Guidelines, both suggest that the structural presumptions in the Supreme Court jurisprudence do not make much economic sense and do not reflect modern economic learning concerning the potential unilateral effects of mergers or the competitive effects of mergers. The Supreme Court may take advantage of this economic consensus and "clean up" troublesome merger decisions. Such a decision would be consistent with the Supreme Court's revealed preference for relying on economic consensus to overturn problematic, if not long-lived, precedents.[83]

In the same spirit, I predict the Roberts Court will overturn *Jefferson Parish's* modified per se rule in favor of the rule of reason, thus eliminating the last vestiges of the hostile approach to vertical contracting practices of antitrust era's past.[84] This is another area that matches the criteria set forth above. Economic theory suggests an overwhelming consensus that, like RPM, the literature is "replete" with pro-competitive explanations for tying. The empirical evidence, if not only in the form of ubiquitous tying in the economy by firms both with and without any market power of antitrust concern, bolsters the case for abandoning the per se rule. Finally, application of the error-cost framework to tying suggests a structured rule of reason approach adopting a presumption of *legality*—certainly not the per se rule of illegality.[85]

A third prediction is that the Court will eventually agree to hear a case challenging patent settlements in the pharmaceutical industry involving "reverse payments," although it did not grant certiorari in *Tamoxifen* this year. One view of the Court's denial of certiorari on reverse payments cases to this point is that the consensus economic and empirical view on these issues is still emerging, as evidenced by the antitrust agencies' disagreement as to the ripeness of reverse payment cases for review. In any case, reverse payments do not present quite the low hanging fruit presented in cases such as *Weyerhaeuser* and *Leegin*. However, a circuit split on these issues is likely to develop, and our empirical knowledge of these settlements is likely to improve over time with increased study, both which militate in favor of certiorari.

I conclude with one area where I am less convinced that the Roberts Court will apply its impressive energies in the antitrust realm: exclusionary pricing in the form of bundled rebates or loyalty discounts. While there is broad consensus that *LePage's* adopted a nonsensical "harm to competitor" standard in lieu of requiring harm to competition, and many have argued that *Brooke Group* or a modified *Brooke Group* approach should apply to all discounting conduct, no real consensus has emerged as to the appropriate test to apply to bundled rebates or loyalty discounts. In addition, the economic literature on bundled rebates and loyalty discounts is growing, with much attention paid to anticompetitive theories that have not yet been subjected to empirical testing and, therefore, may not be "ready for primetime."[86] Even further, economic research exploring pro-competitive justifications for bundled rebates, partial and limited exclusive contracts, and loyalty discounts is still emerging. In the absence of any economic or empirical consensus, and no clear benefit in deviating from the rule of reason approach to exclusionary pricing cases, it is unlikely that the Court will be motivated to address these issues.

Endnotes

1 J. Thomas Rosch, "A New Direction for Antitrust at the Supreme Court?," Remarks Before the Antitrust Section of the Minnesota Bar (March 1, 2007).

2 I omit analysis of *Credit Suisse Sec. (USA) LLC v. Billing,* 127 S. Ct. 2383 (2007). For an excellent discussion of Credit Suisse's potential impact on antitrust activity in regulated industries, *see* Keith Sharfman, *Credit Suisse, Regulatory Immunity, and the Shrinking Scope of Antitrust*, 2007 eSapience Ctr. for Competition Pol'y (arguing that the "clearly incompatible" standard threatens to render mere regulatory overlap a sufficient condition for implied immunity from the antitrust laws).

3 I do not claim that other "schools" of economic thought are not also associated with these themes. My claim, *infra* Part II.B., is that the Chicago School is uniquely associated with this combination of characteristics.

4 *See, e.g.,* Richard A. Posner, *The Chicago School of Antitrust*, 127 U. Penn. L. Rev. 925 (1979); Robert H. Bork, The Antitrust Paradox (New York: Free Press 1978); William E. Kovacic & Carl Shapiro, *Antitrust Policy: A Century of Economic and Legal Thinking*, 14 J. Econ. Persp. 43 (2000); Edmund W. Kitch, The *Fire of Truth: Remembrance of Law and Economics at Chicago, 1932-70*, J.L. & Econ. 163 (1983); Alan J. Meese, *Tying Meets the*

New Institutional Economics: Farewell to the Chimera of Forcing, 146 U. Penn. L. Rev. 1 (1997); William H. Page, *The Chicago School and The Evolution of Antitrust: Characterization, Antitrust Injury, and Evidentiary Sufficiency*, 75 Va. L. Rev. 1221 (1989).

5 Jonathan B. Baker & Timothy F. Bresnahan, *Economic Evidence in Antitrust: Defining Markets and Measuring Market Power* 23-26 (Stanford Law School, Working Paper No. 328, July 2006).

6 *See, e.g.*, Yale Brozen et al., Concentration, Mergers, and Public Policy (1982) (questioning the causal link between market concentration and price and providing alternative, efficiency-based explanations for the correlation); Industrial Concentration: The New Learning (Harvey J. Goldschmid et al. eds., 1974).

7 Professors Demsetz and Armen Alchian are frequently associated with the Chicago School despite the fact that both spent the bulk of their careers at UCLA. As any UCLA economist should note, the antitrust community has allowed the Chicago School to take credit for many of the contributions from UCLA. The contributions of the UCLA economists to antitrust analysis are discussed by former Federal Trade Commission (FTC) Chairman, and UCLA alumnus, Timothy J. Muris. *See* Timothy J. Muris, *Improving the Economic Foundations of Competition Policy*, 12 Geo Mason. L. Rev. 1 (2003).

8 Harold Demsetz, *Two Systems of Belief About Monopoly*, *in* Industrial Concentration: The New Learning (Harvey J. Goldschmid et al. eds., 1974). The contributions of Demsetz and other participants in the famous Airlie House Conference are discussed in Timothy J. Muris, *Economics and Antitrust*, 5 Geo Mason. L. Rev. 303 (1997).

9 *See* The New Palgrave Dictionary of Economics and the Law 227-33, 601-05 (Peter Newman ed., Macmillan Reference 1998); *see also* Sam Peltzman, *Aaron Director's Influence on Antitrust Policy*, 48 J.L. & Econ. 313 (2005).

10 Seminal contributions from the Chicago literature include, but are not limited to, Robert H. Bork, *Vertical Integration and the Sherman Act: The Legal History of an Economic Misconception*, 22 U. Chi. L. Rev. 157 (1954); Aaron Director & Edward H. Levi, *Law and the Future of Trade Regulation*, 51 Nw. U. L. Rev. 281 (1956); Ward S. Bowman, *Tying Arrangements and the Leverage Problem*, 67 Yale L.J. 19 (1957); John S. McGee, *Predatory Price Cutting: The Standard Oil (NJ) Case*, 1 J.L. & Econ. 137 (1958); Lester G. Telser, *Why Should Manufacturers Want Fair Trade?* 3 J.L. & Econ. 86 (1960).

11 Howard P. Marvel, *Exclusive Dealing*, 25 J.L & Econ. 1 (1982); Benjamin Klein & Andres V. Lerner, *The Expanded Economics of Free-Riding: How Exclusive Dealing Prevents Free-Riding and Creates Undivided Loyalty*, 74 Antitrust L.J. 473 (2007).

12 Benjamin Klein & Joshua D. Wright, *The Economics of Slotting Contracts*, 50 J.L. & Econ. (forthcoming 2007); Joshua D. Wright, *Slotting Contracts and Consumer Welfare*, 74 Antitrust L.J. 439 (2007); Joshua D. Wright, *Antitrust Law and Competition for Distribution,* 23 Yale J. on Reg. 169 (2006).

13 Benjamin Klein & Kevin M. Murphy, *Vertical Restraints as Contract Enforcement Mechanisms*, 31 J.L. & Econ. 265 (1988).

14 Continental T.V. v. GTE Sylvania, 433 U.S. 36 (1977).

15 State Oil Co. v. Khan, 522 U.S. 3 (1997).

16 Verizon Commc'ns Inc. v. Law Offices of Curtis V. Trinko, LLP, 540 U.S. 398 (2004)

17 Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209 (1993).

18 *See, e.g.*, Frank Easterbrook, *The Limits of Antitrust*, 65 Tex. L. Rev. 1 (1984).

19 Richard A. Posner, *The Next Step in the Antitrust Treatment of Restricted Distribution: Per Se Legality*, 48 U. Chi. L. Rev. 6 (1981).

20 *See, e.g.*, D. T. Armentano, Antitrust Policy: The Case for Repeal (1986).

21 Posner, *Chicago School of Antitrust*, *supra* note 4, at 928; *accord* Bork, The Antitrust Paradox, *supra* note 4, at 117.

22 *See, e.g.*, Posner, *Chicago School of Antitrust*, *supra* note 4, at 928-29:

> It is still fair to ask why the application of price theory to antitrust should have been a novelty. The answer, I believe, is that in the 1950s and early 1960s, industrial organization, the field of economics that studies monopoly questions, tended to be untheoretical, descriptive, "institutional," and even metaphorical. Casual observations of business behavior, colorful characterizations (such as the term "barrier to entry"), eclectic forays into sociology and psychology, descriptive statistics, and verification by plausibility took the place of the careful definitions and parsimonious logical structure of economic theory. The result was that industrial organization regularly advanced propositions that contradicted economic theory.

23 *See* Bruce H. Kobayashi, *Game Theory and Antitrust, A Post-Mortem*, 5 Geo. Mason L. Rev. 411, 412 (1997) (criticizing the application of game theory in antitrust on the grounds that "game theoretic models of [industrial organization] have not been empirically verified in a meaningful sense"). *See also* David Evans & Jorge Padilla, *Designing Antitrust Rules for Assessing Unilateral Practices: A Neo-Chicago Approach*, 72 U. Chi. L. Rev. 73, 98 (2005) ("it has yet to demonstrate a capacity to produce what we would call identification theorems—useful descriptions of the circumstances determining whether a practice is procompetitive or anticompetitive").

24 *See, e.g.*, Alan J. Meese, *Price Theory and Vertical Restraints: A Misunderstood Relation*, 45 UCLA L. Rev. 143 (1997); Alan J. Meese, *Exclusive Dealing, The Theory of the Firm, and Raising Rivals' Costs: Toward a New Synthesis*, 50 Antitrust Bull. 371 (2005); Alan J. Meese, *Price Theory, Competition, and the Rule of Reason*, 2003 U. Ill. L. Rev. 77; Alan J. Meese, *Market Failure and Non-Standard Contracting: How the Ghost of Perfect Competition Still Haunts Antitrust*, 1 J. Competition L. & Econ. 21 (2005).

25 *See, e.g.*, George J. Stigler, *The Economics of Information*, 69 J. Pol. Econ. 213 (1964) (analyzing the economics of information from a search cost perspective whereas search costs would not exist under perfect competition); Telser, *supra* note 10 (analyzing resale price maintenance); Klein & Murphy, *supra* note 13; Benjamin Klein, *Market Power in Aftermarkets*, 17 Managerial & Decision Econ. 143 (1996); Klein & Lerner, *supra* note 11 (analyzing the role of exclusive dealing contracts in preventing dealer free-riding).

26 *See* Harold Demsetz, *100 Years of Antitrust: Should We Celebrate*?, Brent T. Upson Memorial Lecture, George Mason University School of Law, Law and Economics Center (1991).

27 For example, at a recent antitrust conference at Georgetown University on "Conservative Economic Influence on U.S. Antitrust Policy," the following panel discussions questioning the empirical underpinnings of various assumptions were held: (1) Is the Assumption Valid That Cartels Are Fragile and Temporary - Particularly Because of the Difficulty of Controlling Cheating?; (2) Is It Valid to Assume that Vertical Arrangements (Merger and Distribution) Can Very Rarely Injure Consumer Welfare?; *(3)* Has the "Free Rider" Explanation for Vertical Arrangements Been Unrealistically Expanded?; and (4) Has Merger Enforcement Been Unduly Influenced by Conservative Economic Analysis: Consider Barriers to Entry and Structural Presumptions? A gambler might wager with some confidence that the answers to these questions were likely "No, No, Yes, and Yes," respectively.

28 *See* James C. Cooper et al., *Vertical Antitrust Policy as a Problem of Inference*, 23 Int'l J. Indus. Org. 639 (2005); Francine Lafontaine & Margaret Slade, *Exclusive* Contracts and Vertical Restraints: Empirical Evidence and Public Policy, *in* Handbook of Antitrust Economics (Paolo Buccirossi ed., forthcoming).

29 *See* Brozen, *supra* note 6; Demsetz, *Two Systems of Belief About Monopoly*, *supra* note 8.

30 R.H. Coase, *George J. Stigler*, *in* Essays on Economics and Economists (1994).

31 Harold Demsetz, *George J. Stigler: Midcentury Neoclassicalist with a Passion to Quantify*, 101 J. Pol. Econ. 793 (1993).

32 *See, e.g.*, George J. Stigler, *The Economists and the State*, 55 Am. Econ. Rev. 1 (1965):

> It will become inconceivable that the margin requirements on securities markets will be altered once a year without knowing whether they have even a modest effect. It will become impossible for an import-quota system to evade calculus of gains and costs.... Studies will inevitably and irresistibly enter into the subject of public policy, and we shall develop a body of knowledge essential to intelligent policy formation.

33 George J. Stigler, *The Economic Effects of the Antitrust Laws*, 9 J.L. & Econ. 225 (1966).

34 George J. Stigler, United States v. Loew's*: A Note on Block Booking*, 1963 Sup. Ct. Rev. 152 (1963).

35 George J. Stigler, *The Economies of Scale*, 1 J.L. & Econ. 54 (1958).

36 George J. Stigler, The Memoirs of an Unregulated Economist 97-100 (1988).

37 *See* Muris, *supra* note 7, at 17. The seminal article from Klein and Murphy, *supra* note 13, includes a detailed discussion of Coors' use of vertical restraints to solve dealer free-riding problems.

38 Easterbrook, *supra* note 18.

39 *See, e.g.*, Evans & Padilla, *supra* note 23; C. Frederick Beckner III & Steven C. Salop, *Decision Theory and Antitrust Rules*, 67 Antitrust L.J. 41 (1999); Keith N. Hylton & Michael Salinger, *Tying Law and Policy: A Decision-Theoretic Approach*, 69 Antitrust L.J. 469 (2001); Luke Froeb et al., *Vertical Antitrust Policy as a Problem of Inference*, 23 Int'l J. Indus. Org. 639 (2005).

40 Evans & Padilla, *supra* note 23. Others have applied the error-cost framework in a similar manner. *See supra* note 39.

41 *See* William Kovacic, *The Intellectual DNA of Modern U.S. Competition Law for Dominant Firm Conduct: The Chicago-Harvard Double Helix*, 2007 Colum. Bus. L. Rev. 1 (2007).

42 I Phillip Areeda & Donald Turner, Antitrust Law 31-33 (1978).

43 Barry Wright Corp. v. ITT Grinnell Corp., 724 F.2d 227, 234 (1st Cir. 1983).

44 127 S. Ct. 2705 (2007).

45 Dr. Miles Med. Co. v. John D. Park & Sons Co., 220 U.S. 373 (1911).

46 171 F. App'x 464 (2006) (per curiam).

47 Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36 (1977).

48 *Id.* at 49-50.

49 Business Elecs. Corp. v. Sharp Elecs. Corp., 485 U.S. 717, 723 (1988) (quoting Northwest Wholesale Stationers, Inc. v. Pacific Stationery & Printing Co., 472 U.S. 284, 289-90 (1985)).

50 Importantly, the majority does not limit its discussion of justifications for RPM to the conventional "discount dealer" free-riding story, instead it finds the literature "replete with pro-competitive justifications" and notes the consensus on this point amongst economists. Importantly, the majority also recognizes that RPM might be used to encourage retailer services even where inter-dealer free-riding is not possible. This argument has long been accepted in the economics literature, first introduced by Klein & Murphy, *supra* note 13, and later formalized by Frank Mathewson & Ralph Winter, *The Law and Economics of Resale Price Maintenance*, 13 Rev. Indus. Org. 57, 74–75 (1998). Until *Leegin*, antitrust legal analysis had focused primarily on the narrow "discount dealer" free-riding introduced by Lester Telser, *supra* note 10.

51 Robert T. Miller, *The End of the Road for Dr. Miles: Leegin Creative Leather Products, Inc. v. PSKS, Inc.*, 8 Engage: The Journal of the Federalist Society Practice Group 4, 40-52.

52 Justice Breyer points to a 30-year-old study that compared retail prices across states after the repeal of the Miller-Tydings Fair Trade Act, which found that retail prices were higher by 19-27%, and a statement from a Bureau of Economics Staff Report to the Federal Trade Commission stating that RPM frequently increased retail prices. While this evidence obviously is not sufficient to meet the "always or almost always anticompetitive" standard required for applying the per se rule, it suffers from an even more fundamental flaw. Specifically, the fact that both anticompetitive and pro-competitive theories of RPM predict higher prices implies that one must look at the output effects of RPM in order to make reliable inferences about its competitive impact. Justice Breyer's failure to recognize this rather pedestrian economic point is puzzling in light of his experience with antitrust arguments, his reputation as a savvy antitrust analyst, and the fact that this very point was raised in oral argument.

53 Justice Breyer offered this reminder as a circuit court judge in *Barry Wright*, noting that "unlike economics, law is an administrative system the effects of which depend on the content of rules and precedents only as they are applied by judges and juries in courts and by lawyers advising their clients. Rules that seek to embody every economic complexity and qualification may well, through the vagaries of administration, prove counter-productive, undercutting the very economic ends they seek to serve." Barry Wright Corp. v. ITT Grinnell Corp., 724 F.2d 227, 234 (1st Cir. 1983).

54 127 S. Ct. at 2724.

55 *See* Tad Lipsky & Alexi Maltas, *Leegin and the Future of Resale Price Maintenance*, 2007 eSapience Ctr. for Competition Pol'y.

56 127 S. Ct. 1955 (2007).

57 *See* A. Benjamin Spencer, Plausibility Pleading (working paper, July 30, 2007), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1003874.

58 *See* Einer Elhuage, *Twombly*—The New Supreme Court Antitrust Conspiracy Case, Volokh Conspiracy Blog (May 21, 2007), http://www.volokh.com/posts/1179785703.shtml.

59 127 S. Ct. at 1965.

60 355 U.S. 41 (1957).

61 127 S. Ct. at 1966.

62 No. 06-826, 2007 WL 2253418 (E.D. Pa. Aug. 3, 2007).

63 127 S. Ct. 1069 (2007). The author participated in this case as a signatory to the Law Professors' Amicus Brief in Support of Petitioner (filed Aug. 24, 2006).

64 Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209 (1993).

65 *Weyerhaeuser*, 127 S. Ct. at 1078.

66 *Id.* at 1076 (citing Herbert Hovenkamp, *The Law of Exclusionary Pricing*, 2 Competition Pol'y Int'l, 21, 35 (Spring 2006), and John B. Kirkwood, *Buyer Power and Exclusionary Conduct*, 72 Antitrust L.J. 625 (2005)).

67 *Id.* at 1077 (citing *Brooke Group*, 509 U.S. at 206, for the proposition that "predatory pricing schemes are rarely tried, and even more rarely successful").

68 *Id.*

69 *Id.* at 1077-78.

70 *See* Rosch, *supra* note 1 (documenting the significant experience and written output of the current justices).

71 *See* Kovacic, *supra* note 41, at 67.

72 LePage's Inc. v. 3M Co., 324 F.3d 141 (3d Cir. 2003).

73 John Roberts, *Symposium: Do We Have a Conservative Supreme Court?*, 1994 Pub. Int'l L. Rev. 104 (1994).

74 For the purposes of this essay, I do not address the earlier output of the Roberts Court in *Volvo Trucks North America, Inc. v. Reeder-Simco GMC, Inc.*, 546 U.S. 164 (2006), *Texaco Inc. v. Dagher,* 547 U.S. 1 (2006), and *Illinois Tool Works Inc. v. Independent Ink, Inc.*, 547 U.S. 28 (2006). However, I believe these 2005-06 term decisions are largely consistent with the claim advanced here.

75 *Leegin*, 127 S. Ct. at 2714-15 (citing Brief for Economists as Amici Curiae statement that "In the theoretical literature, it is essentially undisputed that minimum[resale price maintenance] can have procompetitive effects and that under a variety of market conditions it is unlikely to have anticompetitive effects").

76 *Id.* at 2715 (citing T. Overstreet, *Resale Price Maintenance: Economic Theories and Empirical Evidence* 170 (FTC 1983), and Pauline Ippolito, *Resale Price Maintenance: Empirical Evidence From Litigation*, 34 J.L. & Econ. 263, 292-93 (1991)).

77 *Id.* at 2718.

78 *See* George J. Stigler, *The Theory of Oligopoly*, 72 J. Pol. Econ. 44 (1964); *see also* Jonathan B. Baker, *Two Sherman Act Section 1 Dilemmas: Parallel Pricing, the Oligopoly Problem, and Contemporary Economic Theory*, 38 Antitrust Bull. 143, 150 (1993) ("Stigler profoundly changed the way economists understand coordination among oligopolists; and his analysis has also influenced antitrust law.").

79 *Weyerhaeuser*, 127 S. Ct. at 1077.

80 *In re Tamoxifen Antitrust Litig.*, 466 F.3d 187 (2d Cir. 2006), *cert. denied*, 127 S. Ct. 3001 (2007).

81 The Supreme Court is likely to return to the issue of identifying the appropriate measure of cost in predatory pricing cases, evidenced by the fact that it granted certiorari in *Spirit Airlines, Inc. v. Northwest Airlines Inc.*, 431 F.3d 917 (6th Cir. 2005), on this issue, but was taken off the Court's docket because it was not filed before a deadline.

82 United States v. General Dynamics Corp., 415 U.S. 486 (1974).

83 *See, e.g.*, Joshua Wright, *Von's Grocery and The Concentration-Price Relationship in Grocery Retail*, 48 UCLA L. Rev. 743, 773 (2001) ("Beyond the inherent conceptual inconsistencies of the *Von's Grocery* decision and its inability to contribute to modern enforcement of the Sherman Act, failure to overturn *Von's Grocery* results in the very danger that stare decisis and antitrust enforcement agencies have attempted to avoid—unreliability").

84 *See* David S. Evans, *Tying: The Poster Child For Antitrust Modernization* (working paper 2005), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=863031) (making the case that the per se rule should be abandoned).

85 Evans & Padilla, *supra* note 23, apply the error-cost framework to tying and reach this conclusion.

86 *See* Bruce Kobayashi, *Does Economics Provide A Reliable Guide to Regulating Commodity Bundling By Firms? A Survey of the Economic Literature*, 1 J. Competition L. & Econ. 707 (2005).

The End of the Road for *Dr. Miles:* *Leegin Creative Leather Products, Inc. v. PSKS, Inc.*

*By Robert T. Miller**

When Robert Bork published *The Antitrust Paradox* in 1978, he could argue that in antitrust law "the general movement has been away from legislative decision by Congress and toward political choice by courts, away from the ideal of competition and toward the older idea of protected status for each producer, away from concern for general welfare and toward concern for interest groups, and away from the ideal of liberty toward the ideal of enforced equality."[1] Bork was then a leader of the Chicago School of antitrust analysis, which insisted that the exclusive goal of antitrust adjudication should be the maximization of consumer welfare as determined by the most rigorous economic analysis practically available. In the fifteen years following the publication of that book, the United States Supreme Court has adopted the Chicago School arguments with a rapidity and thoroughness that has astonished even the Chicagoans themselves. By 1993, when Bork was writing a new introduction to the book, he could with complete justice speak about a "sea-change" and even a "revolution" in antitrust jurisprudence.[2] Moreover, all this happened largely without ideological rancor. Justices conventionally thought of as liberal as well as those conventionally thought of as conservative adopted Chicago School ideas in antitrust. Antitrust has become one of the least ideological areas of Supreme Court jurisprudence.

The only significant exception to the triumph of the Chicago School, Bork noted in 1993, related to resale price maintenance (RPM).[3] RPM agreements are agreements between a manufacturer and its dealers in which the dealers promise not to resale the manufacturer's products to consumers, except at certain agreed-upon prices. If dealers agree not to resale the manufacturer's products for *more than* a specified price, there is a price ceiling or *maximum* RPM. If dealers agree not to resale the manufacturer's products for *less than* a specified price, there is a price floor or *minimum* RPM. Although the Chicago School argued that RPM, whether maximum or minimum, was always or at least usually efficient, both were illegal *per se* in the pre-modern era of antitrust jurisprudence. In another victory for the Chicago School, the Supreme Court in 1997 overturned a twenty-nine year old precedent to reverse the rule related to maximum RPM.[4] Minimum RPM had been *per se* illegal under the Sherman Act since the Supreme Court's 1911 decision in *Dr. Miles Medical Co. v. John D. Park & Sons Co.*[5] Well known to generations of antitrust students, and regularly excoriated by Chicago School law professors and economists, *Dr. Miles* was the last of the really bad—from the Chicago School point of view—Supreme Court precedents. Its demise had been regularly predicted by Chicagoans at least since Frank Easterbrook did so in 1984.[6]

Last year the Supreme Court granted *certiorari* in *Leegin Creative Leather Products, Inc. v. PSKS, Inc.* on the sole question of whether minimum RPM should continue to be illegal *per se* as *Dr. Miles* held, or whether *Dr. Miles* should be overruled.[7] Everyone in the antitrust community recognized that *Leegin* would probably be one of the most important antitrust cases in a decade. Because of the nearly universal view in academia that *Dr. Miles* was wrongly decided and that RPM is at least often pro-competitive, and because of the nearly unbroken string of triumphs for Chicago School ideas in antitrust over the last thirty years, and especially because of the additions to the Court of presumptive Chicagoans Chief Justice John Roberts and Justice Samuel Alito, the antitrust community almost universally believed that *Leegin* would finally overrule *Dr. Miles*.

As the case played out, this conviction grew stronger. Ted Olson, the former Solicitor General, represented Leegin in the Supreme Court and argued that *Dr. Miles* should be overruled. The United States, as amicus curiae, took the same view and filed an amicus brief in support of Leegin. A group of very highly regarded antitrust economists, eight of whom had served as either Director of the Bureau of Economics at the Federal Trade Commission or Deputy Assistant Attorney General for Economic Analysis at the Antitrust Division of the Department of Justice (the highest-ranking economists at the respective agencies), filed amicus briefs to the same effect. Even William S. Comanor and Frederic M. Scherer, two of the most respected economists with serious reservations about RPM, filed an amicus brief arguing that some kinds of RPM should be treated under the rule of reason. The stars were aligned: it was the end of the road for *Dr. Miles*.

Speculation about the outcome of the case was largely confined to the question of whether the Court would be unanimous (generally thought unlikely), whether Justice Stevens would dissent (generally thought very likely), and whether any other justices would join Justice Stevens in dissent (opinions varied, but many people thought Justice Souter might dissent; fewer that other justices would do so). The oral argument changed little. Most observers agreed that Justice Stevens would almost certainly dissent, and the odds that one or more justices would join him seemed to rise significantly. That the Court would overrule *Dr. Miles*, however, still seemed a sure thing.

And then something very strange happened. The term wore on and on, and the Court announced no decision in *Leegin*. A few observers panicked and started saying that the Court would actually uphold *Dr. Miles*. Finally, on June 28, in the last group of decisions announced in the 2006 term, the Supreme Court issued its opinion in *Leegin*. The majority opinion by Justice Kennedy was exactly what the antitrust community expected: the ninety-six year-old precedent was overruled. Henceforth, RPM would not be illegal per se, it would be judged under the rule of reason, and so be legal or illegal in particular cases depending on whether it enhanced or impaired consumer welfare. The arguments in Justice Kennedy's

** Robert T. Miller is an Assistant Professor of Law at Villanova University School of Law.*

majority opinion were no surprise either; they were the same arguments that had appeared in Chicago School literature for upwards of forty years.[8]

The surprise, however, was that the decision was split five-to-four along the Court's conventional conservative-liberal divide: Justice Kennedy's majority opinion was joined by Chief Justice Roberts, as well as Justices Scalia, Thomas, and Alito, while Justice Breyer dissented, joined by Justices Stevens, Souter, and Ginsburg. Such things had not generally happened in antitrust cases. Many recent antitrust cases had been unanimous,[9] and even when the decisions have been split, the split has not generally been along the usual ideological lines.[10] The most curious aspect, therefore, of the *Leegin* case is why what almost everyone thought should be a non-controversial, non-ideological case in fact split the Court on ideological lines.

In Part I below, I review Justice Kennedy's majority opinion and argue that it fully reflects the broadly shared conclusions in the academic literature. In Part II, I turn to Justice Breyer's dissent and argue that it is surprisingly weak, ignoring well-known results in the economic literature that had been cited to the Court and even falling into an occasional economic fallacy. In Part III, I conclude with some general observations about the case and some speculation as to why it led to such a contentious division on the Court.

I: Majority Opinion of Justice Kennedy

The facts in *Leegin* are entirely straightforward. Leegin manufactured leather goods and distributed them throughout the United States, mostly through independently owned retail establishments, including one owned by PSKS.[11] At least on appeal, Leegin did not dispute that it had entered into minimum RPM agreements with its distributors.[12] When Leegin discovered that PSKS had been marking down its products below the agreed upon prices, Leegin stopped selling to PSKS. PSKS sued, alleging, among other things, that Leegin had violated the antitrust laws by entering into agreements with retailers to charge only those prices fixed by Leegin.[13] Leegin lost in the trial court, and the Court of Appeals for the Fifth Circuit, noting it was bound by *Dr. Miles*, affirmed.[14] The Supreme Court granted certiorari "to determine whether vertical minimum resale price maintenance agreements should be treated as *per se* unlawful."[15]

After recounting these facts, Justice Kennedy divides his majority opinion into three main parts. In the first, he summarizes the law concerning which contractual restraints should be treated as per se illegal and which should be evaluated under the rule of reason. In the second, he considers the economics of RPM and concludes that RPM agreements should be evaluated under the rule of reason. In the third, he considers the stare decisis issues related to overruling *Dr. Miles* and concludes that *Dr. Miles* should be overruled.

A. Rule of Reason Analysis and Per Se Illegality

Justice Kennedy begins his consideration of the relationship between rule of reason analysis and rules of per se illegality by noting that although Section 1 of the Sherman Act prohibits "[e]very contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States,"[16] the Court has for many decades understood Section 1 to outlaw only *unreasonable* contractual restraints on trade.[17] Following the Chicago School, antitrust analysis in the modern era understands *reasonability* in this context exclusively in terms of the challenged agreement's effect on consumer welfare.[18] As Justice Kennedy puts it, the law "distinguishes between restraints with anticompetitive effect that are harmful to the consumer and restraints stimulating competition that are in the consumer's best interest."[19] Inquiries into whether a particular restraint is anticompetitive or procompetitive are called rule of reason analyses. "Under this rule, the factfinder weighs all of the circumstances of a case in deciding whether a restrictive practice should be prohibited as imposing an unreasonable restraint on competition."[20]

But not all restraints are subject to the rule of reason. Rule of reason inquiries tend to be long, complex, and difficult, and so once courts have enough experience with a particular kind of restraint[21] to be able to "predict with confidence that [such restraints] would be invalidated in all or almost all instances under the rule of reason,"[22] the Court will declare such restraints illegal per se. The justification for having per se rules is essentially one of administrative convenience. "The *per se* rule, treating categories of restraints as necessarily illegal, eliminates the need to study the reasonability of an individual restraint in light of the real market forces at work."[23] That is, a class of restraints should be subject to *per se* condemnation only if courts know with a high degree of certainty that almost all such restraints are anti-competitive. Thus, because it is well-known that horizontal agreements among competitors to fix prices or to divide markets are almost inevitably anti-competitive, such agreements are *per se* illegal.[24] Conversely, when courts cannot be sure that a kind of restraint is almost always anti-competitive, they review such restraints under the rule of reason. Justice Kennedy thus quotes from *Khan* to the effect that the Court has "expressed reluctance to adopt *per se* rules with regard to restraints imposed in the context of business relationships where the economic impact of certain practices is not immediately obvious."[25]

Dr. Miles had made illegal per se a vertical agreement between a manufacturer and its distributors to set minimum resale prices. The primary issue in *Leegin*, therefore, is whether the conclusion of *Dr. Miles* was correct, *i.e.*, whether courts know with a high degree of certainty that minimum RPM agreements are always or almost always anticompetitive.

B. The Economics of Resale Price Maintenance

As Justice Kennedy recounts, in *Dr. Miles* the Dr. Miles Medical Co. sold its patent medicines to distributors who agreed to resell them to consumers only at set prices. In finding that the RPM agreements between Dr. Miles and its distributors were illegal per se, Justice Hughes writing for the Court (Justice Holmes dissented) relied on two separate arguments. First, Justice Hughes noted that restraints on the alienation of chattels were generally invalid at common law.[26] Second, he argued that a vertical agreement among Dr. Miles and its distributors regarding resale prices would have the same effect on retail prices as a horizontal agreement among the distributors themselves. Since such a horizontal price-fixing agreement would be illegal per se, so too must the vertical RPM agreement be illegal per se.

Justice Kennedy makes very quick work of the first argument. Noting that it relies on "'formalistic' legal doctrine" rather than "demonstrable economic effect,"[27] Justice Kennedy says, "The general restraint on alienation, especially in the age when then-Justice Hughes used the term, tended to evoke policy concerns extraneous to the question that controls here."[28] This is perfectly true, of course, and considerably gentler than Bork's pointed comment to the same effect, which was that it "is hardly reassuring to learn that the sole basis for antitrust's answer to a modern business problem is the solution given three or four hundred years ago by an English judge who was talking about something else."[29]

For all the fun that has been made of the restraint on alienation rationale in *Dr. Miles*, it is worth remembering that the free alienation of property is generally supported by sound economic reasoning. For, when one party values a property right more than the current owner, it is efficient for that party to purchase the right, and a restraint on alienation can impede this efficient transfer. There is thus slightly more economic substance to the old argument in *Dr. Miles* than might meet the eye. But not enough, of course, to support the holding in the case, for if the efficiency of the rule against restraints on alienation were a sufficient reason to prohibit RPM, it would be a sufficient reason to prohibit *all* restraints on alienation, including all manner of vertical agreements. This, of course, it clearly is not. The economic point is that splitting the bundle of property rights in unusual ways—which is what RPM contracts do—may well be efficient in particular circumstances, and ancient common law rules against splitting up bundles of property rights should not be applied blindly. *Dr. Miles* treated the rule against such restraints "formalistically" precisely because it used the rule to forestall inquiry into whether the restraint was efficient in the particular circumstances in which it was being used.

As to Justice Hughes's argument that RPM agreements have the same effect as horizontal price-fixing agreements among retailers, Justice Kennedy confines himself to the observation that the Court's cases over the last thirty years or so "formulate antitrust principles in accordance with the appreciated differences in economic effect between vertical and horizontal agreements."[30] This observation is, of course, correct. But why it is correct—and why, equivalently, Justice Hughes's argument in *Dr. Miles* was wrong—turns out to be quite important in understanding one of the disagreements between the majority and the dissent in *Leegin*.

Justice Hughes's contention that RPM agreements setting minimum resale prices would have the same effect as horizontal agreements among retailers setting the same prices was right to this limited extent: in both cases, prices paid by consumers would rise from the pre-agreement prices to the agreed-upon minimum prices. It is easy, but fallacious, to conclude that an agreement that raises prices must be anti-competitive. An inference from higher prices to an anti-competitive effect is correct only if *everything else remains the same*. This is almost always what happens in cases of horizontal price fixing. The product being sold remains what it always was and the marginal cost curves of the various cartel members remain what they always were, but the prices paid by consumers rise, and so the effect is anti-competitive. If, however, *not everything else remains the same* when prices increase, the effect may be pro-competitive or anti-competitive. If, for example, the increase in price is accompanied by an increase in quality of the product, it could well turn out that consumers will willingly pay the higher price for the higher quality product and may in the aggregate buy more of the product than before the price increase. In such cases, the change has manifestly increased consumer welfare.

When Justice Hughes argued in *Dr. Miles* that the effect of vertical RPM agreements would be the same as that of horizontal agreements setting the same price, he was thus implicitly assuming that there would be no change in the product being sold. The general argument in favor of RPM in the Chicago School has always been that the increase in retail price allows dealers to provide additional sales services to consumers such as demonstrations, explanations, attractive retail locations, large inventories, etc. If this is correct, then the product being sold—which is not just the physical object but the combination of the physical object and the sales services provided by the dealer—has in fact changed since before the RPM agreement was implemented. The inference from higher prices post-agreement to an anti-competitive effect is thus fallacious. This explanation of the fallacy in Justice Hughes's opinion has been well-known in the literature for a very long time.[31] As we shall see below, a form of the fallacy appears in Justice Breyer's dissent.

Having concluded that the reasoning in *Dr. Miles* is insufficient to support a per se rule of illegality for minimum RPM, Justice Kennedy goes on to the essential question of whether courts can know with a high degree of certainty that minimum RPM is always or almost always anti-competitive. He then states that the "economics literature is replete with pro-competitive justifications for a manufacturer's use of resale price maintenance," and backs up this assertion with an overwhelming string of citations.[32] Although certainly true, this is a much stronger assertion than Justice Kennedy needs to make in order to support the conclusion that RPM agreements ought to be reviewed under the rule of reason. The issue is not whether we know that RPM is generally pro-competitive; the issue is whether we know that RPM is always or almost always anti-competitive. Hence, if we know that RPM is generally or often pro-competitive, then *a fortiori* we do *not* know it is always or almost always anticompetitive, but it is only this much weaker claim that Justice Kennedy needs.[33] For example, if we really had no idea at all whether RPM was generally pro-competitive or anti-competitive, this should suffice to review RPM agreements under the rule of reason.

In this regard, the brief of the amici economists summarizes the state of the economics literature very aptly:

> In the theoretical literature, it is essentially undisputed that minimum RPM can have procompetitive effects and that under a variety of market conditions it is unlikely to have anticompetitive effects. The disagreement in the literature relates principally to the relative frequency with which procompetitive and anticompetitive effects are likely to ensue. The critical issue is the boundaries of that dispute. Some believe that minimum RPM is almost always benign and thus should basically be ignored by antitrust law except when it is part of a cartel case. Others believe that RPM has been demonstrated to be anticompetitive in some cases and thus merits serious antitrust consideration. *The position absent*

from the literature is that minimum RPM is most often, much less invariably, anticompetitive.[34]

In other words, there are a variety of views about the effects of minimum RPM, but everyone agrees that it is sometimes pro-competitive and sometimes anti-competitive; the disagreement is about whether minimum RPM is almost always pro-competitive or only often pro-competitive. What virtually no one believes is that minimum RPM is always or almost always anti-competitive. Moreover, even if this latter position *were* represented in the literature, as long as there remained significant disagreement among economists as to the relative frequency of pro-competitive and anti-competitive effects, it would be impossible to say that the Court *knows* that minimum RPM is always or almost always anti-competitive, and so applying a rule of per se illegality would still be wrong. Unless the economic literature were dominated by the view that minimum RPM is always or almost always anti-competitive, the rule of reason is the correct one, and, as the Economists Brief states, this view, so far from dominating the literature, is not even represented in it.

Justice Kennedy then turns to the economic literature to explain in detail why minimum RPM can be pro-competitive, and in doing so he follows the Economists Brief closely. As the economists explain, given a fixed wholesale price, manufacturers generally want retail margins to be low, for they "want[] the retailing function to be performed as efficiently as possible, with competing retailers, in turn, passing on to consumers the lowest price consistent with retailers' providing desired services and continuing business."[35] At first blush, it is thus unclear why manufacturers would in effect raise retail prices and increase dealer margins by requiring that dealers resell their products only at certain minimum prices. In real-world markets, however, "the incentives facing retailers may be out of alignment with those of manufacturers, to the detriment of the manufacturer's ability to compete effectively with the products of competing manufacturers."[36] That is, although total sales of the manufacturer's product will be maximized by having dealers offer consumers the right mix of price and sales-related services (such as size and quality of retail staff, product demonstrations and explanations, attractiveness and convenience of retail locations, dealer advertising, etc.), nevertheless market conditions may be such that individual dealers will maximize their individual profits by selling at a lower retail price and offering fewer services. In such cases, "minimum RPM can help to align these incentives [of dealers with those of manufacturers] and enhance competitiveness of a manufacturer's product, thereby benefiting consumers."[37] As Justice Kennedy puts it, "Absent vertical restraints, the retail services that enhance interbrand competition might be underprovided."[38]

Justice Kennedy provides some of the most well-known examples of the under-provision of dealer services, starting with the free-rider problem first identified by Lester Telser.[39] "Consumers might learn… about the benefits of a manufacturer's product from a retailer that invests in fine showrooms, offers product demonstrations, or hires and trains knowledgeable employees," or else decide to buy a product "because they see it in a retail establishment that has a reputation for selling high-quality merchandise."[40] As the Economists Brief points out, this phenomenon is likely to be most significant for products that are differentiated and thus sold on the basis of both features and quality as well as price—*e.g.,* complex technological products like digital cameras or fashion items like women's accessories.[41] Although the services that retailers offer in connection with such products enhance consumer welfare, they are susceptible to free-riding by discounting dealers. "If the consumer can… buy the product from a retailer that discounts because it has not spent capital providing services or developing a quality reputation, the high-service retailer will lose sales to the discounter, forcing it to cut back its services to a level lower than consumers would otherwise prefer. Minimum resale price maintenance alleviates the problem because it prevents the discounter from undercutting the service provider."[42]

As the Economists Brief also points out, however, there is some dispute in the literature about how commonly and under what circumstances RPM can eliminate or ameliorate free-riding problems.[43] In some well-known instances, the services a manufacturer would want a dealer to provide cannot easily be free-ridden. Thus, in *Adolph Coors Co. v. FTC*, Coors wanted its retailers to properly refrigerate its beer, and it is quite impossible that a dealer who let the beer assume room temperature could free-ride on the refrigeration provided by another more conscientious dealer.[44] Hence, although the literature does not suggest that eliminating or ameliorating free-riding problems is a rare or aberrational effect of RPM, combating free-riding cannot explain all instances of RPM.

Other instances can be explained, Justice Kennedy points out, by the desire to provide incentives to dealers to facilitate market entry for new firms and brands.[45] Manufacturers entering new markets "can use the restrictions in order to induce competent and aggressive retailers to make the kind of investment of capital and labor that is often required in the distribution of products unknown to the consumer."[46] "New products and new brands," he continues, "are essential to a dynamic economy, and if markets can be penetrated by using resale price maintenance there is a procompetitive effect."[47]

In a final paragraph of the section of the opinion dedicated to pro-competitive justifications for RPM, Justice Kennedy argues that RPM can also increase inter-brand competition by encouraging dealer services that would not be provided even absent free riding.[48] This is really the fundamental point about RPM, and it deserves larger treatment than Justice Kennedy affords it. As noted above, total sales of the product in question will be maximized at a certain combination of price and dealer-provided sales services, and dealers may have incentives to provide mixes of price and services that depart from the optimum. For example, if there are both consumers knowledgeable about the product who do not need sales services and will tend to select a vendor on the basis of price alone and consumers who are not knowledgeable about the product and will tend not to purchase it at all unless the dealer provides educative services about the product, then under some market conditions an individual dealer might maximize its individual profits by concentrating on attracting the knowledgeable consumers and forgoing sales to the unknowledgeable ones by cutting retail prices and reducing sales services. This might be the

case even though total sales by all dealers would be maximized if all dealers charged more and provided more services. By setting an RPM price at the optimum price and calibrating the wholesale price to dealers properly, manufacturers can align the dealers' incentives in such a way that the optimum quantity of services is elicited. Following an influential argument by Klein and Murphy, if monitoring dealer services is difficult, a manufacturer's threat to terminate a non-performing dealer and so deprive such dealer of the quasi-rents available from sales of the RPM product may be sufficient to elicit from the dealer the desired level of services.[49]

If all this is correct, then all the other pro-competitive rationales for RPM—combating free riding, facilitating market entry, maintaining large inventories, etc.—can all be seen as special cases of the general phenomenon of aligning dealers' incentives to provide the optimum mix of price and services. As the Economists Brief puts it, "With minimum RPM, retailers' choice of value-added services are determined and disciplined by market competition with other retailers. By eliminating intrabrand price competition among dealers, minimum RPM effectively shifts intrabrand competition to the non-price arena—that is, retailers compete to find the service package that best drives sales of the product."[50]

After acknowledging that there may be anti-competitive reasons for using RPM, including the familiar theories of RPM as a facilitating device in manufacturer or dealer cartels,[51] Justice Kennedy concludes that "[n]otwithstanding the risks of unlawful conduct, it cannot be stated with any degree of confidence that resale price maintenance always or almost always tends to restrict competition and decrease output. Vertical agreements establishing minimum resale prices can have either procompetitive or anticompetitive effects, depending upon the circumstances in which they are formed."[52] And agreements with such divergent properties are, of course, exactly the kind that ought to be reviewed under the rule of reason.

C. Stare Decisis and Overruling Dr. Miles

Having determined that "were the Court considering the issue as an original matter, the rule of reason, not a *per se* rule of unlawfulness, would be the appropriate standard to judge vertical price restraints,"[53] Justice Kennedy concedes that the Court does "not write on a clean slate, for the decision in *Dr. Miles* is almost a century old" and so "there is an argument for its retention on the basis of *stare decisis* alone."[54] Justice Kennedy divides his treatment of this argument into two parts. In the first, he considers the reasons in favor of overruling *Dr. Miles*, and in the second he rebuts the reasons against doing so.

As to the reasons in favor of overruling *Dr. Miles*, Justice Kennedy begins by noting that, although stare decisis can sometimes justify the retention of a rule later seen to be legally erroneous, the essential point is that stare decisis "is not as significant in this case… because the issue before us is the scope of the Sherman Act,"[55] which the Court has always treated "as a common law statute."[56] Hence, the general presumption that legislative changes should be left to Congress has less force with respect to the Sherman Act,[57] and so just "as the common law adapts to modern understanding and greater experience, so too does the Sherman Act's prohibition on 'restraint[s] of trade' evolve to meet the dynamics of present economic conditions."[58] This is correct, but slightly imprecise as applied to minimum RPM. What has changed since 1911 is not, generally speaking, economic conditions, as if RPM was always or almost always anticompetitive in 1911 but often pro-competitive in 2007. What has changed, rather, is our understanding of economic reality as embodied in the Chicago School revolution in antitrust jurisprudence.

Justice Kennedy then marshals the arguments in favor of overturning *Dr. Miles* understood as a common law precedent. After noting that the argument to this point has shown that the decision in *Dr. Miles* is wrong on the merits, and that "both the Department of Justice and the Federal Trade Commission—the antitrust enforcement agencies with the ability to assess the long-term impacts of resale price maintenance—have recommended that th[e] Court replace the per se rule with the traditional rule of reason,"[59] Justice Kennedy comes to one of the most jurisprudentially important rationales for overruling *Dr. Miles*. That is, the Supreme Court has overruled its precedents "when subsequent cases have undermined their doctrinal underpinnings."[60]

This has happened to the holding in *Dr. Miles* in several distinct ways. First, as noted above, the arguments used in *Dr. Miles* itself have been discredited by the general acceptance of Chicago School methods and conclusions in the modern era of antitrust jurisprudence.[61] Second, in *United States v. Colgate & Co.*,[62] just eight years after *Dr. Miles* was decided, the Court "reined in the decision by holding that a manufacturer can announce suggested resale prices and refuse to deal with distributors who do not follow them."[63] This rule, known as the *Colgate* doctrine, essentially allowed manufacturers to achieve the economic result that RPM would achieve if the manufacturer could do so without in fact *agreeing* with the dealers on resale prices. The reasoning here was that a contract, combination or conspiracy is a necessary element of any Section 1 violation, and so absent a vertical agreement to fix prices there would be no violation of the per se rule in *Dr. Miles.*

This end-run around *Dr. Miles* spawned a large and formalistic body of law (as well as a very active antitrust practice for the bar, with its attendant costs to manufacturers and dealers and thus to consumers) concerning whether the relationship between a manufacturer and a dealer amounted to an agreement within the meaning of Section 1.[64] For example, in 1984 in *Monsanto Co. v. Spray-Rite Service Corp.*, the Court held that antitrust plaintiffs alleging a Section 1 vertical price-fixing conspiracy must present evidence tending to exclude the possibility that the manufacturer and dealer acted independently.[65] Four years later, in *Business Electronics Corp. v. Sharp Electronics Corp.*, the Court held that the *per se* rule of *Dr. Miles* applied only to agreements over price levels, not to agreements between a manufacturer and a dealer to terminate another dealer that was selling at lower prices.[66] Hence, a manufacturer could announce the resale prices it wanted its dealers to adhere to, receive a complaint from a dealer threatening to discontinue its relationship with the manufacturer because another dealer was discounting below the manufacturer's desire price, and then agree with the complaining dealer that, if the dealer continued to sell the manufacturer's products, the manufacturer would terminate its relationship

with the discounting dealer—and all this was perfectly legal. It was also perfectly legal if both manufacturer and dealer expected that the dealer would subsequently continue to sell at the manufacturer's desired prices. But if the manufacturer and the dealer "agreed" that the dealer would sell at the prices the manufacturer desired—even if the agreement was only implied in fact, and even if it was not intended to be legally enforceable—this was per se illegal under *Dr. Miles* and could expose the manufacturer to an action for treble damages.

If the per se rule in *Dr. Miles* were really justified—if, that is, RPM agreements really were known to be always or almost always anticompetitive—then the decisions in *Monsanto* and *Business Electronics* are almost certainly unjustifiable. The text of Section 1 no doubt requires an agreement to establish a violation, but there would be no justification for the Court having made it so difficult, from an evidentiary point of view, to prove the existence of an illegal agreement.[67] Allowing this kind of evasion of *Dr. Miles*—an evasion that almost everyone has long agreed was motivated by misgivings about the economic soundness of the holding in *Dr. Miles*—"is a flawed antitrust doctrine that serves the interests of lawyers—by creating legal distinctions that operate as traps for the unwary—more than the interests of consumers—by requiring manufacturers to choose second-best options to achieve sound business objectives."[68]

There is a final way in which the doctrinal underpinnings of *Dr. Miles* have been eroded. For, minimum RPM agreements are only one kind of vertical restraint, and the Court's treatment of other kinds of vertical restraints are clearly inconsistent, from an economic point of view, with the per se rule in *Dr. Miles*. For example, agreements between a manufacturer and a dealer according the dealer the exclusive right to sell the manufacturer's products in a given geographical area (a so-called "non-price" vertical restraint) generally have economic effects similar to those of RPM agreements.[69] An exclusive dealer of the manufacturer's products can, for instance, set the resale price at whatever level it chooses free from all intra-brand competition, both price competition and service competition. This is a more extreme form of dealer protection than minimum RPM, for minimum RPM insulates dealers only from price competition. Since the Court decided *GTE Sylvania* in 1977, however, non-price vertical restraints have been reviewed under the rule of reason, not a per se rule, precisely because the Court recognized that non-price vertical restraints are often pro-competitive in their effects.[70] Hence, the holding in *GTE Sylvania* is not economically consistent with that in *Dr. Miles*.

Justice Kennedy next turns to the arguments in favor of affirming the per se rule of *Dr. Miles*.[71] The primary argument here is based on the history of congressional action related to RPM since the time of *Dr. Miles*. In 1937, Congress passed the Miller-Tydings Fair Trade Act, which made RPM agreements legal within a state if the state had authorized such agreements by a so-called "fair trade" law.[72] In the McGuire Act of 1952, Congress expanded the Miller-Tydings Act to cover agreements between manufacturers and dealers that provided that, if one dealer in the state agreed to the manufacturer's terms, all dealers in the state would be bound by them.[73] Since at different times more than thirty states had enacted appropriate legislation, under the Miller-Tydings and McGuire Acts RPM agreements were legal throughout large parts of the United States for much of the twentieth century. Generally speaking, the purpose of both acts was to allow states to protect small retail establishments against large-volume discounters,[74] which, as Justice Kennedy notes, is a rationale profoundly at odds with consumer welfare and thus utterly foreign to the Sherman Act, at least as currently understood.[75] In 1975, however, Congress repealed both acts by passing the Consumer Goods Pricing Act, and since that time the per se rule of illegality in *Dr. Miles* has governed RPM agreements in the United States.[76] The plaintiff's argument was that such congressional action ratified the per se rule of *Dr. Miles*.[77]

In response, Justice Kennedy notes that the Consumer Goods Pricing Act did *not* enact the per se rule of *Dr. Miles*, but merely rescinded what amounted to a conditional statutory exemption from it, and rescinding this exemption is not logically equivalent to endorsing the rule. Rather, "Congress once again placed these restraints within the ambit of §1 of the Sherman Act. And, as has been discussed, Congress intended §1 to give courts the ability to develop governing principles of law in the common law tradition."[78] Hence, in Justice Kennedy's view, the Court is respecting the decision of Congress "by analyzing vertical price restraints, like all restraints, in conformance with traditional §1 principles, including the principle that our antitrust doctrines evolve with new circumstances and new wisdom."[79]

Although Justice Kennedy does not say so, it is worth noting in this regard that when Congress passed the Consumer Goods Pricing Act in 1975, the Chicago School revolution in antitrust had not yet begun. That remaking of antitrust law would begin two years later in 1977 with *GTE Sylvania*. Hence, whatever the intention of Congress in 1975, it was premised on an understanding of antitrust law that has now been thoroughly discredited and very largely replaced in the law by Chicago School principles. In my view, that fatally undermines any argument based on what Congress intended in 1975.

Arguments about what conclusions to draw from congressional actions are notoriously inclusive, and the reason for this is that the matter really turns on more general matters relating to a court's treatment of legislation and legislative intent. If a judge thinks he can ascertain with reasonable certainty what a legislative body intended—understanding intention as something that goes beyond what the body literally said or did in enacting the legislation in question—and if the judge is willing to give weight to that expansive intention in interpreting other statutes, then a judge might be impressed with the argument that in enacting the Consumer Goods Pricing Act Congress ratified the per se rule of *Dr. Miles*. If, however, a judge is skeptical about divining the intention of Congress understood in this expansive way, or else thinks that such expansive intentions ought to have no weight in interpreting statutes, then he will likely be unimpressed with such arguments. This kind of disagreement about the proper role of legislative intent in judicial decision-making will not be settled, of course, in anything as mundane as an antitrust case.

Persuasive though Justice Kennedy's arguments in this section of the opinion are, he omits what I think is the most

important point, *viz.*, that the dominant theme of antitrust jurisprudence for the last thirty years has been the overturning, either expressly or by implication, of precedents from the pre-modern era. In the area of vertical restraints alone, before 1977 all three major kinds of vertical restraints—minimum RPM, maximum RPM and non-price restraints—were all illegal per se. Minimum RPM, of course, was illegal *per se* under *Dr. Miles*. Maximum RPM was illegal per se under *Albrecht v. Herald Co.*[80] Non-price vertical restraints were illegal per se under *United States v. Arnold, Schwinn & Co.*[81] In 1977, the Supreme Court overturned *Schwinn* in *GTE Sylvania.* In 1997, it overturned *Albrecht* in *Khan*. This was all part of a larger pattern in which Chicago School ideas triumphed over older theories of antitrust. Overturning *Dr. Miles* is thus just the final step in a revolution in antitrust thinking that began about thirty years ago. Unless someone were prepared to reject all of the Chicago School ideas, there is no substantial reason exempting *Dr. Miles* from the modernization of antitrust law along Chicago School lines. *Dr. Miles* is certainly old and venerable, but as the dissenting justice in *Dr. Miles* famously observed in another context, it is revolting to have no better reason for a rule of law than that it was laid down in the time of Henry IV.[82]

II: Dissenting Opinion of Justice Breyer

Justice Breyer's long dissent is premised primarily on the idea that in overturning *Dr. Miles* the Court is "depart[ing] from ordinary considerations of *stare decisis* by pointing to a set of arguments well known in the antitrust literature for close to half a century."[83] For the reasons given above and as discussed more fully below, I find Justice Breyer's arguments quite unconvincing. Before turning to his arguments about stare decisis, however, I want to discuss two preliminary points. The first concerns Justice Breyer's understanding of the relationship between the rule of reason and per se rules. The second concerns some fairly clear-cut mistakes Justice Breyer makes concerning the economic effects of RPM.

A. Justice Breyer on the Rule of Reason and Per Se Rules

The usual understanding of the relationship between rule of reason analysis and per se rules is as Justice Kennedy stated it. That is, restraints are reviewed under the rule of reason unless the Court, because of their experience with a kind of restraint, can be highly confident that the restraint always or almost always has anti-competitive effects. This, apparently, was also Justice Breyer's understanding in 1998, for he then wrote that "certain kinds of agreements will so often prove so harmful to competition and so rarely prove justified that the antitrust laws do not require proof that an agreement of that kind is, in fact, anticompetitive. An agreement of such kind is unlawful *per se*."[84] This is not the view he takes in *Leegin*.

Justice Breyer begins his discussion of per se rules by describing the law in an imprecise and potentially misleading way. After stating that courts "often" apply a rule of reason, he says, "sometimes the likely anticompetitive consequences of a particular practice are so serious and the potential justifications so few… that courts have departed from a pure 'rule of reason' approach. And sometimes this Court has imposed a rule of per se unlawfulness—a rule that instructs courts to find the practice unlawful all (or nearly all) the time."[85] While this is generally true, this description makes it sound as if there were a class of cases in which the Court has applied per se rules (*i.e.*, departed from the rule of reason) *other than* those cases in which the challenged restraint was known to be one that was always or almost always anti-competitive.[86] This is not correct, at least in the modern era. Justice Breyer's description of the law thus implicitly makes room for a new class of cases to which per se rules might be applied.

Justice Breyer explains his understanding of the relationship between rule of reason analysis and per se rules by referring to administrative concerns arising from the nature of the legal system.[87] Of course, in one clear sense, administrative concerns have always been relevant in distinguishing those restraints that should be reviewed under the rule of reason and those that should be per se illegal. That is, we have per se rules in order to save the time and expense of rule of reason inquiries when we know to a high degree of certainty that such inquiries are unnecessary because we know to a high degree of certainty that the restraint at issue is always or almost always anti-competitive. This is an administrative justification, to be sure. In order to justify a departure from the usual understanding of the relationship between the rule of reason and per se rules, however, Justice Breyer must mean something other than this.

In fact, Justice Breyer writes that "antitrust law cannot, and should not, precisely replicate economists' (sometimes conflicting) views," which in context seems to mean that the difference between per se illegality and rule of reason analysis is not simply whether courts know that, as a matter of economics, a certain kind of restraint is always or almost always anti-competitive.[88] Just how Justice Breyer would make the distinction is not yet clear. He next states that "law, unlike economics, is an administrative system the effects of which depend upon the content of rules and precedents only as they are applied by judges and juries in courts and by lawyers advising their clients."[89] So far, this is unexceptional, but it still does not distinguish a position different from the usual one. That distinction finally begins to appear when Justice Breyer says that the administrative nature of the legal system implies that courts will "sometimes apply[] rules of *per se* unlawfulness to business practices even when those practices sometimes produce benefits."[90] Now, in one sense, this statement is perfectly in accord with the Court's existing antitrust jurisprudence. For, in the usual understanding, *per se* rules are applied to restraints that are known to be always *or almost always* anticompetitive, that is, to conduct that is *sometimes but almost never* procompetitive. Thus, it is true that the Court has in the past applied *per se* rules against restraints that are "sometimes" pro-competitive—if we understand "sometimes" to be the complement of "almost always," that is, in the sense of "sometimes but almost never." In this sense, horizontal price fixing, which is *per se* illegal, is *sometimes* pro-competitive, and Justice Breyer gives precisely this example.[91]

Justice Breyer, however, seems to understand the statement about per se rules being applied to practices that are "sometimes" procompetitive in quite a different sense. He writes, "How often, for example, will the benefits [of RPM] to which the Court points occur in practice? I can find no economic consensus on this point…. Sometimes [free riding

on dealer services] must happen in reality. But does it happen often? … All this is to say that the ultimate question is not whether, but *how much*, 'free riding' of this sort takes place. And, after reading the briefs, I must answer that question with an uncertain 'sometimes.'"[92] In the usual understanding of the relationship between the rule of reason and per se rules, saying that a restraint was pro-competitive in some undetermined percentage of cases that might (for all we currently know) amount to quite often would settle the matter: such a restraint would get rule of reason treatment. For, if we do not know how often the practice is pro-competitive and how often anti-competitive, then in particular we do *not* know the practice to be always or almost always anti-competitive—and hence a per se rule would not apply. Uncertainty implies rule of reason analysis. For Justice Breyer, however, it turns out to be just the reverse: uncertainty, at least with respect to minimum RPM, results in per se illegality. Because we do not know how often minimum RPM is procompetitive, how often anticompetitive, for reasons of administrative convenience, Justice Breyer wants to apply a per se rule and make minimum RPM illegal.

It is critically important to see how Justice Breyer equivocates on the meaning of the word *sometimes*. He began with the unobjectionable point that per se rules are applied against restraints that are *sometimes* pro-competitive in the sense that such restraints are known to be *sometimes but almost never* pro-competitive, and he concludes with the point that a restraint that is *sometimes, with some unknown frequency*, pro-competitive should be subject to a per se rule as well. Justice Breyer's example of the per se rule against horizontal price fixing is very illuminating here. We know that horizontal price fixing is almost always anti-competitive and only rarely ("sometimes") pro-competitive, and so we apply a per se rule. With vertical price fixing, we *do not know* how often the practice is anti-competitive, how often it is pro-competitive. Nevertheless, in a strange updating of Justice Hughes's conflation of the effects of horizontal pricing with those of vertical price fixing, Justice Breyer would apply a per se rule against vertical price fixing just as he would against horizontal price fixing.

Justice Breyer's application of per se rules against conduct that is sometimes, with some unknown frequency, pro-competitive would invert the relationship between the rule of reason and per se rules. The traditional view embodied, as it were, a presumption of legality: conduct would be reviewed under the rule of reason unless and until courts knew it to be always or almost always anti-competitive. Justice Breyer's position reverses this and subjects to per se condemnation both conduct known to be always or almost always anti-competitive and conduct the effects of which are generally unknown. This would create, in effect, a presumption of illegality. To avoid per se combination, an antitrust defendant would have to prove that the conduct in question was known to be at least often pro-competitive. Obviously, this is not a rule that any friend of economic liberty would find attractive.

Besides casting into confusion nearly a century of jurisprudence concerning the relationship between the rule of reason and per se rules, such an expansion of the scope of per se rules is wholly unnecessary. Courts have been well able to analyze horizontal restraints under the rule of reason and make intelligent determinations about which are pro-competitive, which anticompetitive.[93] Justice Breyer recounts all the well-known and acknowledged difficulties of rule of reason analyses,[94] but in order to get to his conclusion he has to give these a weight never before accorded them. He has to say that the costs of undertaking rule of reason analyses justify applying per se rules not only when we know that the restraint in question is always or almost always anti-competitive but even when we know little or nothing about the relative frequency of the restraint's being anti-competitive or pro-competitive. If administrative considerations are to determine such cases, presumably the correct view would be that we should save *all* the potential administrative costs related to RPM cases and just declare RPM *legal per se*. Then there would be no lawsuits at all about RPM, and the administrative costs of such suits would drop to zero.

B. Justice Breyer and the Economic Effects of RPM

There are two serious economic mistakes in Justice Breyer's opinion. The first concerns his understanding of the pro-competitive justifications for minimum RPM. The second concerns the relationship between retail prices and consumer welfare.

With respect to Justice Breyer's understanding of the pro-competitive justifications for minimum RPM, throughout the dissent Justice Breyer speaks as if the only such justification were eliminating or ameliorating the problem of discounting dealers free riding on the services provided by other dealers. Thus, after referring to "the majority's claim that even absent free riding, resale price maintenance may be the most efficient way to expand the manufacturer's [market] share,"[95] he states, "I do not understand how, in the absence of free-riding (and assuming competitiveness), an established producer would need resale price maintenance. Why, on these assumptions, would a dealer not 'expand' its 'market share' as best that dealer sees fit, obtaining appropriate payment from consumers in the process? There may be an answer to this question. But I have not seen it. And I do not think that we should place significant weight upon justifications that the parties do not explain with sufficient clarity for a generalist judge to understand."[96]

The truth is that Justice Breyer had seen the answer to this question, and the answer had been explained in a way that a generalist judge could understand. The argument, recounted above, is that maximizing overall sales of a manufacturer's product requires a particular retail price level and a particular level of dealer services, and that dealers may have incentives to depart from this mix (*e.g.*, can sometimes maximize their own individual profits by cutting prices and reducing services). Minimum RPM can allow manufacturers to set the retail price at the optimum price and then calibrate the level of dealer services by adjusting the wholesale price. It is true that Justice Kennedy does not make the argument very clearly, and he certainly does not make the argument in the detailed way I have here. But the argument was made at greater length in the Economists Brief,[97] and both Justice Kennedy and the Economists Brief cite the well-known articles by Klein and Murphy[98] and by Mathewson and Winter[99] from which the argument derives.

Justice Breyer's failure to appreciate this key argument about the pro-competitive justifications for RPM may have

been a matter of inattention or oversight. His treatment of the effect of RPM on prices and its relationship to consumer welfare, however, is a matter of economic fallacy. Justice Breyer spends several paragraphs arguing that abolishing the per se rule against minimum RPM will result in higher prices to consumers, and he counts this as an argument tending to show that minimum RPM is anticompetitive.[100] He says, for example, that those "who express concern about the potential anticompetitive effects [of RPM] find empirical support in the behavior of prices before, and then after, Congress in 1975 repealed the Miller-Tydings Fair Trade Act."[101] In other words, because minimum RPM raises prices, it tends to be anti-competitive.

This is simply fallacious. Of course *minimum* RPM raises retail prices. If the dealer would sell at or above the price prescribed in the RPM agreement, there would be no need for an agreement providing for a *minimum* resale price. The point, however, is that increasing retail prices is consistent with both pro-competitive and anti-competitive theories of RPM. Both theories predict that RPM will raise retail prices—the anti-competitive theory because RPM is facilitating either a manufacturer or a dealer cartel, for example, and the pro-competitive theory because RPM is providing dealers a higher margin that they are competing away in the form of added dealers services that consumers as an aggregate value more than the aggregate increase in the price of the product. In both the anti-competitive theory and the pro-competitive theory, therefore, retail prices rise, but in the anti-competitive theory output falls and consumer welfare is impaired, while in the pro-competitive theory output increases and consumer welfare is enhanced. To infer an anti-competitive effect from increased prices is simply erroneous. The error is analogous to the infamous one of inferring a price-fixing conspiracy from parallel pricing behavior. In both cases, there are plausible anti-competitive and pro-competitive explanations for the phenomenon, and the mistake, notorious in the history of antitrust, is merely to assume that the anti-competitive explanation is the correct one.

C. Justice Breyer and Stare Decisis

Justice Breyer writes that, abstracting from concerns of stare decisis, the question of whether minimum RPM agreements should be reviewed under the rule of reason or should be per se illegal is a difficult problem and "if forced to decide now, at most I might agree that the per se rule should be slightly modified to allow an exception for the more easily identifiable and temporary condition of 'new entry.'"[102] In saying this, he specifically refers to administrative concerns and so must have in mind his view that restraints whose economic effects are largely unknown or difficult to determine should be treated as per se illegal. But, Justice Breyer continues, "The question before us is not what should be the rule, starting from scratch. We here must decide whether to change a clear and simple price-related antitrust rule that the courts have applied for nearly a century."[103] He provides many arguments tending to show that the rule of *Dr. Miles* ought be affirmed, and the tone becomes heated. Justice Breyer is "not aware of any case in which [the] Court has overturned so well-established a statutory precedent" and does "not see how the Court can claim that ordinary criteria for overruling an earlier case have been met."[104]

Many of the arguments that Justice Breyer adduces follow the same pattern. He will argue that there is a certain kind justification that, if present, would counsel in favor of overturning a case, but that there is no such justification in connection with *Dr. Miles*; he then either states or implies that the conclusion is that *Dr. Miles* ought not to be overturned. For example, he argues that if there had been an important change in the American economy that made the effects of RPM other than what they once were, then overturning *Dr. Miles* might be justified, but that there has been no such change in the economy, and so *Dr. Miles* ought not to be overruled.[105] Arguments of this form, of course, embody the logical fallacy of denying the antecedent—the fallacy that argues from "If P, then Q" and "Not P" to "Not Q"—and so have no force at all. The fact that one potential argument for a conclusion fails does nothing to show that there is no good argument for the conclusion. (Compare the defense attorney who argues that his client ought not to be convicted because there are several people whom he did *not* murder.) For instance, to pursue Justice Breyer's argument, it is perfectly true that there has been no change in the American economy that makes the effects of RPM other than what they always were. In fact, in 1911 as in 2007, those effects are probably very largely pro-competitive, and the diminution in consumer welfare that results from making RPM illegal per se is an excellent reason for scrapping that rule.

The most important of Justice Breyer's arguments to follow this pattern of denying the antecedent is a generalized one that there has been no change—no change in economic circumstances, no change in our understanding of the economics of RPM, no other relevant kind of change—that would justify overturning *Dr. Miles*; hence, *Dr. Miles* ought to stay. In this particular case, however, the argument not only is logically fallacious but also suffers from the fact that its minor premise—that there has been no relevant change—is not only false but obviously and flagrantly so. For, as I noted above, between 1978 and 1993, there was, as Robert Bork put it, a "sea-change" and even a "revolution" in antitrust jurisprudence. One can argue about whether this change has been good or bad, but nobody denies that it occurred, and nobody denies that it amounted to a rather thorough implementation of Chicago School ideas, mostly through the overturning of cases that had established per se rules against conduct that the Chicagoans had argued were in fact at least often pro-competitive. In recent years, many observers have thought that days of the remaining per se rules, including that of *Dr. Miles*, were numbered. For example, many people thought that the per se rule against tying[106] would be overturned if the *Microsoft* case[107] ever reached the Supreme Court.[108] Thus, when Justice Breyer says that there has been no change in our understanding of antitrust economics that would justify overruling *Dr. Miles*, what he means is that there has been a revolutionary change in our understanding of antitrust economics over the last forty years that abundantly justifies overruling *Dr. Miles*, but that the Court did not get around to actually overruling *Dr. Miles* till now; in part, of course, because it was busy overruling other bad antitrust precedents. Some of these, such as *Khan*, which overruled the twenty-nine year old precedent in *Albrecht* that *maximum* RPM was illegal per se, Justice Breyer himself joined.

In a not especially subtle maneuver, Justice Breyer then turns to the factors related to overturning precedents that Justice Scalia discussed in his concurring opinion in *Federal Election Commission v. Wisconsin Right to Life, Inc.*[109] and argues that they imply that *Dr. Miles* not be overruled.[110] Some the arguments Justice Breyer adduces here follow the fallacious pattern identified above. For example, Justice Breyer argues that if a case is constitutional and recently decided, it may more readily be overruled; but since *Dr. Miles* is not constitutional and not recently decided, it ought not to be overruled.[111] Or again, if a case creates an unworkable legal regime, this counsels in favor of overruling it, but since *Dr. Miles* did not create an unworkable regime, it ought not to be overruled.[112] Or yet again, if a case unsettles the law, this argues in favor of overruling it, but *Dr. Miles* did not unsettle the law; hence, it ought not to be overruled.[113] In each case we have the same fallacy of denying the antecedent.

Apart from arguments like these and the argument, discussed above, based on the intent of Congress in repealing the Miller-Tydings and McGuire Acts, Justice Breyer's primary argument is that businesses and consumers have come to rely on the per se rule of Dr. Miles and such reliance counsels against overruling the case. It is difficult to see, however, just what this reliance consists in. We are talking here of making legal conduct that in the past was illegal under the per se. This is not a case in which people could have counted on certain conduct being legal only to discover that the conduct is now illegal, thus frustrating their plans.

The reliance possible in this case would be of quite a different character. It would be reliance, probably by a dealer, that a manufacturer will not require RPM agreements of its dealers. We are talking about the reliance interest, that is, of discounting dealers. For example, Justice Breyer cites the brief of Burlington Coat Factory Warehouse Corp. as amicus curiae to the effect that it and similar businesses have financed, structured, and operated their businesses in part in reliance on the absence of minimum RPM.[114] This is very likely true. But whether or not this reliance interest ought weigh in the Court's determination as to whether it should overturn *Dr. Miles* depends, I think, on whether the conduct in question enhances consumer welfare or not. For, if the conduct in question impairs consumer welfare, the fact that certain businesses have relied on an inefficient rule of law in order to perpetuate inefficient conduct can hardly be a reason to keep in place a rule of antitrust law, for the exclusive goal of antitrust is efficiency. Now if the pro-competitive theories of minimum RPM are generally correct, then the conduct undertaken in reliance on the per se rule of *Dr. Miles* usually impairs consumer welfare; if the anti-competitive theories are generally correct, then it usually enhances consumer welfare. Hence, whether or not this reliance interest should be given weight in the decision to overturn *Dr. Miles* cannot be settled until we know whether minimum RPM is generally pro-competitive or generally anti-competitive. Probably, RPM is often pro-competitive. In any case, before this reliance interest could be a major factor in determining whether *Dr. Miles* ought be overruled, we would have to have good reason to believe that minimum RPM is much more often than not anti-competitive, and, as noted above, this is the one position that is absent from the economic literature.

III: Observations and Conclusions

So after ninety-six years, it is the end of the road for *Dr. Miles*. The Chicago School of antitrust analysis has won the last majority victory there is to win, and there is no likelihood that Chicago School ideas will be seriously challenged in the Supreme Court's antitrust jurisprudence in the foreseeable future. This is all as most observers think it should be and as virtually all observers thought it would be. The only mystery left is why the remaking of the Court's antitrust jurisprudence in accordance with Chicago School ideas—a project that has been going on for three decades and has enjoyed unusually broad support among Supreme Court justices of all ideological persuasions—should, in this last major decision, suddenly produce a five-to-four split along conventionally conservative-liberal ideological lines. This is a regrettable blemish on an otherwise very credible episode in the Court's history.

As I explained above, the argument for overruling *Dr. Miles* was extremely strong. The economic literature teems with pro-competitive theories of minimum RPM, and there is virtually no support any more for the idea that RPM is always or almost always anti-competitive. Since the Court by default applies rule of reason analysis and employs per se rules only when a practice is known to be always or almost always anticompetitive, it is elementary that minimum RPM should be reviewed under the rule of reason. True, *Dr. Miles* was a venerable statutory precedent, but the antitrust laws are common law statutes, and the whole history of antitrust in the modern era has consisted in the Court's overturning antiquated rulings the justifications for which had been demolished by Chicago School analysis. Almost everyone assumed that, sooner or later, *Dr. Miles* would go the way of any number of other bad decisions that had already been overturned.

Making the best of a bad business, Justice Breyer is forced in dissent to adopt positions very difficult to defend. He says that per se rules should be applied not only to conduct known to be always or almost always anti-competitive but also to conduct the competitive effects of which we do not yet know with certainty. He admits he cannot find, or else did not understand, a key economic argument that had been cited to the Court. He argues that price increases from RPM are anti-competitive when in fact it is well known that both pro-competitive and anti-competitive explanations of RPM entail that prices will rise. He premises his dissent largely on stare decisis considerations and produces a great many arguments along such lines, but most of them are infected by an elementary formal fallacy. He finds himself in the very odd position of arguing that there has been no significant change in an area of law famous for having undergone an intellectual revolution in the last thirty years. How could this have happened?

Now, unless a justice happens to tell us that he or she had reasons other than those stated in the opinion for voting one way or another in a particular case, we have to take the stated reasons at face value. Going beyond them and searching for other reasons is pure speculation. Such speculation is, however, human, and the unexpected ideological split on the Court, along with the timing of the decision and other events

in the 2006 term, suggest an explanation that would appeal to someone inclined to speculation. That explanation is that the liberals on the Court—Justices Stevens, Souter, Ginsburg, and Breyer—have felt, and disliked, the power of the new conservative majority of Chief Justice Roberts and Justices Scalia, Kennedy, Thomas, and Alito. Such five-to-four splits have marked the most acrimonious cases this terms, including *Parents Involved in Community Schools v. Seattle School District No. 1*[115] on affirmative action, *Federal Election Commission v. Wisconsin Right to Life, Inc.*[116] on campaign finance, and *Gonzalez v. Carhart* on abortion.[117] All of these cases either overruled prior decisions or at least significantly undercut them. It is easy to imagine that the liberal justices saw *Leegin* as being more of the same.

If this correct, then, in dissenting so vigorously in *Leegin*, the liberals were emphasizing the ideological division of the Court and were attempting to persuade their conservative colleagues not—in the liberals' view—to overplay an admittedly strong position. By politicizing a case that by rights ought not to have been very political, the liberals may have been warning—some might say threatening—about possible politicization of the Court in the future. I offer no opinion on the merits or demerits of such a maneuver, for one's view of it will be determined very largely along the same ideological lines that divide the Court.

However that may be, *Dr. Miles* is overruled, a significant area of the law has been rationalized, and American consumers are likely to benefit. For results like those, a little more controversy about the Supreme Court is a price worth paying.

Endnotes

1 Robert H. Bork, The Antitrust Paradox 10-11 (1978) (*hereinafter* The Antitrust Paradox).

2 *Id.* at xi, x.

3 *Id.* at xiii.

4 State Oil Co., v. Khan, 522 U.S. 3 (1997), overruling *Albrecht v. Herald, Co.*, 390 U.S. 145 (1968).

5 220 U.S. 373 (1911).

6 Frank Easterbrook, "Vertical Arrangements and the Rule of Reason," mimeo, Law School, University of Chicago ("the days of Dr. Miles are numbered").

7 551 U.S. ___ (2007). Citations to *Leegin* are to the pagination used in the Supreme Court's slip opinion.

8 In the final examination in my antitrust class at Villanova Law School in the spring of 2007, I asked my students to write the majority opinion in the then-pending *Leegin* case; the answers were generally close student approximations to what Justice Kennedy in fact later said.

9 *E.g.*, Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber Co., Inc., 127 S.Ct. 1069 (2007) (unanimous opinion by Justice Thomas holding that plaintiff in predatory-bidding claim under § 2 of Sherman Act must prove that predatory bidding led to below-cost pricing of predator's outputs and that predator had dangerous probability of recouping losses through exercise of monopsony power); Illinois Tool Works, Inc. v. Independent Ink, Inc., 547 U.S. 28 (2006) (unanimous opinion, except for Justice Alito who took no part, by Justice Stevens holding that patent does not necessarily confer market power); Verizon Communications, Inc. v. Law Offices of Curtis V. Trinko, LLP, 540 U.S. 398 (2004) (opinion by Justice Scalia, joined by Chief Justice Rehnquist and Justices O'Connor, Kennedy, Ginsburg and Breyer, and concurring opinion by Justice Steven joined by Justices Souter and Thomas, that telephone company had no duty under Section 2 of Sherman Act to share network facilities with competitors); Nynex Corp. v. Discon, Inc., 525 U.S. 128 (1998) (unanimous opinion by Justice Breyer that per se rule against group boycotts does not apply to decision by single purchaser to buy from one seller rather than another even when such decision is not justified by usual competitive reasons); State Oil Co., v. Khan, 522 U.S. 3 (1997) (unanimous opinion by Justice O'Connor overruling *Albrecht v. Herald, Co.*, 390 U.S. 145 (1968), to hold that maximum resale price maintenance is not illegal *per se* but should be judged under rule of reason).

10 *E.g.*, Bell Atlantic Corp., v. Twombly, 127 S.Ct. 1955 (2007) (majority opinion by Justice Souter, joined by Chief Justice Roberts and Justices Scalia, Kennedy, Thomas, Breyer and Alito; dissent by Stevens, joined in part by Justice Ginsburg); California Dental Assoc. v. FTC, 526 U.S. 756 (1999) (majority opinion by Justice Souter, joined by Chief Justice Rehnquist and Justices O'Connor, Scalia and Thomas; opinion by Justice Breyer concurring in part and dissenting in part, joined by Justices Stevens, Kennedy and Ginsburg).

11 *Id.* at 2.

12 *Id.* at 4.

13 *Id.* at 4.

14 171 Fed. Appx. 464 (2006) (*per curiam*).

15 *Leegin* at 5.

16 15 U.S.C. § 1.

17 *Leegin* at 5.

18 *E.g.*, The Antitrust Paradox¸ *supra* note 2, at xi.

19 *Leegin* at 5-6.

20 *Id.* at 5 (quoting Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36, 49 (1977)).

21 *See* Broadcast Music, Inc, v. Columbia Broadcasting System, Inc., 441 U.S. 1, 9 (1979) (referring to considerable experience with kind of restraint needed before subjecting it to per se treatment).

22 *Leegin* at 6. *Cf.* Arizona v. Maricopa County Medical Soc., 457 U.S. 332, 344 (1982).

23 *Id.*

24 *Id.*, referring to *Texaco, Inc. v. Dagher*, 547 U.S. 1, 5 (2006) and *Palmer v. BRG of Ga., Inc.*, 498 U.S. 46, 49-50 (1990) (*per curiam*).

25 *Id.* at 6-7 (quoting from State Oil Co. v. Khan, 522 U.S. 3, 10 (1997)).

26 *Dr. Miles*, 220 U.S. 373, 404-405 (1911).

27 *Leegin* at 7 (quoting Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36, 58-59 (1977)).

28 *Id.* at 8.

29 The Antitrust Paradox, *supra* note 2, 285.

30 *Leegin* at 8.

31 *E.g.*, The Antitrust Paradox, *supra* note 2, 289, 296-297.

32 The authorities Justice Kennedy cites include the economists appearing at *amici*, Brief for Economists as *Amici Curiae* 16; the ABA Section of Antitrust Law, Antitrust Law and Economics of Product Distribution 76 (2006); H. Hovenkamp, The Antitrust Enterprise: Principle and Execution 184-191 (2005); Bork, The Antitrust Paradox, *supra* note 2, 288-291; and F.M. Scherer & D. Ross, Industrial Market Structure and Economic Performance 558 (3d ed., 1990).

33 *See Leegin* at 14 ("Notwithstanding the risks of unlawful conduct, it cannot be stated with any degree of confidence that resale price maintenance always or almost always tend[s] to restrict competition and decrease output") (citations omitted).

34 Brief for Economists as *Amici Curiae* 16 (citations omitted and emphasis added) (*hereinafter* Brief for Economists).

35 Brief for Economists 5.

36 *Id.*

37 *Id.*

38 *Leegin* at 10.

39 Lester G. Telser, *Why Should Manufacturers Want Fair Trade?*, 3. J.L. & Econ. 86 (1960).

40 *Leegin* at 11.

41 Economists Brief 6.

42 *Leegin* at 11.

43 Economists Brief 5.

44 497 F.2d 1178 (10th Cir. 1974), *cert. denied* 419 U.S. 1105 (1975).

45 *Leegin* at 11.

46 *Id.* at 55 (quoting from *GTE Sylvania*).

47 *Id.* at 11-12.

48 *Id.* at 12.

49 Benjamin Klein & Kevin M. Murphy, *Vertical Restraints as Contractual Enforcement Mechanisms*, 31 J.L. & Econ. 263 (1988).

50 Economists Brief 10-11 (citations omitted).

51 *Leegin* at 12-14.

52 *Id.* at 14.

53 *Id.* at 19.

54 *Id.*

55 *Id.*

56 *Id.* at 20.

57 *Id* at 19-20 (citing State Oil Co. v. Khan, 522 U.S. 3, 20 (1997)).

58 *Id.* at 20.

59 *Id* . at 20-21.

60 *Id.* at 21 (quoting Dickerson v. United States, 530 U.S. 428, 443 (2000)).

61 *Id.*

62 United States v. Colgate & Co., 250 U.S. 300 (1919)

63 *Leegin* at 21, referring to United States v. Colgate & Co., 250 U.S. 300 (1919).

64 *See, e.g.*, United States v. Parke, Davis & Co., 362 U.S. 29 (1960) (finding manufacturer defendant had entered into RPM agreement); Monsanto Co. v. Spray-Rite Service Corp., 465 U.S. 752 (1984) (holding that antitrust plaintiffs alleging Section 1 price-fixing conspiracy must present evidence tending to exclude possibility that manufacturer and distributor acted independently); Business Electronics Corp. v. Sharp Electronics Corp., 485 U.S. 717 (1988) (holding per se rule of *Dr. Miles* applied only to agreements over price levels, not to agreement between manufacturer and dealer to terminate another dealer selling at lower prices).

65 465 U.S. 752 (1984).

66 485 U.S. 717 (1988).

67 *Leegin* at 21-24.

68 *Id.* at 25.

69 *Id.* at 24.

70 Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36 (1977).

71 *Leegin* at 25-28.

72 89 Stat. 801.

73 66 Stat. 632.

74 California Retail Liquor Dealers Assn. v. Midcal Aluminum, Inc., 445 U.S. 97, 102 (1980); *Leegin* at 26.

75 *Leegin* at 26.

76 89 Stat. 801.

77 *Leegin* at 25.

78 *Id.* at 25-26 (internal quotation marks and citations omitted).

79 *Id.* at 26 (internal quotation marks and citations omitted).

80 Albrecht v. Herald, Co., 390 U.S. 145 (1968).

81 388 U.S. 365 (1967).

82 *See* Oliver Wendell Holmes, *The Path of the Law*, 10 Harv. L. Rev. 457, 469 (1897).

83 *Leegin*, Dissenting Opinion of Justice Breyer 1.

84 Nynex Corp. v. Discon, Inc., 525 U.S. 128, 133 (1998) (citations omitted).

85 *Leegin*, Dissenting Opinion of Justice Breyer 2.

86 There is, of course, the "quick-look" analysis that the Court has sometimes used, but this is really just an abbreviated rule of reason analysis, not a third class of review, and its existence does not help to make sense of what Justice Breyer says here. *See* California Dental Assoc. v. FTC, 526 U.S. 756 (1999) (discussing quick-look analysis).

87 *Leegin*, Dissenting Opinion of Justice Breyer 3.

88 *Id.* at 8.

89 *Id.*

90 *Id.*

91 *Id.* (citing F.M. Scherer & D. Ross, Industrial Market Structure and Economic Performance 335-339 (3d ed. 1990)).

92 *Id.* at 8-9.

93 *E.g.,* Broadcast Music Inc. v. CBS, Inc., 441 U.S. 1 (1979); Arizona v. Maricopa County Medical Soc., 457 U.S. 332, 342-343 (1982); Nat'l Soc. of Prof. Engineers v. United States, 435 U.S. 679 (1978); Nat'l Collegiate Athletic Association v. Bd. of Regents of the University. of Oklahoma, 468 U.S. 85 (1984); California Dental Assoc. v. FTC, 526 U.S. 756 (1999).

94 *Leegin*, Dissenting Opinion of Justice Breyer 9-11.

95 *Id.* at 14 (internal quotation marks and citations omitted).

96 *Id.* at 14-15.

97 Economists Brief 10-11.

98 Benjamin Klein & Kevin M. Murphy, *Vertical Restraints as Contractual Enforcement Mechanisms*, 31 J.L. & Econ. 263 (1988).

99 G. Franklin Mathewson & Ralph A. Winter, *The Law and Economics of Resale Price Maintenance*, 13 Rev. Indus. Org. 57, 74-75 (1998).

100 *Leegin*, Dissenting Opinion of Justice Breyer 5-6.

101 *Id.* at 5.

102 *Id.*at 11.

103 *Id.*

104 *Id.* at 12.

105 *Id.* at 15.

106 Jefferson Parish Hospital District No. 2 v. Hyde, 466 U.S. 2 (1984).

107 United States v. Microsoft Corp., 253 F.3d 34 (D.C. Cir. 2001), *cert. denied* 534 U.S. 952 (2001).

108 In fact, the Supreme Court denied certiorari, 534 U.S. 952, and affirmed a remarkable decision of the Court of Appeals for the District of Columbia Circuit that acknowledged the per se rule against tying but went on to hold that the per se rule would not apply to certain kinds of software products. United States v. Microsoft Corp., 253 F.3d 34, 89-95 (D.C. Cir. 2001).

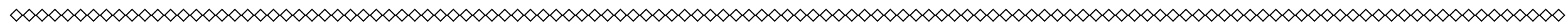
109 Federal Election Commission v. Wisconsin Right to Life, Inc., 551 U.S. ___ (2007).

110 *Leegin*, Dissenting Opinion of Justice Breyer 17-23.

111 *Id.* at 17.

112 *Id.* at 17-18.

113 *Id.* at 18.



114 *Id.* at 18. Justice Breyer goes on to speculate about what might be called consequential reliance, asking, "What about malls built on the assumption that a discount distributor will remain an anchor tenant? What about home buyers who have taken a home's distance from such a mall into account?" *Id.* at 19. The spectacle of malls closing and home prices plummeting at a result of overturning *Dr. Miles* is, in my view, utterly fanciful.

115 Parents Involved in Community Schools v. Seattle School District No. 1, 551 U.S. ___ (2007).

116 Federal Election Commission v. Wisconsin Right to Life, Inc., 551 U.S. ___ (2007).

117 Gonzalez v. Carhart, 551 U.S. ___ (2007).

Majority Voting: Acting in the Best Interest of Stockholders, or Following the Herd?

*By Donald R. Moody & J. Ammon Smartt**

In 2003 the Securities and Exchange Commission proposed a rule that would, under certain circumstances, allow stockholders greater access to proxy materials in director elections.[1] The proposal was supported by many institutional investors, but met with great opposition in the issuer community. To date, the SEC has not adopted the proposed rule.[2] Stockholder activists, however, continue to push for reform in director elections.

Historically, directors have almost always been elected under a plurality system. Since 2004, however, there has been a dramatic shift toward majority voting. As of February 2007, over 52% of S&P 500 companies, and over 45% of companies in the Fortune 500, adopted some variation of majority voting in director elections.[3] Feeding this trend, Delaware amended Sections 141 and 216 of its corporation law to address certain issues that arise in the context of majority voting.[4] Other states also have taken pro-active measures to ensure that applicable state laws do not conflict with the adoption of a majority voting standard.[5]

There are different variations of the majority voting standards adopted by companies, but, as a general rule, the principal difference between a majority voting standard and a plurality standard is the fact that a majority of votes are required to elect a director under a majority voting standard. This differs from a plurality standard in that, under a plurality standard, a director can be elected by any number of votes so long as the director receives more votes than any other director—even if the margin is a single vote.

This development has left management and the boards of directors at many companies wondering if they should adopt a majority voting standard, and, if so, what form to adopt. The answer to that question can be complicated, and will vary depending on the company and its stockholder constituency. This article addresses some of the more popular formulations of majority voting standards that corporations have adopted, the changes in Delaware law that have facilitated the adoption of majority voting, and some of the arguments for and against adopting a majority voting standard in director elections.

I. Majority Voting Standards

Companies have adopted various approaches in formulating a majority voting standard. Some have adopted policies,[6] while others have adopted bylaw amendments.[7] A few companies have amended their charters to provide for majority voting.[8] To date, variations of two majority voting models have emerged as the preferred approach by companies in addressing this issue. The majority voting systems adopted by Pfizer, Inc. and Intel Corporation are the best examples of the two models that have emerged.

**Donald R. Moody is a Partner and J. Ammon Smartt is an Associate in the Nashville, Tennessee office of Waller Lansden Dortch & Davis, LLP. Both men practice in the areas of mergers and acquisitions, securities law, and general corporation law, including corporate governance matters.*

A. Pfizer

Pfizer, Inc. was one of the early adopters of a *policy* to provide for majority voting.[9] Pertinent portions of the policy are as follows:

> In an uncontested election, any nominee for Director who *receives a greater number of votes "withheld" from his or her election than votes "for"* such election (a "Majority Withheld Vote") shall promptly tender his or her resignation following certification of the shareholder vote.[10]

The Pfizer approach has come to be known as the "plurality-plus" or "modified plurality" model. This model proved attractive to many companies because adopting a policy, which can be easily amended by the board of directors, provides greater flexibility. The Pfizer model also proved attractive to companies because a director must resign only if the director receives more "withheld" votes than "for" votes. A "true" majority voting standard would require a director to be elected by an affirmative majority of all stockholders or votes cast.[11]

More recently, the attractiveness of the Pfizer model has diminished. Institutional Shareholder Services, which advises institutional holders on how to vote in shareholder votes, originally indicated that while it generally supported a move by companies to a true majority voting standard, it would also consider "meaningful and equivalent alternative[s]" to a true majority voting standard.[12] Under its guideline, ISS "looked for a policy that included at a minimum: articulation of the decision-making process, prompt disclosure, and independent director involvement."[13] Much to the dismay of many companies trying to appease their stockholder constituencies, only one company in 2006 met the ISS standards for an "alternative structure" to true majority voting.[14] Ultimately, ISS abandoned its original intent to accept "meaningful and equivalent alternative[s]" with a declaration that ISS would only consider supporting "true majority voting standard[s] that... include a majority default rule and modification of the Holdover Director Rule…."[15]

With this backdrop, many companies considered alternative models that went beyond the modified plurality exemplified by the Pfizer approach. Intel Corporation is one of the leading examples of a company that adopted an alternative to Pfizer's model by setting an early standard for what constitutes a true majority voting model.

B. Intel

Intel Corporation adopted a majority vote bylaw on January 19, 2006, and further amended it on January 18, 2007. Intel's bylaw currently reads, in pertinent part, as follows:

> [E]ach director shall be elected by the vote of the *majority of the votes cast* with respect to the director at any meeting for the election of directors at which a quorum is present, provided that if as of a date that is fourteen (14) days in advance of the date the corporation files its definitive proxy statement (regardless of whether or not thereafter revised or supplemented) with the Securities and Exchange Commission the number of nominees exceeds the number of directors to be elected, the directors shall

> be elected by the vote of a plurality of the shares represented in person or by proxy at any such meeting and entitled to vote on the election of directors. For purposes of this Section, a majority of the votes cast means that the number of shares voted "for" a director must exceed the number of votes cast "against" that director.[16]

The Intel model is considered a true majority voting standard because of the language requiring a "majority of the votes cast" to elect a director. Additionally, Intel adopted this model as a bylaw, as opposed to a policy, which makes it somewhat more difficult for the Intel Board of Directors to modify for the reasons discussed below.

During 2006, there was a significant increase in the number of companies choosing to adopt the Intel model over the Pfizer model. While difficult to predict, this trend towards true majority voting will likely continue. Notably, some companies are beginning to adopt bylaw provisions that explicitly provide that only the stockholders can amend the majority voting bylaw.[17] A smaller number of companies are addressing the concept of majority voting in their charters.[18]

C. Policy vs. Bylaw vs. Charter

If the ultimate goal of majority voting reform is to provide the stockholder with a greater voice in the direction of the companies in which they invest, it makes sense that stockholder activists would prefer the adoption of majority voting standards that are not easily changed by a board of directors. Most companies have adopted such standards through policies,[19] but an increasing number of companies are adopting bylaw amendments to incorporate majority voting in director elections.[20] As a general rule, bylaws can also be changed without stockholder approval. For public companies, however, any amendment to the bylaws must be disclosed to stockholders in a filing on SEC Form 8-K.[21] Therefore, a majority voting standard adopted by a public company pursuant to a bylaw amendment provides transparency, even though it may be modified in the future without stockholder approval. This transparency should provide a greater sense of stability to stockholders because a board of directors will think twice about amending its bylaws if the change must be made public. Transparency, however, does not translate into control by stockholders.

Some companies, despite the fact that state law allows boards of directors to amend bylaws without stockholder approval, are adopting bylaw provisions that provide that the majority-vote bylaw can not be changed without stockholder approval. Delaware recently amended Section 216 of the Delaware General Corporation Law to provide that "[a] bylaw amendment adopted by stockholders which specifies the votes that shall be necessary for the election of directors shall not be further amended or repealed by the board of directors."[22] Obviously, either of these two scenarios gives the stockholders the ultimate say in when or how a majority voting bylaw should be amended or repealed.

A smaller number of companies are enshrining majority voting in their charters. Charter amendments generally require stockholder approval to be adopted.[23] ISS perceives the adoption of majority voting through a charter amendment to be an important component in achieving their "gold standard."[24]

Given that this trend toward majority voting is a recent phenomenon, a question that companies must grapple with when considering the adoption of a majority voting standard is whether to adopt a policy, bylaw, or charter amendment. There are many arguments for and against adopting a policy as opposed to a bylaw or a charter amendment. The ultimate determination of what is best for the company will depend upon a factual analysis after considering the totality of the circumstances surrounding the perceived need to adopt a majority voting policy.

For some companies, the flexibility of a policy may be more appropriate. For example, if a company is uncertain about the application of particular state laws, or if the company anticipates dealing with "empty voting" as described below, it may need the flexibility of a policy to act quickly in order to protect the overall interests of its stockholders. For other companies, particularly those that are targeted by stockholder activists, a charter or bylaw amendment might be the best fit in keeping with a company's corporate governance goals and stockholder constituency. For many companies, retention of the plurality standard in director elections may be the best alternative.

Undoubtedly, some companies have adopted majority voting standards to "follow the herd." Such a mentality on the part of management is rarely beneficial to stockholders. Where possible, companies should give the market time to develop the right approach. While ISS may recommend true majority voting coupled with a charter amendment, there are many questions as to how such a standard may affect the governance of that company.

II. Concerns with Majority Voting

With any major shift in corporate governance, there will always be uncertainty with respect to how those changes will affect the overall health of companies competing in the marketplace. The majority voting trend is no exception to that rule. Some examples that have already been identified include the holdover problem, compelled resignation of directors, broker non-votes, and general policy considerations as to the need for majority voting in the first place.

A. The Holdover Problem

Over 90% of the majority voting standards adopted by companies include a carve-out for contested elections.[25] Such a carve-out typically provides that in the event of a contested election the traditional plurality standard will apply.

The dynamics change in the majority voting paradigm for uncontested elections. Most state laws provide that a director "holds over" and remains in office until a "successor is elected and qualified."[26] But what happens in an *uncontested* election in the case of an incumbent director who receives more "withheld" votes than "for" votes, or who does not receive a majority of votes cast? Under the holdover rule, the director continues to stay in office until a successor is elected and qualified. Of course, the board of directors or stockholders can take action to call a special meeting to replace the director, but at what cost in terms of money and time? Moreover, how should the board of directors proceed if the director refuses to resign?

B. Director Resignation

Most companies have addressed the holdover issue by adopting policies or bylaws that indicate what action the board of directors and the incumbent director must take in the event of a failed election. Intel Corporation, for example, adopted in its January 19, 2006, bylaw amendment the following:

> If a director is not elected, the director shall offer to tender his or her resignation to the Board. The Corporate Governance and Nominating Committee will make a recommendation to the Board on whether to accept or reject the resignation, or whether other action should be taken. The Board will act on the Committee's recommendation and publicly disclose its decision and the rationale behind it within 90 days from the date of the certification of the election results. The director who tenders his or her resignation will not participate in the Board's decision….[27]

Intel, similar to many other companies, requires their directors to resign in the event of a failed election. In the Intel example, the board of directors also retains some discretion as to whether or not they will accept the mandatory resignation of the director. In 2007, Intel modified its bylaws and adopted a policy to require advance director resignations that become effective upon a failed election.[28]

Mandatory director resignations, however, have been the cause of some concern. Many have raised questions as to whether the requirement that a director resign is tantamount to director removal. Under the laws of many states, only stockholders can remove directors and then sometimes only for cause.[29] Delaware addressed this issue by amending Section 141 of the Delaware General Corporation Law. Section 141 now provides, in pertinent part, the following:

> A resignation is effective when the resignation is delivered unless the resignation specifies a later effective date or an effective date determined upon the happening of an event or events. A resignation which is conditioned upon the director failing to receive a specified vote for reelection as a director may provide that it is irrevocable.[30]

As a result, Delaware corporations can require their directors to tender advance resignations contingent upon future events, and such resignations can be irrevocable.

Not all states have followed Delaware's lead. Consequently, any company considering adoption of a majority voting model must seriously consider the ramifications of failed director elections. For example, the practical reality that an incumbent director may refuse to resign in the context of an uncontested election must be taken into account. If the losing director is forced to resign, it is possible that a court could invalidate the board's decision depending on the wording of the director removal statute in the state in question. Moreover, even if state laws are modified to help companies deal with these issues, companies must also consider the internal political ramifications of asking sitting directors to tender advance resignations that may become irrevocable.[31]

Assuming a company adequately considers failed director elections and applicable state law supports a shift toward majority voting, other political and legal considerations remain. While requiring all directors to tender advance resignations contingent upon a failed election may solve the holdover problem, it does not necessarily address the broader issue of certain consequences that may follow once a director's resignation is accepted by the board of directors. For example, what if the director required to resign in a failed election is an independent director whom the board needs in order to comply with certain listing rules? What if the resigning director triggers a change in control provision in the corporation's debt instruments? What if the director is also the CEO (which may breach the CEO's employment agreement triggering significant severance payments)?

Intel's model addresses many of the questions that arise in the context of failed elections by giving the board discretion in considering the acceptance or rejection of a director's resignation in a failed election. Only time and experience will tell, however, if the models adopted most recently by companies, including Intel, will be sufficient to address all of the legal and practical challenges that may arise in the coming years.[32]

C. Broker Non-Votes

Recently the New York Stock Exchange proposed an amendment to its Rule 452, effective January 1, 2008, eliminating discretionary voting in director elections.[33] Under Rule 452, brokers can vote shares on "routine" matters in instances where the beneficial owner of the shares has not provided instructions on how to vote.[34] If the proposed amendment is approved by the Securities and Exchange Commission, director elections will no longer be considered "routine."[35]

This Rule 452 development is important because, historically, brokers followed management's recommendations on how to vote in director elections. Beginning in 2008, however, brokers may no longer be able to follow the advice of management in director elections if they have not received specific voting instructions from the beneficial owners of the shares. This opens the door to an increase in the potential number of failed elections, for several reasons. For example, under the Pfizer approach, a director could receive more "withheld" votes than "for" votes because of a failure of certain beneficial owners to provide instructions on how to vote. As such, the proposed amendment to Rule 452 is cause for some concern among companies considering adoption of a majority voting standard.

D. Policy Considerations in Majority Voting

Even if the standards adopted by most companies prove to be adequate in addressing many of the issues that arise in the context of majority voting, do the standards adopted adequately address relevant policy considerations, and will such standards prove to enhance stockholder value in the future? The discretion left to boards in both majority voting models to accept or reject the resignation of a director is one such consideration. The board can determine not to accept the resignation of a director, but must publicly disclose its reasons for so acting. Public disclosure, however, does not effectuate the vote of stockholders if the board decides to leave a director in office after a failed election. If the majority voting movement's purpose is to give a meaningful voice to the stockholders, what purpose does such a standard serve if the board can summarily determine not to honor the vote of the stockholders?

Recent developments in corporate law also call into

question whether a stockholder or a director is better situated to act in the overall best interests of the company. The quintessential corporate governance requirement that directors in a corporation owe a fiduciary duty to *all* of the stockholders should also be considered in the majority voting paradigm. Stockholders owe no such duty of loyalty or care to companies in which they invest. Historically, the assumption has been that stockholders will always act in the best interests of the company because at the end of the day the best interests of the company will yield the best stockholder returns. This assumption, however, has come under attack with recent developments in "empty voting" or "vote morphing."

On January 22, 2007, SEC Commissioner Paul S. Atkins highlighted the potential problems with "empty voting," or votes by stockholders that have no economic interest in the stock they vote.[36] Such trends highlight the potential undermining of the basic assumption that stockholders vote in their best economic interest. It also places a renewed emphasis on the fact that directors legally owe a fiduciary duty to *all* stockholders, not to certain segments or special interest groups found within certain stockholder constituencies. As applied to the majority voting paradigm, in some cases it may be better for a company to retain a director under a plurality standard, because of the applicable fiduciary duties that apply, instead of opening the board room to stockholders who may have no similar duty.

Another concern in adopting a majority voting standard is that of reducing the number of potential directors in the market. It is counterintuitive to expect that directors or potential directors will not react to the possibility that they may face potential ouster from a board of directors in fulfilling their fiduciary duties to a company. Such concerns among the potential director pool could further reduce the availability of potential directors from a pool that has already been diminished as a result of other market movements.[37] If a reduced director pool is one of the unanticipated results of majority voting, the potential for harm to companies generally could be significant. The goal of every company should be to attract the most qualified directors in the market. If the market yields sub-par directors, companies, and stockholders will pay the cost.

Some examples in the marketplace also demonstrate that the plurality system of previous years has been as effective as majority voting in giving a voice to stockholders. For example, in 2004, Disney replaced Michael Eisner as Chairman when 43% of votes were "withheld" for his election.[38] Merrill Lynch's general counsel summarized the effectiveness of a plurality standard as follows:

> I think that in today's environment, most Boards and managements are extremely sensitive to governance criticisms and the related ugly publicity, and will carefully analyze any significant withhold votes and the message behind it and consider the implications for a particular director's continued service.[39]

Others have similarly argued that stockholders already have the power to remove directors through different mechanisms by "calling special meetings, acting by written consent without meeting in person, or making a motion at the annual shareholder meeting."[40] All of this indicates that a majority voting standard may not be needed in the first place.

Moreover, on July 26, 2007, the SEC adopted a rule titled "Shareholder Choice Regarding Proxy Materials" that could significantly reduce the cost to stockholders in making proxy proposals, such as those that arise in the context of director elections, by expanding the use of the internet to lower the cost of proxy solicitations generally.[41] Because of the potential cost savings under the proposed rule, it could dramatically change the face of proxy challenges by empowering average stockholders with the ability to challenge management and boards of directors under the existing regulatory framework. As such, the policy argument that holds companies should support majority voting to give a greater voice to stockholders may no longer apply, because stockholders will be given a greater voice under the existing framework.

As a practical matter, there are also many companies that remain under the stockholder proposal radar and are not targeted by stockholder activists. As such, the pressure that other companies may feel to adopt a majority voting model may not apply. These companies have the ability to watch the market develop standards that will evolve into best practices. Directors at these companies can make a strong argument that, given the uncertainties in the development of majority voting standards, they are acting in the best interests of the company by waiting to see what will emerge as a best practice in the future.

While there are many elements that should give directors and management pause in considering a majority voting standard, there are other things that should be weighed in the balance to determine the best course for a particular company to follow. Majority voting, despite its uncertainties, may be the best choice for certain companies, depending on the circumstances.

III. Benefits to Majority Voting

Many reasons can be provided in support of majority voting in director elections. This debate is not new. In fact, Commissioner Atkins recently stated that "[t]he question of whether the SEC should mandate shareholder director nominations is one that goes back to the very formative years of the Commission."[42] Board accountability, pressure to adopt an emerging best practice, and the ability to preempt stockholder proposed majority voting standards are some reasons why a company may consider adopting a majority voting model. In addition, ISS considers it a positive factor in the corporate governance score it assigns to public companies.[43]

A. Board Accountability

Many have argued that majority voting gives stockholders a true voice in director elections and that it will cause directors to be more responsive to the needs of stockholders. To date, there is very little quantifiable evidence to support this claim as most of the majority voting models have been adopted recently so that there is not enough data.

Of course, some have also argued that directors are already paying attention even when a "withheld" vote may not be binding.[44] In a 2005 report, ISS observed that "[m]ost directors are highly dedicated and successful men and women who care about their integrity and their reputation. It is only natural that withhold votes, whether symbolic or legally binding, will matter a great deal to them."[45] This was exemplified by the Home Depot board meeting in May 2006 when approximately 30% of

shareholders withheld their votes for the election of ten directors in protest over the CEO's pay.[46] The Home Depot withhold campaign ultimately led to the ouster of Home Depot's CEO in January 2007.[47]

B. Preemption of Stockholder Proposals

One of the advantages directors and management have in the face of this trend towards majority voting is to define the models adopted. A number of the recently adopted majority voting models were adopted in the face of an imminent stockholder proposal.[48] There is no guarantee that pre-emptive action on the part of a board will preclude stockholder action, but it is a good-faith gesture to stockholders that the board is aware of stockholder concerns and is willing to give a voice to the stockholders in the resolution of those concerns. Such a gesture may be enough, depending on the stockholder constituency, to satisfy the demands of stockholders pushing for a majority voting standard, and simultaneously give directors and management more of a voice in determining what that standard will be.

Adding importance to this preemptive strategy, stockholder activists have begun to submit binding stockholder proposals calling for majority-vote bylaws. The Second Circuit ruled in September 2006 that these proposals could not be excluded on the basis that they relate to the election of directors.[49] Consequently, stockholders, at least in the Second Circuit, can now present proposals relating to majority voting bylaw amendments in proxy statements. The SEC, however, issued a proposed rule on July 27, 2007 to clarify its position and to amend the applicable rules to provide that such proposals "may be excluded... if [they]... result in an immediate election contest... *or* [if they] set up a process for shareholders to conduct an election contest in the future by requiring the company to include shareholders' director nominees in the company's proxy materials for subsequent meetings."[50]

C. Best Practice

Given the number of companies that have already adopted majority voting models, one could argue that majority voting has already become an accepted best practice for U.S. public companies.[51] Importantly, however, many of the quirks associated with majority voting must be remedied through the trial of time and experience. Accordingly, boards are not, and should not, currently be required to adopt a majority voting model in all circumstances. But the day may not be that distant when majority voting becomes the default standard and boards will have to provide reasons as to why such a standard should not apply to their company.

Currently, the argument that all companies should adopt a majority voting model fits better under the "herd" mentality discussed above, given the related uncertainties. Once the majority voting models have been tried and tested, boards could be held to a new standard in corporate governance with respect to majority voting. All of this assumes, however, that once the market tests the current majority voting standards they remain viable in the marketplace.

Endnotes

1 SEC Release No. 34-48626 (Oct. 14, 2003).

2 On July 27, 2007, the SEC proposed another rule titled "Shareholder Proposals Relating to the Election of Directors" to amend Rule 14a-8(i)(8) and to clarify that such a proposal "may be excluded under Rule 14a-8(i)(8) if it would result in an immediate election contest . . . or would set up a process for shareholders to conduct an election contest in the future by requiring the company to include shareholders' director nominees in the company's proxy materials for subsequent meetings." SEC Release No. 34-56161 (July 27, 2007). This proposal was largely the result of the Second Circuit's ruling in AFSCME v. Am. Int'l Group, Inc., 462 F.3d 121. *See*, infra, note 49 and accompanying text.

3 Claudia H. Allen, *Study of Majority Voting in Director Elections* (Feb. 5, 2007) *at* http://www.ngelaw.com/files/upload/majority_callen_020707.pdf (last visited July 21, 2007).

4 The applicable portion of Section 141(b), as amended, provides:

> Any director may resign at any time upon notice given in writing or by electronic transmission to the corporation. A resignation is effective when the resignation is delivered unless the resignation specifies a later effective date or an effective date determined upon the happening of an event or events. A resignation which is conditioned upon the director failing to receive a specified vote for reelection as a director may provide that it is irrevocable.

8 Del. C. § 141(b) (2007). The applicable portion of Section 216 provides that "[a] bylaw amendment adopted by stockholders which specifies the votes that shall be necessary for the election of directors shall not be further amended or repealed by the board of directors." 8 Del. C. § 216 (2007).

5 *See*, e.g., Cal. Corp. Code § 708.5 (Deering 2007). *See also*, infra, note 31 and accompanying text discussing changes to the Model Business Corporation Act.

6 *See*, e.g., Pfizer, Inc., Corporate Governance: Principles, Item 7 *at* http://www.pfizer.com/pfizer/are/mn_investors_corporate_principles.jsp (last visited July 21, 2007).

7 *See*, e.g., Intel Corporation, Bylaws, Article III, Section 1, *available at* http://www.intel.com/intel/finance/docs/bylaws.pdf (last visited July 21, 2007).

8 *See*, e.g., Progress Energy, Inc. Form 10-Q filed with the SEC on August 9, 2006, *available at* http://www.sec.gov/Archives/edgar/data/17797/000109409306000296/form10-q212006.htm (last visited July 21, 2007). Progress Energy amended Article 5, Section 4 of its Articles of Incorporation to provide:

> [E]ach director shall be elected by a vote of the majority of the votes cast with respect to the director at any meeting for the election of directors at which a quorum is present, provided that if the number of nominees exceeds the number of directors to be elected, the directors shall be elected by a vote of the plurality of the shares represented in person or by proxy at any such meeting and entitled to vote on the election of directors. For purposes of this Section, a majority of the votes cast means that the number of shares voted "for" a director must exceed the number of votes cast "against" that director.

Id.

9 Pfizer adopted its policy on June 23, 2005. A copy of the policy is *available at* http://www.pfizer.com/pfizer/are/mn_investors_corporate_principles.jsp (last visited July 21, 2007).

10 *Id.* (emphasis added).

11 Most companies that have adopted a "true" majority voting standard require a majority of the votes cast as opposed to a majority of votes outstanding. *See* Allen, *supra*, note 3, at iii.

12 Institutional Shareholder Services, *Director Election Reforms*, (2006) *at* http://www.issproxy.com/pdf/DirectorElectionReform_CommentPeriod.pdf (last visited July 21, 2007).

13 *Id.*

14 *Id.*

15 *Id.*

16 Intel Corporation Form 8-K filed with the SEC on January 18, 2007 (emphasis added), *available at* http://www.sec.gov/Archives/edgar/data/50863/000005086307000009/exh31.txt (last visited July 21, 2007).

17 Allen, *supra*, note 3, at v.

18 Allen, *supra*, note 3, at ii.

19 Allen, *supra*, note 3, at i–vi.

20 Allen, *supra*, note 3, at i–ii.

21 *See* SEC Form 8-K, Item 5.03, *available at* http://www.aspenpublishers.com/secureredbox/crpfrm2.pdf (last visited July 21, 2007).

22 8 Del. C. § 216 (2007).

23 *See*, e.g., 8 Del. C. § 242 (2007).

24 Allen, *supra*, note 3, at ii.

25 Allen, *supra*, note 3, at iii.

26 *See*, e.g., 8 Del. C. § 141(b) (2007).

27 Intel Corporation Form 8-K filed with the SEC on January 19, 2006, *available at* http://www.sec.gov/Archives/edgar/data/50863/000005086306000007/exh31.txt (last visited July 21, 2007). In its Form 8-K filed on January 18, 2007 with the SEC, Intel modified its bylaws by removing a portion of the above referenced language and adopting the following policy:

> Director nominees annually submit a contingent resignation in writing to the Chairman of the Corporate Governance and Nominating Committee to address majority voting in director elections. The resignation becomes effective only if the Director fails to receive a sufficient number of votes for re-election at the Annual Meeting and the Board accepts the resignation.

Intel Corporation Board of Directors Guidelines on Significant Corporate Governance Issues, Board Composition: Advance Resignation to Address Majority Voting, *available at* http://www.intel.com/intel/finance/docs/Corp_Governance_Guidelines.pdf (last visited on July 21, 2007).

28 *See*, supra, note 27, and accompanying text.

29 *See*, e.g., 8 Del. C. § 141(k) (2007).

30 8 Del. C. § 141(b) (2007).

31 For further information on proposed amendments to the Model Business Corporation Act to address certain issues that arise in the context of majority voting, see *Changes in the Model Business Corporation Act—Amendments to Chapter 7 and Related Provisions Relating to Shareholder Action Without a Meeting, Chapters 8 and 10 Relating to Shareholder Voting for the Election of Directors, and Chapter 13, Relating to Appraisal and Other Remedies for Fundamental Transactions*, 61 Bus. Law 1427 (Aug. 2006).

32 Intel has already modified its original bylaw amendment adopted in January of 2006. *See supra*, note 27, and accompanying text.

33 NYSE Group, *NYSE Adopts Proxy Working Group Recommendation to Eliminate Broker Voting in 2008*, *at* http://www.nyse.com/press/1161166307645.html (last visited July 21, 2007).

34 *Id.*

35 *Id.*

36 Commissioner Atkins stated the following in a speech given to the Corporate Directors Forum in early 2007:

> Financial investors today can relatively easily and cheaply augment their voting power or alter voting results, even when they have no ownership interest in the corporation. In one example, a hedge fund owned a significant stake in a company targeted in a stock-for-stock tender offer. The hedge fund stood to profit greatly if the acquisition went through. However, dissident shareholders of the *acquiring* company were complaining that the deal was overpriced and should be abandoned.
>
> So what did this hedge fund do? It purchased a 9.9% stake of the *acquiring* company, but then entered into equity swaps and other transactions to eliminate any of its economic interest in the acquirer. Thus, you had a situation where this hedge fund could vote a significant block of the acquirer's shares to approve a transaction which would, in turn, benefit the hedge fund but arguably hurt the shareholders of the acquirer.

Commissioner Paul S. Atkins, Remarks at the Corporate Directors Forum 2007 (Jan. 22, 2007), *available at* http://www.sec.gov/news/speech/2007/spch012207psa.htm (last visited July 21, 2007)

37 Institutional Shareholder Services, *Majority Voting in Director Elections: From the Symbolic to the Democratic*, at 12 (2005), *available at* http://www.issproxy.com/pdf/MVwhitepaper.pdf (last visited July 21, 2007) (discussing the fact, among other things, that majority voting could affect the potential director pool).

38 Paul R. La Monica, *Eisner out as Disney* chair, CNNMoney.com, March 4, 2004, *at* http://money.cnn.com/2004/03/03/news/companies/disney/index.htm.

39 Institutional Shareholder Services, *Majority Voting in Director Elections: From the Symbolic to the Democratic*, at 14 (2005), *available at* http://www.issproxy.com/pdf/MVwhitepaper.pdf (last visited July 21, 2007).

40 *Id.*

41 *See* SEC Release No. 34-56135 (July 26, 2007). The Release states that "[b]y requiring Internet availability of proxy materials, the amendments are designed to enhance the ability of investors to make informed voting decisions and to expand use of the Internet to ultimately lower the costs of proxy solicitations." *Id.* at 7.

42 Commissioner Paul S. Atkins, Remarks at the Corporate Directors Forum 2007 (Jan. 22, 2007), *available at* http://www.sec.gov/news/speech/2007/spch012207psa.htm (last visited July 21, 2007).

43 Institutional Shareholder Services, *ISS US Corporate Governance Policy 2007 Updates* (2006) *at* http://www.issproxy.com/pdf/2007%20US%20Policy%20Update.pdf (last visited July 21, 2007).

44 *See supra*, note 39, and accompanying text.

45 Institutional Shareholder Services, *Majority Voting in Director Elections: From the Symbolic to the Democratic*, at iv (2005), *available at* http://www.issproxy.com/pdf/MVwhitepaper.pdf (last visited July 21, 2007).

46 *Activist Fund Takes on Home Depot*, The New York Times Deal Book, December 18, 2006, *at* http://dealbook.blogs.nytimes.com/2006/12/18/activist-fund-takes-on-home-depot/ (last visited August 4, 2007).

47 Parija B. Kavilanz, *Nardelli out at Home Depot*, CNNMoney.com, January 3, 2007, *at* http://money.cnn.com/2007/01/03/news/companies/home_depot/index.htm?postversion=2007010319 (last visited August 4, 2007).

48 Allen, *supra*, note 3, at iv.

49 AFSCME v. Am. Int'l Group, Inc., 462 F.3d 121.

50 SEC Release No. 34-56161 (July 27, 2007).

51 *See supra*, note 3, and accompanying text.

Criminal Law and Procedure

Minding Moral Responsibility: The Supreme Court's Recent Mental Health Rulings

*By Steven K. Erickson**

It can be fairly said that American criminal law is based upon a moral consensus about which behaviors are considered right or wrong. This consensus is derived from our cultural and legal traditions, which inseparably hold individual autonomy and freedom in tandem with individual responsibility. As such, which behaviors are considered right or wrong flows not so much from legal precedent but from popular notions of agency, accountability, and the belief in an objective truth demarking good actions from evil ones. Yet, modern times have borne witness to such truly revolutionary advances in psychological science that many question whether these popular beliefs about agency and accountability are in fact true. The emergence of various forms of brain scanning technologies has led scientists to make startling claims. Recent studies have suggested that the brain embarks on a decision before an individual is actually aware of his choice,[1] while others propose that neuroscientists have located an area of the brain, known as the dorsolateral prefrontal cortex, where moral decision-making takes place.[2]

The contrived methodology and novelty of these findings has not prevented many scientists and scholars alike from speculating on the substantial implications these findings suggest.[3] In regards to the former, some conclude that most decisions are not voluntarily made by the actor, while, in the latter, others claim damage to this area may leave some unable to behave in a morally responsible manner.[4] In a similar vein, additional scientific findings ranging from behavioral genetics of psychopathy to predictions of future dangerousness have given fodder to those who argue that historical conceptions of free will are false and most human behavior determined by genetic factors, leaving little room for choice.[5] If such conclusions are correct, we are indeed in a new era of criminal and moral responsibility, whereby individual responsibility itself is anathema to a coherent regimen of criminal law.

But history is a testament to scientific theories once embraced as definitive and later abandoned because they are fraught with flaws and oversimplifications. In the late nineteenth century, Italian criminologist Cesare Lombroso posited the idea that criminal behavior was an immutable genetic trait that set affected individuals down an inevitable road of criminality and vice.[6] Lombroso's work centered on identifying these hapless miscreants, and suggested various physiognomic features could be used to ascertain which people were in fact "natural born killers." This imprudent social Darwinist approach seems laughable today, but was considered serious science at the time and employed statistical analyses which gave Lombroso's theory an appearance of scientific certainty and impartiality. Thus, it is unsurprising that Lombroso's theory was heavily relied upon by the eugenics movement, which advocated substantial coercive governmental policies aimed at establishing better societies through forced sterilization, massive institutionalization of "undesirables," and even genocide by its fanatics. Less stark but no less grievous was Sigmund Freud's enduring theory of psychosexual development and the subconscious, which captivated behavioral scientists and our popular culture with notions of a universal Oedipus complex, repressed memories, and the derivative "schizogenic mother."[7] Similar to Lombroso, Freudian theory was once heralded as the scientific explanation for a variety of social woes including violence, criminality, and delinquency, only to be dismissed in later years as utter pseudoscience.[8]

Despite this sordid past, psychological science has indeed benefited our culture and legal system. Most reputable behavioral scientists now agree that severe mental illnesses like schizophrenia and bipolar disorder are caused by significant biological deficiencies in the brain, and not by frigid parents or an individual's choice to have a "unique" outlook on life, as perpetuated during the 1960s by prominent psychiatrists such as Thomas Szasz and R. D. Laing. We have learned a great deal about the fallibility of eyewitness testimony and confessions conducted under extreme duress.[9] Such accomplishments should be praised for their value in ensuring a just criminal justice system. Since science and law approach the world from differing epistemologies, it is predictable that disagreements will occur about the normative construction of our criminal code. For the law, individual capacity for responsibility is presumed, and the bar for exculpation is set high, as the opposite would surely cause our entire criminal justice system to collapse under the weight of an endless procedural morass of dueling experts and frivolous affirmative defenses.

Indeed, efficiency is both a necessary and legitimate aim for criminal law. The past thirty years have shown just how damaging an inefficient criminal justice system can be if we examine the nearly endless delays seen in death penalty practice. While arguments abound about the deterrent effect of the death penalty, classic criminological theory suggests substantial delays between the time of the crime and imposition of the penalty likely reduces the deterrent effect.[10] Nonetheless, for science, knowledge is cumulative, and consequently the world seems increasingly nuanced. For every scientific question answered, many more follow in its wake. Thus, questions about competency and insanity seem less clear as mounting evidence suggests mental processes are far more complex than previously thought. The upshot of these different approaches that law and science take has been an indulgence of the former with novel scientific claims and an increasing politicization of the latter. Both are dangerous for different reasons. When the law entertains superficial science, it risks becoming ensnared in a false objective reality; when science becomes a political entity, it gives us more of that specious reality.

* *Steven K. Erickson is a forensic psychologist and the Mental Illness Research, Education and Clinical Center Fellow at Yale University Department of Psychiatry, where he conducts research on the intersection of law and science.*

This past term, the Supreme Court decided two cases that entailed questions of psychological science and criminal law. Both cases involved the death penalty and to some degree questions about mental capacity that tread close to the ultimate issue about how our criminal code conceptualizes the responsible agent. In *Panetti v. Quarterman*, the issue surrounded the competency of a defendant with a well-documented history of schizophrenia who represented himself during a capital trial and was subsequently sentenced to death.[11] After exhausting his state appeals and denied federal habeas corpus relief, he filed a second federal habeas corpus appeal arguing that he was incompetent to be executed because he lacked a rational understanding of why the death penalty was being imposed. Unlike the seminal case of *Ford v. Wainwright*, which prohibited the execution of an incompetent defendant who lacked a factual awareness of his punishment, it was uncontested that Panetti knew that the death penalty was being imposed upon him by the state for a crime in which it held him responsible.[12] Rather, Panetti's delusional belief was that the real reason for his death sentence was because the government wished to prevent him from preaching the gospel of Jesus Christ. Thus, the question presented rested on whether the Constitution requires a defendant to posses a factual and rational awareness of the criminal process in order to be considered a competent for punishment. Since the Constitution is silent about competency, Supreme Court cases have repeatedly engaged in a quasi-orginalist analysis by examining historical traditions of English and American common law regarding mentally impaired defendants. Such analyses always harken back to Blackstone's famous commentaries, which held it immoral to pursue criminal prosecution against "madmen," since such defendants are unable to understand why they are being accused, and thus unable to assist in their defense.[13] Blackstone's moral precept is congruent with our modern criminal justice system, which holds the individual ultimately responsible, and charged with his own defense, in a court of law.

And this is exactly where the problem lies. In the legendary case of *Faretta v. California*, the Supreme Court held that a defendant has a constitutional right to represent himself during a criminal proceeding.[14] As Justice Blackmun aptly noted in his dissent, the Court created a constitutional right to be a fool.[15] Panetti did exactly that, and by any measure his defense was both laughable and tragic. His attempts to subpoena Pope John Paul II, John F. Kennedy, and Jesus Christ were only outdone by his incoherent ramblings and nonsensical testament about his fearless approach to death. Despite ample evidence of his raging psychosis, which was surely exacerbated by his refusal to take his prescribed antipsychotic medicine, the trial court never questioned Panetti's competency once the trial commenced. While the right to be competent to stand trial is firmly rooted in American criminal law and affirmed by no less than four major Supreme Court decisions during the past forty years, *Faretta* seems to have trumped that right in the *Panetti* case. Nonetheless, in writing for the majority, Justice Kennedy held that the state could not execute Panetti since he lacked a rational appreciation for why the state wished to put him to death. Pointing to *Ford*, the 5-4 Court opinion concluded that it was inhuman to send an insane person to death, while engaging in a tortuous analysis of what *Ford* meant in its fragmented plurality opinion by holding that defendants must have "awareness" to be considered competent. While the Court in *Ford* complained about the "subtleties and nuances" of psychiatric evidence as reason for an abundance of caution when approaching mental health issues, in *Panetti* the Court appears to eagerly agree with the amicus brief filed by the American Psychological Association supporting Panetti's appeal and the certitude of his diagnosis of schizophrenia as sound science. Despite *Ford's* plurality status, the Court in *Panetti* readily deferred to its holding as the basis for its decision.[16]

However, precedent is flexible matter. In *Ford*, the Court held that when there is a "substantial threshold showing of insanity" a defendant is entitled to an evidentiary hearing on his competency, including production of his own expert witnesses and cross-examination of the state's expert witnesses.[17] Although Panetti's trial transcript is ample proof of his florid psychosis during trial, the Court relies upon a one-page affidavit by a psychologist and law professor who interviewed Panetti for one hour as satisfying *Ford's* "substantial showing" of incompetence. On its face, these seem apposite since "substantial showing" presumably means that the presumption of competency is given significant deference and overcoming this presumption requires considerable evidence to the contrary. Yet psychosis is a difficult matter. On the one hand, interviewing someone in the throws of florid psychosis is often a fruitless endeavor and one may not need more than a modicum of interaction to ascertain that the afflicted is seriously impaired. On the other hand, if rationality is such an intractable and subtle construct to determine as *Ford* and *Panetti* seem to indicate, than one wonders whether such a brief interview could accurately illuminate Panetti's true competence. Nonetheless, since the Court did not point to anything within the one-page report as a definitive showing of incompetence, a sensible interpenetration of the Court's ruling suggests that the *Ford* precedent of "substantial showing" may not be as high of a burden for defendants as the vernacular suggests.

Of course, precedent is a matter that the Court has the liberty of defining, but Acts of Congress are usually left to "plain meaning" and the intentions of Congress. Thus, the second issue resolved in *Panetti* was whether Panetti's second federal habeas corpus petition was a "second or successive" petition prohibited under the Antiterrorism and Effective Death Penalty Act (AEDPA).[18] Since Panetti had not claimed during his first habeas petition that he was incompetent to be executed, his second habeas petition seems a clear violation of the Act. Not so, held the Court, holding that the Act's phrase "second or successive" was not "self-defining" but instead in harmony with prior case law.[19] As Justice Thomas noted, however, in his dissent, the Court's opinion neither cites to a pre-AEDPA case that defined "second or successive" nor any pre-AEDPA case in which a subsequent habeas application challenging the same state-court judgment was considered anything but a "second or successive" petition. The best the Court can do is distinguish *Panetti* from its prior holding in *Stewart v. Martinez-Villareal*, where the Court affirmed a defendant's right to raise a second *Ford* claim when his initial claim had been dismissed as unripe. In justifying its novel finding that Panetti's claim was not second or successive the majority suggests that the logical consequence

of adopting a literal interpretation of the AEDPA's prohibition of second or successive petitions would equate to a tacit approval by the Court for all defendants to raise Ford claims at the outset of appeal in an effort to preserve the issue.[20] What follows in the *Panetti* opinions is a debate between the majority and the dissenting Justices about whether allowing Panetti's second habeas petition promotes or detracts from judicial efficiency, with the implicit difference of opinion resting on whether district courts can easily dismiss illegitimate *Ford* claims if the literal meaning of AEDPA's prohibition is followed. And what was lost in this procedural disputation is that while judicial efficiency is surely a legitimate and important consideration, none of this would be at issue if our judicial system was more concerned about the illegitimacy of an obviously incompetent defendant allowed to proceed at trial and less concerned about the intricacies of the legal lexicon behind words like "second or successive." Of course, Panetti's competence at trial was not at issue, and thus the Court ruled only on the narrow issue before it; namely, whether Panetti was competent to be executed. Applying narrow holdings to constricted fact patterns is what courts do, after all.

This narrow approach of court holdings is both a blessing and a curse. In *Panetti*, Justice Kennedy, writing for the majority, assures us that the mental state at issue is a narrow one. In attempting to foreclose any debate about which mental states may qualify for the newly created *Panetti* exception to AEDPA's ban on "second or successive" petitions, the Court holds that "the beginning of doubt about competence in a case like petitioner's is not misanthropic personality or amoral character. It is psychotic disorder."[21] It is reasonable to assume that the Court wished to prevent opening a floodgate of claims by death row inmates claiming incompetence based on the widely used psychiatric classification manual, the Diagnostic and Statistical Manual for Mental Disorders (DSM), published by the American Psychiatric Association. Now in its fourth edition, the DSM was first published in 1952 and contained 106 mental disorders at a length of 130 pages.[22] It was also the first time the term mental disorder was widely used in place of terms like mental illness or mental disease. Since then, the number of diagnosable mental disorders has blossomed, with the current edition at a hefty 886 pages containing 297 mental disorders.[23] Included among these 297 mental disorders are such mental phenomena as "Breathing-Related Sleep Disorder",[24] "Hyposexual Desire Disorder",[25] and the various personality disorders.[26] Recently, the American Medical Association tabled consideration of the addition of Video Game Addiction Disorder.[27] There are many critics of the DSM and many valid criticisms of the numerous behaviors now considered mental disorders, and hence under the purview of behavioral experts. What is telling, however, is how the Court approaches these various mental disorders and how it applies its skepticism of psychiatric "subtleties and nuances." In *Panetti*, the Court reassured us that the point of departure was Panetti's psychotic disorder and not anything else. Thus, the Court seemed reasonably confident that psychotic disorders were legitimate mental illnesses worthy of special consideration in criminal proceedings. In this sense, the Court appeared to be making a judgment that psychoses like Panetti's schizophrenia are legitimate while "misanthropic personalities" are not. Yet the second case to be discussed in this essay casts some doubt on this supposed bright line.

In *Schriro v. Landrigan*, a 5-4 Supreme Court reversed the Ninth Circuit Court of Appeals ruling that defendant Landrigan's attorney provided ineffective assistance of counsel by failing to investigate and provide mitigating evidence of Landrigan's purported antisocial personality disorder.[28] The defendant, Jeffery Landrigan, had a long history of antisocial behavior that stretched back to his childhood which included drug use during his early teen years. In 1982, Landrigan was convicted of murder and while serving a sentence for that crime, was convicted for repeatedly stabbing another inmate. In 1989, Landrigan escaped from custody and subsequently murdered another man. During the ensuing sentencing for that murder, Landrigan's counsel sought to introduce mitigating evidence in accord with *Wiggins v. Smith* by having Landrigan's mother testify regarding Landrigan's troubled childhood, and his ex-wife about being a good father for their child.[29] But Landrigan would have none of it. He repeatedly interrupted his counsel and stated "if you want to give me the death penalty, bring it on."[30] Nevertheless, the trial judge asked Landrigan whether he had instructed his attorney not to bring *any* mitigating evidence forward. Landrigan responded in the affirmative and the judge subsequently sentenced Landrigan to death. On direct appeal, the Arizona Supreme Court affirmed the sentence. As expected, a federal habeas claim was filed next and was subsequently denied by the District Court which was unanimously affirmed by the Ninth Circuit panel. However, the full court granted a hearing en banc and reversed, holding that Landrigan had made a colorable claim under *Strickland v. Washington*.[31] The Supreme Court reserved, holding no violation of *Strickland* had occurred and found that Landrigan had clearly waived his right to present mitigating evidence.

In a curious dissent, Justice Stevens wrote that Landrigan's purported anti-social personality disorder was a "serious organic brain syndrome" and should have been further investigated and presented by Landrigan's counsel.[32] Furthermore, Justice Stevens suggested that Landrigan had not waived his right to present mitigating evidence despite being asked directly by the trial judge whether he had instructed his attorney not to present any mitigating evidence. In no less than seven instances, the dissenting opinion referred to Landrigan's antisocial personality disorder as a "serious organic brain syndrome" despite the fact that the DSM does not refer to it in such a manner, instead classifying it as a personality disorder. A reasonable reading of the term "serious organic brain syndrome" suggests that the dissent was implying that Landrigan was unable to control his behavior because such behavior was the product of a brain disease on par with epilepsy. In fact, the DSM reserves the term "organic" for mental disorders, such as delirium, mental retardation, and dementia—although organic simply means originating from living organisms. Indeed, the dissent points to a psychological report prepared for the defense claiming that Landrigan's violence likely stemmed from a genetic disposition, leaving him unable to control his behavior.

While some may find the dissent's characterization of Landrigan's anti-social personality disorder bordering on the

absurd, such claims are not uncommon as of late. Numerous stories in the popular media discussing recent brain imaging findings regarding antisocial personality disorder, psychopathy, and impaired moral decision-making have generated considerable attention. In fact, a recent article in the *New York Times* discussed the vast increase in the number of legal claims based on brain imaging technology and impaired moral reasoning which have been enthusiastically received by numerous law professors and neuroscientists.[33] Likewise, a recent law school symposium was entirely dedicated to brain scanning as the possible "next big thing," implying a radical transformation of our cultural and legal traditions is at hand.[34]

Such enthusiasm is both undeserved and troubling. While brain-scanning technology has allowed scientists unprecedented access to living brains, fundamentally all they can describe is brain anatomy and physiology. Of course, these are vitally important matters, but whether a certain location of the brain appears inactive during a task performed while confined within the enormous magnetic field that constitutes an MRI machine tells us little about such global constructs as free will, agency, and accountability. The evidence from these limited studies simply cannot be generalized to the larger concepts of concern in criminal law. Moreover, the impressive pictures that these technologies produce are often based on extremely small sample sizes that undermine any statistical analyses performed to assess for differences between the supposed affected individuals and the normal controls. Additionally, they rely entirely upon complicated mathematical algorithms which employ "image smoothing" techniques that impute missing values, and hence add artificial values for missing data. Such procedures are entirely legitimate, as most computer monitors and televisions employ similar methods, but the notion that brain scans take a "picture" of the brain implies a precision they do not deserve.

Nonetheless, the findings from the various brain imaging technologies have provided additional evidence that behavior is a complex phenomenon, and there is no doubt that as the technology improves further evidence will come forward suggesting certain brain areas are associated with certain behaviors—even legally relevant behaviors. But scholars who adopt such findings to suggest that free will and agency are myths simply misunderstand the science behind the findings. For instance, the dissent in *Landrigan* implied that because Landrigan's behavior was the product of biology—a possible broken brain—he was entitled to presentation of this evidence, because it probably would have mitigated his culpability. But, of course, all behavior is biologically derived. Entrenched American legal traditions hold that individuals may be exculpable for criminal offenses in only limited circumstances where a mental illness so severely and substantially undermines an actor's rationality that holding such defendants culpable would offend common notions of decency. In fact, an equally compelling argument can be made that anti-social personality disorder should be exclusively considered an aggravating factor, since, absent a psychotic illness, such a diagnosis is associated with a rational actor who has a substantially elevated risk of future criminality and dangerousness. Mental disorder and madness are not synonymous, and not all mental disorders are created equal. Irrespective of their biological origins, some mental disorders deserve our pity, while others rightfully signal little compassion by the public. Mitigation rests not on a mechanical determination that some brains are biologically impaired but from the moral precept that substantial mental impairments can undermine individual accountability within our criminal justice system. Thus, our criminal justice system allows mitigation for those limited defendants because it reflects our collective belief in mercy for those unable to defend themselves before the law. As the majority aptly said in the *Panetti* case, it is cruel to send an insane person to death. Such pronouncements have little to do with the actual defendant suffering the penalty, since arguably psychoses like Panetti's would probably obscure realization of the impending penalty. Rather, the cruelty Justice Kennedy refers to in *Panetti* speaks volumes about our social norms, including our desire to punish only those who understand why the state wishes to exact the ultimate price for "amoral" behaviors. As the Court noted in *Panetti* when speaking of competency, the doubt begins with psychosis and is limited to very few other mental disorders.

The problem with our behavioral sciences lies with the misnomer of mental abnormality. When the DSM abandoned the term mental disease in favor of mental disorder few perhaps understood at that time the future it would unleash. But, as psychological science has moved further away from the constricted dimension of disease in describing abnormal behavioral phenomena towards the almost boundless construct of disorder, an increasingly greater number of behaviors have fallen under the authority of behavioral experts. As it remains entirely unclear what makes a cluster of behaviors a disorder, behavioral science experts and our culture have become complacent with the very idea that bad behavior is caused by bad biology. Such thinking has infected our beliefs about all behaviors that we view as undesirable; hence, we have respected national medical organizations seriously contemplating whether playing video games too much should be called and *thought of* as addictions. Likewise, our legal system has adopted the notion of indefinite civil commitment for sex offenders not because they have a disease as conventionally construed but because it is postulated that some hereto unknown mental abnormality causes them to engage in the worst behavior imaginable against our most helpless citizens. While the incapacitation of such offenders is desirable and understandable, the means of achieving that end have hastened our journey down the road of biology run amuck with any biological abnormality located in the brain as sufficient evidence for jettisoning our long-standing traditions of holding individuals responsible for their behaviors irrespective of their individual idiopathic differences. American criminal law has always set the bar high for diminished capacity and exculpatory defenses, not because it is ignorant of the individual differences people have but because it demands equal compliance of the law from everyone, irrespective of those differences. Only under that regimen can we have a comprehensible and effective criminal code that assures the biologically gifted and the biologically deficient that all citizens they encounter are expected not to murder, rape, and otherwise engage in wrongful conduct against them.

Such expectations by the public are not only intuitive but also wise. As certain as Lombroso was that he could identify

the biologically determined criminal over 100 years ago, such hubris is evident with those who confidently proclaim free will a myth because of the brain scans of a select few. Many scoff at the popular skepticism of insanity defenses and claims of childhood maltreatment as an excuse for adult criminality, but such skepticism insulates our legal and moral codes from the corrupting influence of fashionable scientific claims proffered with much certitude early on only to fall into disfavor in ensuring years. Science is about testing hypotheses in the empirical world, but it infrequently proves anything definitively. Even the hard sciences, such as physics, have yet to provide a unified theory of the material universe, and are subject to the flavor-of-the month effect. String theory, for example, has been heralded as the final piece of the puzzle in subatomic physics, but has recently been called into question as an insufficient explanation.[35] When we deal with the products of the metaphysical mind, such as anti-social behavior, caution and a healthy dose of skepticism are in order. At the same time, science does tell us much about the material world in which we live, and our desire to pick and choose which scientific findings we wish to entertain is foolish. Physics may lack a unifying theory, but there is little doubt that gravity exists, just as it is unquestionable that schizophrenia is a severe brain disease. Folks like Scott Panetti deserve our mercy just as Blackstone decreed many years ago, not so much because he deserves it, but because we and our future generations deserve a just and merciful society. The past can tell us much about how to navigate the future ahead of us. Any thoughtful reflection of the past invariably humbles the wise as it becomes apparent how little we really know about our world. Humility is a virtue the behavioral sciences and our criminal-moral code could surely benefit from; the former because it is so certain it has the answers for many questions, and the latter because it is so unwilling to listen to any of those answers.

Endnotes

1 Benjamin Libet, et al., *Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential): The Unconscious Initiation of a Freely Voluntary Act*, 106 Brain 623 (1983).

2 *See, e.g.,* Joshua D. Greene, et al., *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 Sci. 2105 (2001).

3 *See, e.g.,* Deborah W. Denno, *Crime and Consciousness: Science and Involuntary Acts*, 87 Mich. L. Rev. 269 (2002).

4 *See generally,* Peter Arenella, *Convicting the Morally Blameless: Reassessing the Relationship Between Legal and Moral Accountability*, 39 UCLA L. Rev. 1511, 1611 (1992).

5 *See, e.g.,* Matthew Jones, *Overcoming the Myth of Free Will in Criminal Law: The True Impact of the Genetic Revolution*, 52 Duke L.J. 1031 (2003).

6 *See* Arthur Herman, The Idea of Decline in Western History 110-113 (1997).

7 *See* E. Fuller Torey, Freudian Fraud: The Malignant Effect of Freud's Theory on American Thought and Culture (1992).

8 *See* Richard Webster, Why Freud Was Wrong (1995).

9 *See generally* Neil Brewer & Kipling D. Williams, Psychology and Law: An Empirical Perspective (2005).

10 *See* Jeremy Bentham, "Principles of the Civil Code: Principles of Penal Law," *in* 1 The Works of Jeremy Bentham (John Bowring ed., 1962).

11 Panetti v. Quaterman, 127 S. Ct. 2842 (2007).

12 477 U.S. 399 (1986).

13 William Blackstone, 4 Commentaries *20-33.

14 422 U.S. 806 (1975).

15 *Id* at 852.

16 *Ford* at 426.

17 *Id.*

18 Antiterrorism and Effective Death Penalty Act 28 U.S.C. § 2254.

19 *Panetti* at 2853.

20 Stewart v. Martinez-Villareal, 523 U.S. 627 (1998).

21 *Panetti* at 2862.

22 American Psychiatric Association, Diagnostic and Statistical Manual for Mental Disorders (1st ed., 1952).

23 American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders - Text Revision (4th ed. 2000).

24 *Id.* at 622.

25 *Id.* at 539.

26 *Id.* at 835-730.

27 *See AMA Takes Action on Video Games*, American Medical Association, *at* http://www.ama-assn.org/ama/pub/category/17770.html (last modified July 27, 2007).

28 127 S. Ct. 1933 (2007).

29 Wiggins v. Smith, 539 U.S. 510 (2003).

30 *Landrigan* at 1937.

31 Strickland v. Washington, 466 U.S. 668 (1984).

32 *Landrigan* at 1945.

33 Jeffery Rosen, *The Brain on the Stand*, N.Y. Times, March 11, 2007.

34 *The Law and Ethics of Brain Scanning: The Next Big Thing Coomig Soon to a Courtroom Near You?*, Sandra Day O'Conner College of Law, *at* http://www.law.asu.edu/?id=658 (April 13, 2007).

35 Peter Woit, Not Even Wrong (2006).

Genetic Information and Privacy Interests: The DNA Fingerprinting Act

*By Ronald J. Rychlak**

In the early 1980s, geneticist Alec Jeffreys of Leicester University discovered that DNA, the nucleic acid molecule found in all living organisms, could be used to isolate a "genetic marker" unique to each individual. This marker can be found in hair, blood, saliva, and other parts of the body. Not all testing procedures evaluate every aspect of DNA, but DNA evidence can be important in the establishment of paternity, the determination of familial relations for inheritance purposes, and the identification of criminal suspects.

The first reported use of DNA evidence by an American court came in 1988 and it received widespread public attention in the 1995 O.J. Simpson criminal trial, at which the jury rejected the prosecution's DNA evidence.[1] In criminal cases, DNA evidence serves primarily to confirm the presence of suspects at crime scenes.[2] The match of crime scene DNA to an individual by comparison through a database is known as a "cold hit." There are various claims about the number of rapes and murders that could be prevented with cold hits from a well-developed DNA register, but the evidence is conjectural. Nevertheless, it is clear that a larger catalog of DNA samples increases the value of DNA evidence.

Perhaps the strongest case for expansive DNA sampling comes from England, which over the past ten years has undertaken the world's most aggressive DNA gathering effort. Not only do English authorities take DNA samples from arrestees but since 2001 they have been permitted to retain those samples even when the arrest results in an acquittal. Moreover, in 2004 British police were given the authority to collect DNA from mere suspects.[3] Because of these liberal policies, England has been able to acquire and maintain over four million DNA samples, about six percent of the population--more than ten times the percentage of DNA samples maintained in the United States.

The British database has matched nearly 600,000 suspects to crimes.[4] For several years, most American states have had legislation regarding the collection of DNA, usually involving convicted felons.[5] These laws vary significantly from one state to another, but as of today all of the states share their DNA information with a national database, the FBI Laboratory's Combined DNA Index System ("CODIS").[6] CODIS began as a pilot project in 1990, serving fourteen state and local laboratories. Over time, with strong backing from most police departments, many states joined the pool. In 2000, with the enactment of the DNA Analysis Backlog Elimination Act ("Backlog Act"), individuals convicted of murder, manslaughter, sexual abuse, child abuse, kidnapping, robbery, burglary, or any attempt or conspiracy to commit such crimes, could be compelled to submit a DNA sample.[7] Federal, state, and local law enforcement could also input DNA samples to CODIS, and compare crime scene DNA to the samples collected from potential suspects and other crime scenes.

Despite the potential of the Backlog Act, the DNA database did not result in a large number of cold hits. For a DNA database to be truly effective in identifying perpetrators of many crimes, it has to contain many, many samples. Senator Jon Kyl has taken this up as a cause and has been promoting the DNA Fingerprinting Act as a tool to use in preventing crime. The Senator's web page gave the following example:

> In early 1993, [Andre] Crawford was arrested for felony theft. Under the DNA Fingerprint Act, DNA could have been taken from him at that time and kept in [the national DNA database]. Because it was not, when Crawford murdered a 37-year-old woman in September 1993, although he left DNA at the scene, he could not be identified as the perpetrator. Over the next six years, Crawford went on to commit one rape and to murder ten more women between the ages of 24 and 44. If Crawford's DNA sample had been taken and kept in NDIS after his March 1993 arrest, he could have been identified and arrested after the September 1993 murder, and ten more murders and one rape would have been prevented.[8]

Kyl argued that taking a DNA sample was no more invasive or complex than taking a fingerprint from a suspect.

The DNA Fingerprinting Act & Privacy Concerns

The DNA Fingerprinting Act passed the House of Representatives in 2005 as stand-alone legislation. It was incorporated into the Senate's reauthorization of the Violence Against Women Act, and passed in that form. President George W. Bush signed it into law on January 5, 2006. The Act authorizes the collection of DNA from anyone convicted, charged, or arrested for a felony or crime of violence; and from any non-U.S. citizen who is merely detained by a federal agency.[9] It further provides for the DNA samples to be entered into the CODIS system. Those arrestees or detainees who end up being exonerated, having their charges dropped, or against whom charges are never filed, may have their DNA fingerprint removed from the CODIS system, provided that the FBI receives a certified final court order relating to each charge.[10] (Obviously, this is complicated in the case of detainees who were never charged.)

At least ten states have already passed "sample on arrest" laws for some crimes.[11] But the DNA Fingerprinting Act was designed to remove concerns states may have had regarding taking samples and putting them into the CODIS database. With the DNA Fingerprinting Act now law, it is likely that more states will soon adopt these laws.[12] Supporting such a move, Arizona Governor Janet Napolitano said, "DNA in many respects is the new fingerprinting and when people get arrested now they usually get fingerprinted. To me this is just an evolution of that process."[13]

But the involuntary extraction of DNA raises special privacy concerns, particularly as regards the Fourth Amendment's protection against unreasonable searches and seizures. DNA can reveal genetic predispositions and health issues. It also reveals

** Ronald J. Rychlak is Professor of Law and Associate Dean for Academic Affairs at the University of Mississippi. This paper was originally presented at the annual meeting of the Southeastern Association of Law Schools on July 30, 2007.*

information not just about individuals but about their families.[14] As such, the major objections to a DNA sample on arrest policy are the threat it poses to the constitutionally guaranteed personal privacy of individuals, and the Fourth Amendment bar against unreasonable searches and seizures.

Virtually every case challenging DNA collection has recognized that a compelled collection is a search or seizure.[15] However, most courts have found the searches to be reasonable--the searches are almost always related to convicts or probationers/parolees. And the collecting process itself is not burdensome. As the *Arizona Republic* put it, "The collection method is hardly more intrusive than inking fingers to get a set of prints: a swab with a piece of filter paper is rubbed against the inside of a person's mouth to pick up some cells for sampling. (All of which is a lot less stressful than the urine test for drugs that employers routinely require of job applicants.)"[16]

Taking DNA from arrestees, or even detainees, raises much more difficult issues. Unlike convicts, arrestees and detainees have not tested the evidence and been found guilty of the charged crimes. Of course, even with them, the Fourth Amendment is implicated only if obtaining a DNA sample constitutes an unreasonable search. That essentially boils down to a determination of whether arrestees and detainees have a legally recognized privacy interest that is violated by the collection of their DNA.

In *United States v. Dionisio*, the U.S. Supreme Court upheld a grand jury subpoena for a voice exemplar on the theory that the subpoena itself was not a seizure of the person, and that a person's voice cannot be considered private.[17] The Court explained:

> The physical characteristics of a person's voice, its tone and manner, as opposed to the content of a specific conversation, are constantly exposed to the public. Like a man's facial characteristics, or handwriting, his voice is repeatedly produced for others to hear. No person can have a reasonable expectation that others will not know the sound of his voice, any more than he can reasonably expect that his face will be a mystery to the world.[18]

Similarly, the Supreme Court has long recognized that fingerprinting suspects in the course of booking, photographing them for the purpose of generating evidence, or taking handwriting exemplars in a criminal case is such a minimal intrusion on privacy that probable cause is unnecessary. Unless DNA fingerprinting is fundamentally different from these other matters, the Fourth Amendment should not invalidate the DNA Fingerprinting Act.

The most powerful argument for Fourth Amendment protection is that the DNA strands have the potential to reveal far more information than does a fingerprint. "Given the wealth of genetic material in the famous double helix, we should be cautious. It's sensible to be concerned about, say, employers and insurers getting sensitive information that could cause people to lose their jobs or health coverage."[19] The problem is that this argument does not stand up when applied to the way DNA evidence is collected, kept, and used in the CODIS database.

It is true that DNA molecules, if analyzed at certain loci that are not typically used for identification purposes, could reveal the existence of rare diseases or indicate a predisposition to more common ones. Most police laboratories, however, are not even equipped to do such testing.

DNA profiling for law enforcement purposes is so tightly focused that there's no extra information beyond identification. The profile looks at 13 bits of genetic coding that are uniquely combined in each individual—but have nothing to do with predicting susceptibility to disease or other inherited traits. Each entry in the DNA database is just a string of 13 pairs of numbers (from the mother's and father's side).

That DNA code has no use outside the forensic system, said one scientist—unlike a Social Security number, which is a gold mine for identity thieves.[20]

In fact, "there is currently no known potentially compromising genetic information contained among the thirteen CODIS locations other than the fact that they serve as a unique DNA fingerprint that can also confirm familial relationships."[21] Most consumers willingly provide much more private information in exchange for a few cents when they use discount cards at their grocery stores.

If the retained biological samples collected by law enforcement were to be further analyzed for anything other than CODIS loci, a significant personal privacy issue could come into being.[22] This, however, is unlikely. Even though the labs typically retain the actual cell samples, in case extra tests are needed or technology changes, for security reasons those samples "are identified only by bar code and are stored and handled with the same protection as crime evidence." [23] Moreover, even if private information could be extracted from these samples, legislation similar to that used to protect confidential material in the hands of the IRS could be enacted to protect the privacy of the affected individuals.

There is simply no compelling evidence that DNA samples will compromise the privacy of arrestees. On the other hand, adding arrestee DNA to CODIS will almost certainly generate increased cold hits and ultimately reduce crime. The FBI currently claims a cold hit rate of only 22% using CODIS.[24] In the United Kingdom, where law enforcement officials follow a "sample on arrest" policy, the cold hit rate is almost 40 percent.[25] Similar statistics can be found in New Zealand, where sample on arrest policies have been in place for years.[26] Unless and until DNA fingerprinting is proven to reveal more about an arrestee than a unique and unchanging identification code, the value of DNA fingerprinting seems to far outweigh the privacy intrusion on the affected individuals.

Perhaps the hardest question relates to taking DNA samples from mere suspects. This applies only to non-citizens, and several commentators have suggested that it should be seen in light of the recent debates over illegal immigration from Mexico.[27] According to Deborah Notkin, former president of the American Immigration Lawyers Association: "It's so broad, it's scary. It is a terrible thing to do because people are sometimes detained erroneously in the immigration system." Truthfully, however, collection of DNA only becomes a concern if it results in a cold hit, or if the DNA is entered into CODIS and matches up with a future crime. In either of these cases, the equities are with the authorities. Lynn Parrish, spokeswoman for the Rape, Abuse and Incest National Network, pointed

to the case of Angel Resendiz, a Mexican immigrant who committed at least fifteen murders and numerous rapes in the United States. Deported seventeen times before finally being executed in Texas, Ms. Parrish said, "If he had been identified as the perpetrator of the first rapes, it would have prevented later ones." Regarding the DNA Fingerprinting Act, she said, "If this had been implemented years ago, it could have prevented many crimes. Rapists… don't just rape, they also murder."[28]

A slightly different concern is that if a minority racial group is arrested or detained more often than other racial groups, DNA samples will be taken from that minority more frequently, and the DNA database will contain a higher percentage of their DNA.[29] That, of course, would suggest that members of that minority would end up being identified more often through the database. But the same problem is true of traditional fingerprints, and relates not to the DNA collection but to the reasons for the arrests; they might or might not be valid. Moreover, if most crime takes place within racial groups, the minority group might actually end up benefiting, because it will have a safer community. In any case, this is not a reason to hold that the DNA Fingerprinting Act is unconstitutional.

Perhaps the most important remaining question, then, relates to the appropriate role of governmental agencies. There will almost certainly be pressure to expand the CODIS program. It is not hard to imagine a time when DNA samples are taken from all children shortly after birth. This could be justified on a safety basis. DNA fingerprinting will help to exonerate the innocent, convict the guilty, and protect the children. But are governmental agencies competent to handle this much authority? If CODIS expands to the point where it covers all citizens, it may come to embody the "big brother" government so feared by generations of Americans.

DNA evidence is thought by many to be foolproof, but it is only as reliable as the people and processes by which it is collected and analyzed. As with all governmental programs, there will be instances of poor management, budget shortages, and corruption in the CODIS program. Mistakes will happen.[30] Moreover, the DNA Fingerprinting program is "certain to bring a huge new workload for the F.B.I. laboratory that logs, analyzes, and stores federal DNA samples." [31]

Thus far, fortunately, neither CODIS nor the other genetic databases have been subject to significant acts of fraud or data compromise. The downside risk related to errors or abuse with the current CODIS system is not significantly different than with traditional fingerprints or other investigatory techniques. Unless and until that changes, DNA fingerprinting will likely continue to be seen as an important tool in the search for justice.

CONCLUSION

The clear legislative intent behind DNA fingerprinting is to generate investigative leads and improve the accuracy of the criminal justice system. The collection of the DNA is not overly burdensome or embarrassing, and the data are useful, reliable, and effective. The risk of harm to the innocent is minimal. Moreover, DNA data significantly increase the accuracy of the criminal justice system. The Innocence Project, which uses DNA to try to win the release of the wrongfully convicted, has already helped exonerate over 200 American convicts.[32] Presumably, none of them would have served any time if DNA evidence had been used to find the real culprit at the time of their trial. Concern about the development of a big-brother-type database is legitimate, but that is more a matter of political will than constitutional constraint. As such, the balance between individual privacy and government interest points to the reasonableness of the collection and use of DNA evidence without a judicial warrant.

Endnotes

1 Andrews v. State, 533 So.2d 841 (Fla. Dist. Ct. App. 1988). By 1994, a federal appellate court found that it was appropriate to take judicial notice of DNA testing. United States v. Martinez, 3 F.3d 1191 (8th Cir. 1993), *cert. denied*, 114 S.Ct. 734 (1994).

2 Human Genome Project Information, DNA Forensics, http://www.ornl.gov/sci/techresources/Human_Genome/elsi/forensics.shtml (last visited June 19, 2007).

3 Aubrey Fox, *The Move to Expand DNA Testing*, The Gotham Gazette, May 21, 2007 <<http://www.gothamgazette.com/article/crime/20070521/4/2186>>.

4 Anna Gosline, *Will DNA Profiling Fuel Prejudice?*, New Scientist, April 25, 2005 <<http://www.newscientist.com/channel/opinion/mg18624944.900-will-dna-profiling-fuel-prejudice.html>>.

5 *See, e.g.*, 42 U.S.C. § 14135 (2000); CAL. PENAL CODE § 296 (West 2006); LA. REV. STAT. ANN. § 15:609 (2006); NEB. REV. STAT. § 29-4103 (2006); N.J. STAT. ANN. § 53:1-20.20 (West 2007); TEX. GOV'T CODE ANN. § 411.1471 (Vernon 2006); VA. CODE ANN. § 19.2-310.2:1 (2006).

6 According to the FBI, all fifty states and the U.S. Army cooperate with CODIS <<http://www.fbi.gov/hq/lab/codis/partstates.htm >>

7 DNA Analysis Backlog Elimination Act of 2000, Pub. L. No. 106-546, § 3, 114 Stat. 2726, 2728 (codified as amended at 42 U.S.C. § 14135a).

8 Press Office, Press Release, *Judiciary Committee Adds Kyl DNA Bill to Violence Against Women Act* (September 8, 2005), *available at* http://kyl.senate.gov/record. cfm?id=245432.

9 Several states also collect DNA samples from arrestees who are never charged (*e.g.*, detainees). Ron Scherer, *Should DNA be collected from all criminals?*, The Christian Science Monitor, May 19, 2006.

10 The burden of seeing that the order gets to the FBI would appear to be on the detainee/arrestee. Jonathon Krim, *Bill Would Permit DNA Collection from All Those Arrested*, The Washington Post, September 24, 2005.

11 *A-OK on DNA, ASAP*, The Arizona Republic, Jun. 27, 2007 ("Ten states and the federal government have some requirement for collecting DNA from arrestees. The United Kingdom has taken DNA from arrested suspects for years. "). *See, e.g.,* CAL. PENAL CODE § 296 (West 2006); LA. REV. STAT. ANN. § 15:609 (2006); NEB. REV. STAT. § 29-4103 (2006); N.J. STAT. ANN. § 53:1-20.20 (West 2007); TEX. GOV'T CODE ANN. § 411.1471 (Vernon 2006); VA. CODE ANN. § 19.2-310.2:1 (2006).

12 *See A-OK on DNA, ASAP*, The Arizona Republic, Jun. 27, 2007.

13 The Associated Press, *Governor Backs Collecting DNA from Arrestees*, Tucson Citizen, June 21, 2007.

14 This has led to a controversy between a Colorado District Attorney, who wants more information about a non-perfect DNA match, and the California Attorney General who sees this as a privacy issue. *See* Richard Willing, *DNA 'near matches' spur privacy fight*, USA Today, Aug. 3, 2007, at 3A.

15 *See, e.g.,* U.S. v. Kincade, 379 F.3d 813, 837–38 (9th Cir. 2004); Rise v. Oregon, 59 F.3d 1556, 1560 (9th Cir. 1995); Landry v. Attorney Gen., 709 N.E.2d 1085, 1092 (Mass. 1999).

16 *A-OK on DNA, ASAP*, The Arizona Republic, June 27, 2007.

17 410 U.S. 1 (1973).

18 *Id*. at 14.

19 *A-OK on DNA, ASAP*, The Arizona Republic, Jun. 27, 2007.

20 *Id*.

21 Patrick Haines, Note: *Embracing the DNA Fingerprint Act*, 5 J. on Telecomm. & High Tech. L. (2007) (citing John M. Butler, Forensic DNA Typing: Biology and Technology Behind STR Markers (2001)). At least one commentator has argued that the family relations that can be discerned through DNA presents a problem because entire families may become suspect due to a DNA sample, even though there is no exact match on file. Ronald Bailey, *Criminal Kinship: Slouching Toward a DNA Database Nation*, Reasononline, May 19, 2006.

22 *Haines*, *supra* note 21.

23 *A-OK on DNA, ASAP*, The Arizona Republic, Jun. 27, 2007.

24 *See* Federal Bureau of Investigation – Combined DNA Index System (CODIS) Home Page, http://www.fbi.gov/hq/lab/codis/index1.htm (last visited Apr. 2006) (noting that 27,700 cold hits had been generated from 124,285 crime scene DNA samples as of Nov. 2005).

25 Haines, *supra* note 21.

26 *See*, *e.g.*, S.A. Harbison et al., *The New Zealand DNA Databank: Its Development and Significance as a Crime Solving Tool*, 41 Sci. & Just. 33, 36 (2001) (reporting that 77% of reported database matches in New Zealand originated from burglaries); David Werrett, "The Strategic Use of DNA Profiling," Address to the 18th International Congress of Forensic Haemogenetics (Aug. 19, 1999).

27 *See* National Immigration Law Center, *How Does the Kyl Amendment to the VAWA Reauthorization Bill Affect Immigrants?*, October 2005 <<www.nilc.org>> ("For Immigrants, the provision is stunningly egregious and overreaching. It would cast immigrants – both documented and undocumented – as criminals, requiring them to submit to seizure of their DNA….").

28 Julia Preston, *U.S. Set to Begin a Vast Expansion of DNA Sampling*, N.Y. Times, Feb. 5, 2007.

29 *British police targeting young blacks – report,* The Jamaica Gleaner, June 19, 2007 <<http://www.jamaica-gleaner.com/gleaner/20070619/lead/lead6.html>>. *See also* Anna Gosline, *Will DNA Profiling Fuel Prejudice?*, New Scientist, April 25, 2005 <<http://www.newscientist.com/channel/opinion/mg18624944.900-will-dna-profiling-fuel-prejudice.html>>.

30 *See, e.g.,* Adam Liptak, *You Think DNA Evidence is Foolproof? Try Again*, N.Y. Times, Mar. 16, 2003, at D5.; Jennifer L. Mnookin, *Fingerprint Evidence in an Age of DNA Profiling*, 67 Brook. L. Rev. 13, 49–50 (2001).

31 *Preston*, *supra* note 28.

32 In 1992, attorneys Barry Scheck and Peter Neufeld created a center at the Benjamin N. Cardozo School of Law of Yeshiva University to help those convicts who were able to conclusively prove their innocence with DNA evidence. This evolved into a number of "Innocence Projects" around the nation. This author is currently involved in the establishment of the Mississippi Innocence Project.

Sanctuary Cities and the Second Circuit's Challenge in *New York v. U.S.*

*By Marc M. Harrold**

The existence of "sanctuary cities" in the United States falls in line with the old adage that "all government is local": many cities whose electorate would consistently, even adamantly, favor strict enforcement of immigration laws and border security policies have nonetheless pragmatically enacted local "sanctuary" policies believed to be necessary for the municipal government to provide police services to all its citizens, residents, inhabitants, and visitors.[1]

Sanctuary cities are municipalities that have express ordinances or policies that prohibit local law enforcement officers from inquiring about immigration status generally or reporting this information, if discovered, to federal authorities. Strong arguments for and against this type of policy exist. However, in this article, I will not delve into the substance of these arguments; for my purposes, what is crucial is that, for whatever reason, and regardless of the soundness of the underlying logic, a particular state or municipality has chosen to deem itself a "sanctuary city" thus pitting itself in violation of federal law (explained below).

Most likely, had only a few small communities deemed themselves "sanctuary cities," we would not see the national debate, at times outrage, that has since erupted.[2] But the list of sanctuary cities include three of the most populous cities in the United States: New York, Los Angeles, and Houston, as well as prominent cities like Anchorage, Phoenix, San Diego, San Francisco, Chicago, Baltimore, Detroit, Minneapolis, Albuquerque, Austin, and Seattle.[3] Two states have also enacted sanctuary policies: Alaska in 2003 and Oregon (a pioneer in this regard) in 1987.

Two sections of federal law prohibit "sanctuary city" policies generally and, more specifically, the broad regulation enacted in New York, Executive Order 124. In part, Order 124 "prohibits [New York employees] from voluntarily providing federal immigration authorities with information concerning the immigration status of any alien."[4] In 1996, two federal laws were enacted that contained provisions at odds with Order 124,[5] eventually resulting in the Second Circuit Court of Appeals' ruling in *New York v. U.S.*[6]

Section 434 of the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 ("Welfare Reform Act")[7]:

> Notwithstanding any other provision of Federal, State, or local law, no State or local government entity may be prohibited, or in any way restricted, from sending to or receiving from the Immigration and Naturalization Service information regarding the immigration status, lawful or unlawful, of an alien in the United States.

** Marc M. Harrold is Senior Counsel at the National Center for Justice and the Rule of Law and Visiting Professor of Law at the University of Mississippi. He is a former City of Atlanta Police Officer. Those interested in immigration issues should read Margaret Stock's review of The Founders on Citizenship and Immigration in the Book Reviews section.*

Section 642 of the Illegal Immigration Reform and Immigrant Responsibility Act of 1996 ("Immigration Reform Act")[8]:

> (A) IN GENERAL.—Notwithstanding any other provision of Federal, State, or local law, a Federal, State, or local government entity or official may not prohibit, or in any way restrict, any government entity or official from sending to, or receiving from, the Immigration and Naturalization Service information regarding the citizenship or immigration status, lawful or unlawful, of any individual.
>
> (B) ADDITIONAL AUTHORITY OF GOVERNMENT ENTITIES.—Notwithstanding any other provision of Federal, State, or local law, no person or agency may prohibit, or in any way restrict, a Federal, State, or local government entity from doing any of the following with respect to information regarding the immigration status, lawful or unlawful, of any individual:
>
> (1) Sending such information to, or requesting or receiving such information from, the Immigration and Naturalization Service.
>
> (2) Maintaining such information.
>
> (3) Exchanging such information with any other Federal, State, or local government entity.
>
> (C) OBLIGATION TO RESPOND TO INQUIRIES.—The Immigration and Naturalization Service shall respond to an inquiry by a Federal, State, or local government agency, seeking to verify or ascertain the citizenship or immigration status of any individual within the jurisdiction of the agency for any purpose authorized by law, by providing the requested verification or status information.

The Second Circuit expressly acknowledged that their holding was in light of the fact that the City of New York was making a facial challenge to Sections 434 and 642 noting:

> [a] facial challenge to a legislative Act is…the most difficult challenge to mount successfully, since the challenger must establish that no set of circumstances exists under which the Act would be valid.[9]

The Second Circuit distinguished the Sections 434 and 642 from the federal programs at issue in *Printz v. United States*[10] and *New York v. United States*[11]:

> Unlike Sections 434 and 642, the federal programs in *Printz* and *New York* conscripted states (or their officers) to enact or administer federal regulatory programs. *See Printz*, 117 S. Ct. at 2376 (distinguishing federal directives to states that "require only the provision of information to the Federal Government" from those that "force [ ] [the] participation of the States' executive in the actual administration of a federal program," even though both kinds of directive leave states with no "choice" but to comply). The central teaching of these cases is that "even where Congress has the authority under the Constitution to pass laws

requiring or prohibiting certain acts, it lacks the power directly to compel the States to require or prohibit those acts." *New York*, 505 U.S. at 166, 112 S. Ct. 2408. Congress may not, therefore, directly compel states or localities to enact or to administer policies or programs adopted by the federal government. It may not directly shift to the states enforcement and administrative responsibilities allocated to the federal government by the Constitution. Such a reallocation would not only diminish the political accountability of both state and federal officers, *see New York*, 505 U.S. at 168, 112 S. Ct. 2408; *Printz*, 117 S. Ct. at 2382, but it would also "compromise the structural framework of dual sovereignty," *Printz*, 117 S. Ct. at 2383, and separation of powers, *see id.* at 2378 ("[T]he power of the President would be subject to reduction, if Congress could act as effectively without the President as with him, by simply requiring state officers to execute its laws.")....

In the case of Sections 434 and 642, Congress has not compelled state and local governments to enact or administer any federal regulatory program. Nor has it affirmatively conscripted the states, localities, or their employees into the federal government's service. These Sections do not require or prohibit anything. Rather, they prohibit state and local governmental entities or officials only from directly restricting the voluntary exchange of immigration information with the INS. *See Printz*, 117 S. Ct. at 2376.[12]

The Question Left Open in *New York v. U.S.*

The Second Circuit, after rejecting the Tenth Amendment sovereignty argument and "republican form of government claim," seemingly threw down the gauntlet and invited New York (or some other sanctuary city) to challenge §§ 434 and 642 "as applied" (not facially) and instead on the affect these provisions of federal law have on the "performance of legitimate municipal functions."[13] This article focuses on police services as "a legitimate municipal function."

Confidentiality and Police Services: What the Cities Fear

The City's concerns are not insubstantial. The obtaining of pertinent information, which is essential to the performance of a wide variety of state and local governmental functions, may in some cases be difficult or impossible if some expectation of confidentiality is not preserved.[14]

Order 124, originally issued by Mayor Ed Koch in 1989, was re-issued by Mayor Rudy Giuliani. It seems highly unlikely that the New York City government sought to create a haven for those in disregard of national immigration laws. The primary reason that cities have enacted policies forbidding the police from inquiring into immigration status is to foster an atmosphere of trust between the local police and the immediate community they serve.[15] Intuitively, we can imagine a plethora of situations where crime victims and witnesses might be reluctant or afraid to communicate with police if they expect the police to inquire into their immigration status and turn them over to the Department of Homeland Security.

An incident in Houston provides an example of the legitimacy of local and state law enforcement's concern:

In July 2002, three people were shot and killed inside a Vietnamese restaurant in Houston. Most of the witnesses fled the scene immediately. Due to their fear of being implicated and the fact that many did not have legal status in the U.S., the witnesses were not willing to talk to police. A police officer from the Vietnamese community asked a local Vietnamese language radio program to interrupt a popular program and let him speak to the Vietnamese community. After assuring witnesses, on the air, that the police only wanted information about the shooting, and not the immigration status of witnesses, more than five witnesses came forward.

Professor David A. Harris, Balk Professor of Law and Values at the University of Toledo College of Law, testified before Congress:

If local police are forced to become de facto immigration agents, people in their neighborhoods will simply stop talking to them. They will fear officers and hide from them, instead of communicating with them about the problems, the issues, and the wrongdoers in their neighborhoods. Even worse, when they are victims of crimes, they will fear reporting the offenses. This can lead only to increased fear and less safe streets, as predators exploit this fear and repeatedly prey on not only immigrants, but anyone in these neighborhoods.[16]

A final example comes from the statement of Joseph Estey, Chief of Police in Hartford, VT and former President of the International Association of Chiefs of Police (IACP):

Many leaders in the law enforcement community have serious concerns about the chilling effect any measure of this nature [including requiring non-federal police to inquire into immigration status] would have on legal and illegal aliens reporting criminal activity or assisting police in criminal investigations. This lack of cooperation [between police and the immigrant community] could diminish the ability of law enforcement agencies to police effectively their communities and protect the public they serve.[17]

Can §§ 434 and 642 survive "a constitutional challenge in the context of generalized confidentiality policies that are necessary to the performance of legitimate municipal functions"?

This is the question posed by the Second Circuit in its ruling in *New York v. U.S.* The "legitimate municipal function[s]" at issue for our purposes here is the execution of police services. The "context of generalized confidentiality policies that are necessary" for the performance of this "function" is the non-federal government's interest in fostering trusted communication between its representatives and the immigrant community they are responsible to serve. Obviously, this type of trust and communication will be diminished if police inquire into the immigration status of victims and witnesses and then forward this information to the federal authorities.

Specifically, New York argued that the federal laws at issue caused "disrupt[tion] [of] the actual operation of state and local government."[18] The Second Circuit was clear that its ruling only considered the facial challenge and not any "as applied" challenge under the Tenth Amendment:

Nevertheless, the City has chosen to litigate this issue in a way that fails to demonstrate an impermissible intrusion on state and local power to control information obtained in the course of official business or to regulate the duties and responsibilities of state and local government employees. On the present record, the only state and local policy proffered by the City as disrupted by Sections 434 and 642 is the Executive Order and that Order alone.[19]

Clearly, in certain instances, a state or local government could "demonstrate" that §§ 434 and 642 ("as applied") do cause "an impermissible [federal] intrusion" on state and local governments.

CONCLUSION

I conclude that the Second Circuit was correct in not invalidating the federal laws at issue (that these federal laws, on their face, did not "commandeer" the NYC government); however, the question left open in *New York v. U.S.* should be answered in opposite fashion: that, due to "generalized confidentiality policies that are necessary to the performance of legitimate municipal functions that include federal immigration status," the federal laws at issue do violate the Tenth Amendment to the United States Constitution and cannot be, consistent with tenets of federalism, allowed to stand.[20]

When examined "as applied," the federal government should be constitutionally barred from enacting provisions such as §§ 434 and 642. As such, cities should be able to designate themselves as "sanctuary cities." The degree of "sanctuary," (with respect to citizenship and immigration status as discussed herein) a particular local or state entity chooses to offer should be determined by the local or state government in light of their compelling interest in providing police services and protection to all of the people within their respective jurisdiction and the equally compelling need to foster an environment of trust between the police and the community they serve.[21]

Endnotes

1 Although other types of services are at issue, and affected, by local sanctuary policies, this article will focus on municipal police services.

2 An example of this "outrage" can be seen in popular FOX News personality Bill O'Reilly's television show, website, and blog. (O'Reilly's blog used as a resource to locate and determine the existing "Sanctuary" cities. *See* www.billo reilly.com/blog?action=viewBlog&blogID=88083313171143568 (last visited April 20, 2007)). Further, Congressman Tom Tancredo has received acclaim from certain segments of the American population for highlighting this issue in his bid for the 2008 Republican Presidential nomination.

3 The remaining "sanctuary cities" as of August 2006 include: Fairbanks, AK; Chandler, AZ; Fresno, CA; National City, CA; Sonoma County, CA; Evanston, IL; Cicero, IL; Cambridge, MA; Orleans, MA; Portland, ME; Takoma Park, MD; Ann Arbor, MI; Newark, NJ; Trenton, NJ; Durham, NC; Aztec, NM; Rio Arriba, County, NM; Sante Fe, NM; Ashland, OR; Gaston, OR; Marion County, OR; Katy, TX; Virginia Beach, VA; and Madison, WI.

4 New York v. United States, 179 F.3d 29, 31 (2d Cir. 1999).

5 Executive Order 124 provides in pertinent part:

> Section 2. Confidentiality of Information Respecting Aliens.
>
> a. No City officer or employee shall transmit information respecting any alien to federal immigration authorities unless
>
> (1) such officer's or employee's agency is required by law to disclose information respecting such alien, or
>
> (2) such agency has been authorized, in writing signed by such alien, to verify such alien's immigration status, or
>
> (3) such alien is suspected by such agency of engaging in criminal activity, including an attempt to obtain public assistance benefits through the use of fraudulent documents.
>
> b. Each agency shall designate one or more officers or employees who shall be responsible for receiving reports from such agency's line workers on aliens suspected of criminal activity and for determining, on a case by case basis, what action, if any, to take on such reports. No such determination shall be made by any line worker, nor shall any line worker transmit information respecting any alien directly to federal immigration authorities.
>
> c. Enforcement agencies, including the Police Department and the Department of Correction, shall continue to cooperate with federal authorities in investigating and apprehending aliens suspected of criminal activity. However, such agencies shall not transmit information respecting any alien who is the victim of a crime.

6 179 F.3d 29 (2d Cir. 1999).

7 Pub. L. No. 104-193, 110 Stat. 2105 (1996). The "Immigration and Naturalization Service" referred to has been replaced by the Immigration and Customs Enforcement, an agency of the new Department of Homeland Security. *See* http://www.ice.gov/about/index.htm.

8 Pub. L. No. 104-208, 110 Stat. 3009 (1996). The "Immigration and Naturalization Service" referred to has been replaced by the Department of Homeland Security.

9 *New York*, 179 F.3d at 33 (quoting United States v. Salerno, 481 U.S. 739, 745 (1987)).

10 521 U.S. 898 (1997).

11 505 U.S. 144 (1992).

12 *New York*, 179 F.3d at 34-35. ("We therefore hold that states do not retain under the Tenth Amendment an untrammeled right to forbid all voluntary cooperation by state or local officials with particular federal programs.")

13 *Id.* at 37 ("[W]e find on this record that §§ 434's and 642's interference with the city's Executive Order is entirely permissible and in no way alters the form of New York City's government.").

14 *Id.* at 36.

15 Obviously, other pragmatic and political reasons could also exist that spurs a particular governmental entity to enact "sanctuary" provisions.

16 Testimony of Prof. David A. Harris, Senate Judiciary Subcommittee on Immigration, CLEAR Act Hearings (April 22, 2004).

17 InfoNet Doc. No. 04120163 (AILA)

18 *New York*, 179 F.3d at 34.

19 *Id.* at 36.

20 Generally speaking, enforcement of immigration laws has been seen exclusively as a role of the federal government; further the judiciary has seen this authority as a "plenary power" and has rarely interfered through judicial edict. *See, e.g.,* Takahashi v. Fish and Game Commission, 334 U.S. 410, 419 (1948).

> The Federal Government has broad constitutional powers in determining what aliens shall be admitted to the United States, the period they may remain, regulation of their conduct before naturalization, and the terms and conditions of their naturalization.

Id.

21 As indicated earlier in article, my only contention is that "as applied" §§ 434 and 642 could (upon the proper record) violate the Tenth Amendment and tenets of federalism. Simply because a city or state government *could* enact laws designating themselves as "sanctuary" cities (or states) does not mean they necessarily *should*. The merit of this type of designation is outside the scope of this article.

Making Prisons Compete:
How Private Prisons Enhance Public Safety and Performance

*By Geoffrey F. Segal & Alexander McCobin**

The American public is growing increasingly frustrated with the escalating cost of managing our vast prison system. Facilities are experiencing increased failures in safety, education, and health and training services. There is rising concern over the rates at which inmates are returning to jail, rather than successfully reentering the community. Institutional barriers and resistance to change among existing, public, bureaucratic management cultures have prevented correctional systems from adopting the most effective management practices and inmate rehabilitation programs.

Furthermore, spending on corrections continues to climb, eating up more tax dollars and limited resources. In many ways, spending on corrections prevents public investments in other critical infrastructure, such as transportation and water; both of which need hundreds of billions, if not trillions, in investment over the next twenty years. State governments spent $42.9 billion on corrections in 2005, and it is estimated that federal and state governments will need as much as $27 billion in new spending over the next five years. Additionally, spending is estimated to increase by 6.5% over the Fiscal 2005 level.[1]

Percentage of State Government Spending on Corrections (All Sources)[2]	
2001	3.7
2002	3.6
2003	3.5
2004	3.5
2005	3.5
2006	3.6

Percentage of State Government General Fund Spending on Corrections[3]	
2001	7
2002	6.9
2003	7
2004	7
2005	7.2
2006	7

While corrections remain a fairly small percentage of total government spending (see Table x and y), the size of state spending has reached a level where any growth in spending on corrections represents a substantial nominal sum. Over the past five years nominal state spending year to year has experienced significant growth, including an estimated 8.1% increase in Fiscal 2006.[4]

Exacerbating this problem is the fact that larger numbers of people are being sent to prison. At the end of June 2006, 2,245,189 inmates were incarcerated in federal, state, or local facilities; representing an increase of nearly 3% over the previous year.[5] Unless sentencing reform takes hold quickly, incarceration numbers will only go up, driving costs higher, and placing further pressure on already limited government budgets.

Indeed, prisons have become part of the problem. As commentators have long recognized, prisons provide little in the way of rehabilitation, serving largely as warehouses for criminals, many of whom return soon enough to American streets. More alarmingly, some evidence suggests that many prisons serve as training grounds for criminals where younger, less experienced, and less violent prisoners learn the tools of the trade from more hardened criminals.[6]

Speaking of the Commission on Safety and Abuse in America's Prisons, U.S. Senator Tom Coburn correctly noted that "the experiences inmates have in prison—whether violent or redemptive—do not stay within prison walls, but spill over into the rest of society. Federal, state, and local governments must address the problems faced by their respective institutions and develop tangible and attainable solutions."[7] For these reasons—the prison system's fiscal burden, the growth of the prison population, and the abject failure of the prison system to enhance public safety—reform is urgently needed.

I. A Brief History of Private Prisons

Private prisons are not new to the U.S., or the world, for that matter. The first private prison opened in 1985. Since then, some thirty-four states and the federal government have begun contracting with private companies.[8] The United Kingdom has an official policy that all new prisons will be commissioned from the private sector. Germany, France, Japan, Israel, Brazil, and Netherlands, to name a few countries, have all opened in recent years, or will open, new private prisons in the not-too-distant future.[9]According to the Association of Private Correctional and Treatment Organizations, there are more than 250 private facilities, or more than 157,000 functioning beds under private operation—representing 7% of the U.S. prison population. While the growth has slowed, private prisons have continued to enjoy a modest increase in number of facilities and rated capacity.

II. Economics of Private Prisons

The economics of private prisons are simple: private prisons save money. A comprehensive review of the privatization literature by the Reason Foundation examined twenty-eight research reports that compared cost data for private prisons to government-operated facilities. Of those studies, twenty-two (79%) found significant budget savings, conservatively estimated to be between 5 and 15%, due to privatization.[10]

** Geoffrey F. Segal is the Director of Government Reform at the Reason Foundation, and was recently appointed to the Commonwealth Competition Council in Virginia. Alexander McCobin is a Research Assistant at the Reason Foundation.*

More importantly, private prisons have a competitive effect that helps control corrections spending. Two recent studies highlight the importance of competition in the field of corrections. Two professors from Vanderbilt University found that the use of private prisons in a state resulted in the reduction of daily incarceration costs for the public corrections system by 4.45% annually.[11] The Rio Grande Foundation in New Mexico compared per-prisoner department of corrections budgets across forty-six states. By measuring an entire department's spending rather than just a particular prison's spending, the study accounts for the cost savings public prisons can achieve in response to private competition. The study uses the percentage of prisoners under private management as its measurement of the extent of private prisons in each state.

Holding other factors constant, this study found that states with 5% of their prison population in private prisons spent about $4,804 less per prisoner in 2001 than states without any private prisons. As the extent of private participation increases, so do savings.[12] New Mexico, for example, has 45% of its prison population under private management; it spent $9,660 less per prisoner in 2001 than did counterpart states without any private prisons. New Mexico has gone farther down the private prison road than any other state, saving $51 million in 2001 alone, according to the Rio Grande study.[13]

This experience is only solidified with real world experience and data. As one of the largest users of private corrections, the state of Texas provides us a unique look at the long-term benefits of competition to a correctional system. For more than a dozen years the Texas Criminal Justice Policy Council and now the Texas Legislative Budget Board have conducted a biannual review of the average cost per day of government facilities and the average contract price at private facilities. The first study was published in 1991. Subsequent studies have been conducted every other year since, with the latest published in 2007.[14] This data represents the best longitudinal evidence of the benefits of competition. During that sixteen-year period, Texas' in house per diem cost has gone down, *i.e.,* they have gotten more efficient—on both the contracts themselves and avoiding costs—because competition has made them.

In addition, a wealth of studies performed by government agencies, universities, auditors, and research organizations have examined the relative quality of private prisons compared with government-run prisons. A 2002 Reason Foundation paper reviewed all available studies comparing public and private facilities.[15] Seventeen studies were identified that used various approaches to measure the relative quality of care at correctional facilities managed by government versus private firms. Fifteen of the studies examined by the Reason Foundation demonstrate that quality at private facilities is as-good or better than at government-run facilities.

The major charge against privatization is that, by reducing costs, quality and security are sacrificed. Yet, there is clear and significant evidence, as demonstrated by more than a dozen comparison studies, that private facilities provide at least the level of service that government-run facilities do. Private correctional facilities have fared well against government-run facilities in almost all measures of quality, including a wide range of quality comparison studies.

III. Legal Issues and Private Prisons

With the increasing use of private prisons both domestically and abroad, the legal and judicial issues affecting private prisons in the U.S. today are increasingly important. The first legal issue private prisons face is simply whether they are legal. Federal, state, and local officials have all recognized the need for legal authority to delegate correctional responsibilities to non-governmental entities. It is possible to define imprisonment as a uniquely governmental function that cannot be delegated. However, this interpretation is rare. The responsibility for *sentencing* individuals to be confined is certainly a purely governmental function, but the mechanics of holding someone in confinement are not.

At the federal level, this is recognized in the language of 18 U.S.C. Sec. 4082(b), which remands all federal offenders to confinement in "any available, suitable, and appropriate institution or facility, whether maintained by the Federal Government or otherwise." The Bureau of Prisons has interpreted this to mean there is authority to contract with private prisons.

State and local governments deal with the legal authority to contract for correctional services in their own ways. States that currently have a private prison in operation or under construction obviously have legal authority to do so, as have a number of states that do not yet have a private facility. One very common method state and local governments use to assure legal authority is to pass enabling legislation. Others seek a determination by the state attorney general that there is no law forbidding contracts for private prisons.

Perhaps the second most important legal issue concerns inmate treatment and rights. Court rulings in *Richardson v. McKnight*[16] and *Correctional Services Corp. v. Malesko*[17] have held private prisons to at least as high a standard as public prisons (at least in the context of qualified immunity). Put simply, lawsuits may be brought against private prison guards by inmates. However, inmates are not able to sue the private prison for violations of their civil rights.[18]

A more recent legal issue has developed as states try to cope with overcrowding and escalating costs. States like California, which lack explicit authority to operate private prisons in-state, send, or attempt to send, inmates to private prisons in other states. The conditions inside California's prison system illustrate the important role private prisons play.

The California Department of Corrections and Rehabilitation has exceeded its maximum inmate capacity, to the point where its prison system housed 172,000 inmates for a 100,000 inmate design.[19] More than 15,000 inmates have been placed in prison areas never designed for housing including gymnasiums, dayrooms, and program rooms. In addition, 4,000 state correctional officer vacancies are not expected to be filled for at least five years, exacerbating security threats related to overcrowding.[20] A trial court in California found that "[p]rison overcrowding in California is a crisis creating conditions of extreme peril to the safety of persons and property within the state…"[21] In the words of U.S. District Court Judge Lawrence Karlton, "people are dying" because of the state of California prisons.[22]

Owing to this severe overcrowding, Governor Arnold Schwarzenegger declared an emergency in twenty-nine California prisons on October 4, 2006. As a temporary solution to alleviate the housing problems, Governor Schwarzenegger chose to transfer inmates to out-of-state prison facilities. Two contracts with out-of-state, private prison facilities, one with The Geo Group, Inc. ("GEO") and the other with the Corrections Corporation of America (CCA), were entered into on October 19, 2006. In response, on October 30, 2006, the prisons guard union (California Correctional Peace Officers' Association) filed a lawsuit to prevent the inmate transfer. On April 2, 2007, the trial court ruled that the Governor exceeded his authority in ordering the transfer of inmates, and so mandated he revoke his order, and excluded any contracts to transfer inmates to out-of-state prisons. In part, the court ruled that the contracts violated the California Constitution, which prohibits the hiring of non-civil servants for such positions.[23] While exceptions to this constitutional provision have been made for the sake of emergency situations, the same trial court that declared the overcrowding to be a "crisis" did not find that the current case qualified under the requirements. However, upon appeal to the California Court of Appeals, the trial court's order was stayed, allowing for the transfer of inmates out-of-state pending a decision on the appeal.[24] As of July, the state has begun to transfer inmates to out-of-state private prison facilities.[25]

The future of Governor Schwarzenegger's decision to outsource prison services to out-of-state private facilities is uncertain, but the case raises questions surrounding each legal question. While California prohibits the use of private prisons except for emergencies, the need for private prison facilities to alleviate an obvious emergency in the state is indicative of a wider problem with prohibition of private prison services. Without outsourcing to these prisons, California would have no way of achieving basic constitutional standards for inmates. As well, California has made a specific distinction between public employees and employees of private facilities performing government services. Most importantly, the humanitarian threats to inmates living in California correctional facilities has prompted the need to look to private facilities as alternative remedies that provide better quality-of-life.

CONCLUSION

Significant evidence demonstrates that private prisons do save money. It is remarkable that such a wide variety of approaches spanning over a decade and half of research conducted in states across the nation repeatedly demonstrate time and again that privatization does not reduce quality.

One might argue that, at the end of the day, citizens do not care who is providing a service as long as it is being provided effectively. Taxpayers ultimately care about results and performance; it matters not whether a private or public employee does the work. In order to generate genuine discussions about improving performance in state and local correctional systems, policymakers need to get away from ideology and partisanship. Policy debates need to focus on results, performance, and achieving the best outcome with the limited resources available. Failing programs, whether public or private, should be halted in favor of better-performing programs. The debate should move away from public versus private and toward performing versus non-performing.

Neither a fully public nor a fully private corrections system is necessarily ideal. But, by introducing competition and focusing on outcomes and results, governments can provide the greatest incentives to innovate, and provide the high-quality services citizens expect for their tax dollars. Regardless of whether public agencies or private companies ultimately win the contract, the threat of losing "business" to competitors ensures that operators do their best to meet or exceed predetermined benchmarks.

Competition between public agencies and private companies provides greater accountability and incentives to achieve than the status quo bureaucratic approach. In the event of a contract, agencies may enhance these incentives for maximum effect by including provisions for performance bonuses and penalties, and even cancellation (for failure to achieve) for every institution, regardless of whether it is publicly or privately run. Competition, in conjunction with performance-based budgeting, is the best means of ensuring that correctional managers will continue to seek improved prison performance, reduced recidivism rates, and the best uses for taxpayer dollars.

Ultimately, focusing on performance and goals enables elected officials, the public, and government officials to know that they are receiving the maximum value for the funds appropriated to corrections. Governments will be forced to become more efficient, and taxpayers will benefit from fewer repeat offenders, reduced crime, and reduced criminal justice system costs.

Endnotes

1 Nat'l Ass'n of State Budget Officers, 2005 State Expenditure Report 56 (2006), *available at* http://www.nasbo.org/Publications/PDFs/2005%20State%20Expenditure%20Report.pdf.

2 *See generally* Pub. Safety Performance Project, The Pew Charitable Trusts, Public Safety, Public Spending: Forecasting America's Prison Population 2007-2011 (2006), *available at* http://www.pewpublicsafety.org/pdfs/PCT_CorrectionsReport_3-1.pdf.

3 *See generally id.*

4 Nat'l Ass'n of State Budget Officers, *supra* note 1, at 56.

5 Bureau of Justice Statistics, U.S. Dep't of Justice, *Prison Statistics*, *available at* http://www.ojp.usdoj.gov/bjs/prisons.htm (last visited Aug. 14, 2007).

6 *Juvenile Crime, Juvenile Justice,* Comm'n on Behavioral and Social Sciences and Education (Joan McCord, et al., eds.) (2001). *Available at* http://www.nap.edu/openbook.php?isbn=0309068428

7 Commission on Safety and Abuse in America's Prisons, *available at* http://www.prisoncommission.org (last accessed September 1, 2007).

8 P. Harrison & A. Beck, U.S. Dep't of Justice, Bureau of Just. Stats. Prisoners in 2004. (U.S. Government Printing Office, Oct. 2005).

9 Segal, G. (February 2006). Unpublished letter to the Salt Lake Tribune Editor. Reason Foundation.

10 Geoffrey F. Segal & Adrian T. Moore, Weighing the Watchmen: Evaluating the Costs and Benefits of Outsourcing Correctional Services—Part 1: Employing a Best-Value Approach to Procurement; Part II: Reviewing the Literature on Costs and Quality Comparisons (January 2002). Policy Studies 289 & 290. Reason Public Policy Institute, *available at* http://www.reason.org/policystudiesbydate.shtml#2002.

11 James F. Blumstein & Mark A. Cohen, The Interrelationship Between Public and Private Prisons: Does the Existence of Prisoners under Private Management Affect the Rate of Growth in Expenditures on Prisoners under Public Management, Ass'n of Private Corr. and Treatment Organizations (April 2003).

12 Matthew Mitchell, *The Pros of Privately-Housed Cons: New Evidence on the Cost Savings of Private Prisons*, Rio Grande Foundation 1 (Tijeras, New Mexico: March 2003).

13 *Id*.

14 Texas Criminal Justice Policy Council, *Texas Correctional Costs 1989-1990* (January 1991); Texas Criminal Justice Council, *Texas Correctional Costs per Day 1991-1992* (March 1993); Texas Criminal Justice Policy Council, *Texas Correctional Cost Per Day, 1993-94* (February 1995); Texas Criminal Justice Policy Council, *Apples to Apples: Comparing the Operational Costs of Juvenile and Adult Correctional Programs in Texas* (January 1997); Texas Criminal Justice Policy Council, *Oranges to Oranges: Comparing the Operational Costs of Juvenile and Adult Correctional Programs in Texas* (January 1999); Texas Criminal Justice Policy Council, *Limes to Limes: Comparing the Operational Costs of Juvenile and Adult Correctional Programs in Texas* (January 2001); Texas Criminal Justice Policy Council, *Mangoes to Mangoes: Comparing the Operational Costs of Juvenile and Adult Correctional Programs in Texas* (January 2003); Texas Legislative Budget Board, *Criminal Justice Uniform Cost Report Fiscal Years 2003-2004* (January 2005); Texas Legislative Budget Board, *Criminal Justice Uniform Cost Report Fiscal Years 2004-2006* (January 2007).

15 Segal & Moore, *supra* note 8, part 2, *available at* http://www.reason.org/ps290.pdf.

16 521 U.S. 399 (1997). *See id.,* at 401, 412 (holding that private prison guards cannot claim qualified immunity in §1983 suits).

17 534 U.S. 61, 66 (2001) (holding that companies operating private correctional facilities, like agencies operating analogous public facilities, are not subject to civil rights suits under *Bivens v. Six Unknown Named Agents of Federal Bureau of Narcotics*, 403 U.S. 388 (1971)).

18 *Id*.

19 Associated Press, *California's Overcrowded Prisons Subject of Court Hearing*, SFGate.com, June 7, 2007, *available at* http://sfgate.com/cgi-bin/article.cgi?f=/n/a/2007/06/07/state/n174816D90.DTL.

20 Edmund G. Brown, Jr., et al., Petition for Writ of Supersedeas or Other Stay Order; Memorandum of Points and Authorities, Apr. 3, 2007, 4.

21 *Id*. at 1.

22 Andy Furillo, *Hearing Looks at Limit on Inmates*, The Sacramento Bee, June 28, 2007, *available at* http://www.sacbee.com/111/story/245775.html. (last visited Aug. 16, 2007).

23 Cal. Const. art. 7 § 7 ("A person holding a lucrative office under the United States or other power may not hold a civil office of profit.")

24 Cal. Corr. Peace Officers' Ass'n v. Arnold Shwarzenegger, (Cal. Dist. Ct. App. 3 filed Apr. 12, 2007).

25 By June 14, 400 inmates were voluntarily transferred to facilities in Arizona and Tennessee. *Controversy Continues Over Moving Prisoners Across State Lines* (NBC11 broadcast in Sacramento, Cal. June 14, 2007), *available at* http://www.nbc11.com/news/13507931/detail.html.

Can the Golden State Catch A Greenhouse Waiver?

*By Jonathan H. Adler**

In 2004, the California Air Resources Board (CARB) adopted the nation's first regulations limiting the emission of greenhouse gases from new motor vehicles. The California rules require automakers to reduce greenhouse gas emissions from new motor vehicles sold in the state, beginning with the 2009 model year. By 2016, new vehicle emissions must decline by nearly 30 percent. Several other states have announced their plans to follow California's lead, adopting the regulations as their own. Before these regulations can take effect, however, California must obtain a waiver of preemption under the Clean Air Act (CAA) from the U.S. Environmental Protection Agency (EPA). California politicians demand such a waiver forthwith, and EPA Administrator Stephen Johnson has promised a decision on the request by the end of 2007.

Many assume California is entitled to a waiver, and only Bush Administration intransigence stands in the way. Testifying before the Senate Committee on the Environment and Public Works in May 2007, California Attorney General Edmund G. Brown, Jr. declared "If EPA follows the law, there's no question that it must grant California's waiver."[1] California has sought and obtained numerous waivers of CAA preemption in the past without much trouble. Why should this case be any different? Because this would be the first waiver authorizing state regulation of vehicular emissions of greenhouse gases. EPA approval of California's waiver request is not—and should not be—automatic. A CAA waiver for greenhouse gas emission reductions raises distinct legal and policy issues. In prior cases, California sought permission to adopt more stringent pollution controls to facilitate its efforts to combat urban smog and other localized pollution problems. Now California seeks to reduce emissions that contribute to global warming. The problem California seeks to address with these rules is not localized, or even national; global climate change is a *global* phenomenon. As a consequence, it is not entirely clear that the EPA is required, or even permitted, to grant California the waiver it wants. While the Bush Administration may face tremendous political pressure to grant the Golden State its wish, there are grounds to pause before assuming it is legally obligated to do so.

This article provides an overview of the legal and policy issues raised by California's request for a waiver of federal preemption of its new greenhouse gas emission regulations. After summarizing the legal requirements for obtaining a waiver of preemption under the CAA, this article explains why it may be more difficult for California to obtain a waiver for greenhouse gas emission regulations than it has been for prior state regulations governing traditional air pollutants. The article then considers how California's waiver request fits into a broader policy framework dividing responsibility for environmental protection between the federal and state governments. Accepting there are strong arguments for greater state flexibility in environmental law, this article assesses the relative strength of California's demand for greater freedom to set its own greenhouse gas emission control policies.

Vehicle Emission Controls under the Clean Air Act

The CAA establishes a baseline of uniform emission controls for new motor vehicles. Section 209(a) of the CAA provides that no state may adopt or enforce "any standard relating to the control of emissions from new motor vehicles or new motor vehicle engines" subject to regulation under the Act.[2] The purpose of this provision is to maintain a national market for motor vehicles by providing for uniformity in vehicle emission standards. Any automobile that rolls off an assembly line meeting federal emission control requirements should be able to be sold anywhere in the United States. A uniform national standard prevents the balkanization of the national automobile market that could result if automakers were required to design and sell different vehicles in different states.[3] Indeed, the major automakers supported adoption of federal vehicle emission controls in order to preempt the proliferation of variable state standards.[4]

The CAA contains one exception to this general policy of preemption. Recognizing California's particularly severe air pollution problems, and the Golden State's pioneering efforts to control mobile source air pollution—efforts that predated adoption of vehicle emission controls at the federal level—Congress adopted Section 209(b), authorizing a waiver of preemption for California.[5] This provision effectively grandfathered California's pre-existing emission controls and authorized a potential exemption for additional emissions controls adopted there in the future.[6] Once the EPA grants a waiver, other states are permitted to adopt California's regulations as a part of their own air pollution control programs under CAA Section 177, but they are never allowed to adopt vehicle emission standards of their own.[7] As a consequence, there can never be more than two sets of vehicle emission standards—those set by the EPA, and those set by California.

California's Greenhouse Gas Emission Controls

In 2002, the California state legislature enacted Assembly Bill 1493, directing the CARB to "develop and adopt regulations that achieve the maximum feasible and cost-effective reduction of greenhouse gas emissions from motor vehicles."[8] In September 2004, CARB approved regulations amending its existing "Low-Emission Vehicle (LEV II)" program to establish declining fleet average greenhouse gas emission standards. The regulations apply to four greenhouse gases: carbon dioxide, methane, nitrous oxide, and hydrofluorocarbons, but are enforced by reference to the carbon-dioxide equivalent of a

....

** Jonathan H. Adler is a Professor of Law and the Director of the Center for Business Law & Regulation at Case Western Reserve University School of Law.* *This essay is based upon testimony given before the Senate Committee on Environment and Public Works, "Examining the Case for the California Waiver," May, 22, 2007. He would like to thank Tai Antoine and James Weikamp for their research assistance.*

vehicle's emissions. Under CARB's regulations, the emission standards take effect beginning with the 2009 vehicle model year, and decline in subsequent model years through 2016.

California adopted its greenhouse gas regulations out of a stated concern about the state's contribution and vulnerability to global climate change. California is responsible for a substantial share of U.S. greenhouse gas emissions, and over half of the state's emissions come from transportation. "Climate is a central factor in California life," according to CARB, and "climate change threatens California's public health, water resources, agricultural industry, ecology, and economy."[9]

California policymakers assumed that CARB would apply for, and the EPA would grant, a waiver of preemption under CAA Section 209(b). Accordingly, CARB submitted a waiver request to the EPA in December 2005. Up until now, the EPA has never completely denied a waiver request under Section 209(b).[10] Up until now, however, California had always sought waivers for emission control measures that related to the state's efforts to control its notoriously severe local air pollution problems.

When California adopted its greenhouse gas regulations, there was some legal uncertainty as to whether a waiver request was even possible. At the time, the EPA denied that it had any authority to regulate automotive emissions of greenhouse gases, so there was no basis upon which the EPA could approve a waiver request. The Supreme Court's recent decision in *Massachusetts v. EPA* resolved some of this uncertainty, holding that greenhouse gases were "pollutants" subject to regulation under the CAA.[11] The EPA must now reconsider its refusal to adopt federal regulations limiting vehicular greenhouse gas emissions—and such rules are almost certain to be adopted.[12] If the EPA may regulate greenhouse gases, there is no question that California must obtain a waiver from the EPA if it is to maintain and enforce its greenhouse gas emission regulations—but it does not mean that a waiver must be forthcoming. Following the *Massachusetts* decision, however, the EPA announced it would formally consider California's waiver request.[13]

Waivers of Preemption under Section 209(b)

The CAA waiver provision is not a blank check. Section 209(b) imposes some limitations on the EPA's authority to approve a waiver of preemption for California's vehicle emission standards.[14] EPA review of a California waiver request is fairly deferential. "California's regulations, and California's determination that they comply with the statute... are presumed to satisfy the waiver requirements and [] the burden of proving otherwise is on whoever attacks them."[15] Nor does the EPA have any authority to consider criteria beyond those enumerated in Section 209(b) when making a waiver determination. Judicial review of the EPA's waiver determination is also deferential, and the EPA is presumed to have acted properly when ruling on a waiver request, particularly if it approves the waiver.[16]

Section 209(b)(1) provides that before California can receive a waiver, it must make a threshold determination that its proposed standards "will be in the aggregate, at least as protective of public health and welfare as applicable Federal standards."[17] Once California has made such a determination, and seeks a waiver, Section 209(b) provides that the EPA *must* deny a waiver if the EPA finds that:

> (A) the determination of the State is arbitrary and capricious,
>
> (B) such State does not need such State standards to meet compelling and extraordinary conditions, or
>
> (C) such State standards and accompanying enforcement procedures are not consistent with section 202(a) of this part.[18]

An EPA finding that any one of these three criteria is met is grounds for denying California's waiver request.

The first criterion is unlikely to present much difficulty for California's waiver request. CARB has substantial experience developing and modeling vehicle emission control programs, even if not for greenhouse gases. California maintains an extensive vehicle emission control program and CARB's expertise in this area rivals that of the U.S. EPA.

When adopting its greenhouse gas regulations, CARB analyzed the proposed greenhouse gas emission reductions and concluded that they are no less protective than applicable federal standards. The EPA is obligated to give substantial deference to this determination. It can only reject CARB's conclusions on this point if it concludes that this determination was arbitrary and capricious, and this is unlikely. It is not enough if the EPA disagrees with CARB's conclusion. To reject the waiver request on this basis, the EPA must actually conclude that CARB's conclusion was arbitrary and capricious.

In support of its waiver application, CARB noted that, "since U.S. EPA has declined to set federal standards for greenhouse gases, California's Greenhouse Gas Regulations are unquestionably at least as protective as the applicable federal standards since the latter do not exist."[19] This is true. Some standard is necessarily as, if not more, protective than no standard. After *Massachusetts v. EPA*, the EPA is reconsidering whether to adopt some emission controls, but unless the EPA adopts regulations more stringent than California's—effectively making the Golden State's rules irrelevant—the California standards must be at least as protective as the federal standards.

It is conceivable that the EPA could conclude that CARB underestimated or unreasonably discounted the effect of the greenhouse gas emission standards on fleet turnover, and therefore underestimated the extent to which such emission controls could retard reductions in other air pollutants. It is also possible that further analyses could reveal that the California standards would actually impair other air pollution control efforts if they result in increased driving (due to increased fuel economy) resulting in greater emissions of other pollutants.[20]

The third criterion is also unlikely to provide a sufficient legal basis for rejecting the California rules. The EPA is unlikely to conclude that the California regulations governing greenhouse gas emissions are inconsistent with CAA Section 202(a).[21] CARB appears to have given adequate consideration to the technological feasibility of, and required lead time for, its greenhouse gas emission reduction standards. CARB adopted the rules several years before they took effect. Further, CARB maintains that its emission standards may be met with "off-

the-shelf" technologies. This is also a matter with which the Golden State's regulators have significant experience. Unless opponents of California's standards can demonstrate with clear and compelling evidence that CARB's conclusions are inaccurate, the EPA would be unlikely to deny a waiver on these grounds.

The Second Criterion

If the EPA were to deny California's waiver request—or if a federal court were to overturn an EPA decision to grant a waiver—it is most likely to be because California's regulation of greenhouse gas emissions from motor vehicles is not necessary to meet "compelling and extraordinary conditions." As noted above, Section 209(b) requires EPA to reject California's waiver request if the Agency determines that California "does not need such State standards to meet compelling and extraordinary conditions."[22] Global climate change, unlike the local and regionalized forms of air pollution that have been the focus of California's regulatory efforts to date, may not satisfy this standard.[23]

In the past, California has been able to argue that more stringent controls on vehicular emissions regulated by the EPA were necessary due to California's uniquely severe urban air pollution problems, the difficulty some California metropolitan areas would otherwise have meeting applicable National Ambient Air Quality Standards, and the comparatively large contribution mobile source emissions made to California's air pollution problems. None of these arguments are applicable in the context of global climate change.

CARB submits that the EPA must show as much, if not more, deference to California's policy determination that greenhouse gas emission reductions are necessary, as it would to other emission control policies.[24] There is little basis for this argument. If anything, the EPA is less likely to defer to California's determination because climate change is not an environmental problem that presents a "compelling or extraordinary" threat to California—as distinct from the nation as a whole.

The argument that climate change cannot satisfy the second criterion of Section 209(b) is not based upon any skepticism or denial of human contributions to climate change. Nor is it dependent upon rejecting that a modest increase in global temperature brought about by anthropogenic emissions of greenhouse gases could have negative ecological and other effects in California. As a coastal state, California may be threatened by sea-level rise in a way that land-locked states cannot be. California's unique geography and ecological conditions further mean that temperature increases will trigger different types of secondary effects there than elsewhere. But this may not be enough to satisfy Section 209(b)(1)(B).

Section 209(b) almost certainly requires that California do more than show that anthropogenic emissions are causing an increase in atmospheric concentrations of greenhouse gases that, in turn, contribute to a gradual warming of the climate, and that such warming could have negative effects. To read the second criterion in this way is to make it wholly redundant with the CAA standard for setting federal emission standards in the first instance, and thus a dead letter.

CAA Section 202(a)(1) requires the EPA to adopt controls on emissions from new motor vehicles that, in the judgment of the EPA Administrator "cause, or contribute to, air pollution which may reasonably be anticipated to endanger public health or welfare."[25] This is the standard for adopting federal controls on vehicular emissions of a given pollutant. Section 209(b) provides for a waiver of preemption for California regulations controlling pollutants that are already subject to regulation under this standard pursuant to Section 202(a). Therefore, to justify a waiver under Section 209(b), California must demonstrate something more than the fact that the accumulation of greenhouse gases in the atmosphere will contribute to global warming that, in turn, may have some deleterious effects in California. If Section 209(b) authorized the EPA to grant a waiver for any pollutant that could have negative effects it would, by definition, apply to every pollutant for which there is a standard under Section 202(a). Assuming the language of Section 209(b) serves some purpose within the Act, it must create a different and more demanding standard than that which triggers federal regulation under Section 202(a).

The most sensible reading of Section 209(b) is that California must be able to show that *California needs* more stringent standards than those provided by the EPA to meet specific conditions or concerns *in California*. This gives the Golden State a problem. Because global climate change is, by definition, a global phenomenon, California cannot claim that it faces a unique problem as distinct from that faced by the nation as a whole. Nor can California claim that it *needs* its greenhouse gas emission controls to address the threat posed by climate change, as adoption of these measures will not have any meaningful (if even measurable) effect on global climate change, let alone the specific effects of climate change about which Californians are concerned.

California cannot maintain that its regulatory program, even if adopted by a dozen other states, will have any meaningful effect on future projections of climate change. Dr. T.M.L. Wigley of the National Center for Atmospheric Research demonstrated that were *all developed nations*—those on "Annex B" of the Kyoto Protocol—to fully comply with the greenhouse gas emission reduction targets established by the Kyoto Protocol, and maintain such controls through 2100, this would only change the predicted future warming by 0.15ºC by 2100.[26] The reductions modeled in the Wigley study are several times greater than the complete elimination of all greenhouse gas emissions from the U.S. transportation sector, let alone any realistic estimate of emission reductions to be achieved from the imposition of regulatory controls on new motor vehicles nationwide (let alone in a handful of states). Thus, California cannot plausibly maintain that its vehicle emission controls would do much of anything to address any threat posed by climate change to the state.

CARB argues that "California need not demonstrate... that the state faces unique threats from greenhouse gas emissions," and that it is enough that the state faces "extraordinary and compelling conditions generally."[27] The basis for this argument is that the EPA has traditionally evaluated California waiver requests as applied to emissions control *programs* as opposed to individual *standards*. According to CARB, "the relevant inquiry

under Section 209(b)(1)(B) is whether California needs its own emission control program to meet compelling and extraordinary conditions, not whether any given standard is necessary to meet such conditions."[28] Because it is clear that California does experience the sort of "compelling and extraordinary conditions" that justify a California-specific emissions control program, CARB reasons, the greenhouse gas emission controls must be permitted as well.

There is some merit to CARB's argument. As interpreted by the EPA, Section 209(b) does not authorize or require the EPA to analyze separately each individual component of each program for which CARB seeks a standard. Rather, the EPA may look at programs as a whole, recognizing that the waiver provision is designed to enable California regulators to tailor a set of standards to California-specific pollution concerns and make different trade-offs than those embodied in relevant federal standards. But this does not mean—indeed cannot mean—that once California had adopted its first vehicular emissions control program, it would be able to adopt any and all emission control standards from that point forward that satisfied the remaining 209(b) criteria. Here California seeks to adopt a new set of standards to address a previously unregulated environmental concern. Surely California's preexisting emission control program, for which preemption was waived, does not require a waiver for the new standards as well; and the EPA would be wholly justified—if not required—to ensure that California's greenhouse gas emission controls satisfy Section 209(b)(1)(B), as have those measures adopted before it.

Because California cannot demonstrate that controls on vehicular emissions of greenhouse gases are *necessary* to meet any "compelling and extraordinary conditions" in California, the EPA would have ample justification for denying California's waiver request. This does not mean that the EPA is necessarily obligated to deny the waiver, however. The CAA gives the EPA some amount of discretion, and federal courts give substantial deference to such agency determinations if they are supported by a reasonable explanation. Insofar as the language of Section 209(b) is at all ambiguous, courts would defer to the EPA's construction of the standard. Therefore, an EPA decision to grant a waiver would necessarily be vulnerable to court challenge.

Potential EPCA Preemption

The CAA is not the only federal law that may preempt California's greenhouse gas emission standards for new motor vehicles. Therefore, even were the EPA to grant California's request for a waiver, this would not necessarily preclude the federal preemption of California's greenhouse gas emission controls under other laws, such as the Energy Policy and Conservation Act (EPCA). By its own terms, Section 209(b) only provides for a waiver of preemption under the Clean Air Act, not by other statutes.[29]

The EPCA establishes federal standards for automotive fuel efficiency, known as Corporate Average Fuel Economy (CAFE) standards.[30] The EPCA also explicitly preempts any state or local regulations "related to fuel economy standards."[31] This language adopts a fairly broad standard for preemption. It preempts all state rules "related to" fuel economy, and not simply those that explicitly seek to improve automotive fuel efficiency.

The National Highway Transportation and Safety Administration (NHTSA), the agency charged with implementing the fuel economy standard provisions of the EPCA, has argued with some force that federal fuel economy standards preempt state regulation of greenhouse gas emissions. This is because regulations limiting the vehicular emission carbon dioxide, the primary greenhouse gas emitted from automobiles, are, in effect, regulations of automotive fuel economy. As NHTSA explained:

> Congress has explicitly preempted any state laws or regulations relating to fuel economy standards. A State requirement limiting CO_2 emissions is such a law or regulation because it has the direct effect of regulating fuel consumption. CO_2 emissions are directly linked to fuel consumption because CO_2 is the ultimate end product of burning gasoline.[32]

Although the California regulations apply to four greenhouse gases, and not just carbon dioxide, compliance is measured by converting emissions of the other three greenhouse gases into "CO_2-equivalent." On this basis, NHTSA concluded that California's regulations are "based primarily" on carbon dioxide emissions.[33] Unlike the CAA, the EPCA does not contain a provision authorizing a waiver of preemption for California or any other state.

In *Massachusetts v. EPA* the Supreme Court held that the existence of federal fuel economy program under EPCA did not preclude the conclusion that the EPA had authority to regulate greenhouse gas emissions under the CAA. The Court concluded that these two obligations did not inherently conflict. It does not necessarily follow that *state* efforts to control greenhouse gas emissions are not preempted by federal fuel economy rules, however. In rejecting the EPA's argument that NHTSA's authority to set automobile fuel economy standards under EPCA precluded EPA authority to regulate greenhouse gas emission standards, the Court explained that "[t]he two obligations may overlap, but there is no reason to think *the two agencies* cannot both administer their obligations and yet avoid inconsistency."[34] Two federal agencies can be expected to communicate and coordinate overlapping regulatory efforts. The same level of cooperation cannot be assumed between federal and state agencies. Where two federal regulatory programs "overlap," such coordination may prevent one agency's regulation from preempting another's. When federal and state law "overlap," however, it is common to find that the federal rule preempts potentially conflicting state regulations. Therefore, *Massachusetts* should not be read to preclude EPCA preemption of California's greenhouse gas regulations.

The Broader Policy Context

To complement the legal analysis above, it is worth briefly considering the relevant policy considerations raised by California's waiver request. The CAA may impose some limits on the EPA's ability to grant California's waiver request, but this does not mean the balance struck by the CAA is the correct one. There is a strong argument that states should be given greater flexibility in meeting environmental standards than they receive under current law. [35]

The principle of "subsidiarity"—the principle that problems should be addressed at the lowest level practicable—is particularly appropriate in the context of environmental policy. Because most environmental problems are local or regional in nature, there is a strong case that state and local governments should be given the flexibility to design and implement their preferred approaches to such environmental concerns. This case is particularly powerful in the context of local or regional environmental concerns, such as land-use patterns, local drinking water quality, and the siting of polluting facilities. Decentralizing environmental policy in this way could spur ecological innovation and better "fit" between environmental protections and given environmental concerns. The waiver of applicable federal standards, particularly those that could preempt variable state programs, is one way of providing for such flexibility.[36]

A preference for decentralization or subsidiarity does not mean there should be no federal environmental regulation, however. It simply creates a rebuttable presumption toward decentralization—a presumption that can be overcome with a demonstration that more centralized action is necessary or likely to produce a more optimal result. For example, the presumption may be overcome where there is an identifiable *federal* interest, or some reason to believe that state and local governments will be systematically incapable or unwilling to adopt publicly desired environmental measures.[37]

The regulation of greenhouse gas emissions from motor vehicles touches upon two identifiable federal interests that could justify uniform federal regulation. First, global climate change is, by definition, a global concern. It affects the nation, indeed the world, as a whole. Emissions of greenhouse gases, such as carbon dioxide, disperse throughout the atmosphere. Because global climate change requires measures that address global atmospheric concentrations of greenhouse gases, there is little reason to believe that state governments are capable of adopting effective or efficient polices in this area.[38] Because climate change is a global concern, the most effective policies must be developed at a "higher" level—through international institutions or the cooperation of national governments, rather than independent actions by states.

State governments are simply incapable of adopting policies that will have a significant impact on climate change trends. This does not mean that state governments should always be precluded from adopting localized climate measures. State climate policies may encourage greater policy innovation, as it has in other contexts.[39] But given the relatively minimal benefits that such policies are capable of producing, there should be adequate attention paid to the potential of such policies to externalize costs on to other jurisdictions. If states wish to adopt various measures demonstrating their commitment to reducing greenhouse gases, they should be allowed to do so. States should not necessarily have the freedom to adopt policies that impose disproportionate costs on other jurisdictions.

Another national interest that could justify federal preemption of state standards is the national market for motor vehicles. There are economies of scale in the manufacture of products produced for and distributed to national markets that may make a single federal standard more efficient, and more beneficial for consumers and producers alike than a multiplicity of state standards.[40] Specifically, a single set of regulations may make more sense for a single, integrated national economy. This argument is strongest in the case of product regulation. Where a given product is bought and sold in national markets, and will travel throughout interstate commerce, it is less costly to design and produce so as to conform to a single national standard. While it is not clear why siting standards for a pulp mill in Vermont should match those for one in Oregon or Mississippi, if commercial goods are going to be produced on a national scale for national markets, producers may be best served if there is a single product standard that applies nationwide. In addition, consumers may benefit from national product standards, insofar as lower compliance costs result in lower consumer prices.

Allowing states to adopt more stringent product standards of their own poses the risk of one state externalizing the costs of its environmental preferences onto out-of-state market participants. For instance, if California and several northeastern states adopt more stringent emission standards for automobiles, and this produces a de facto national standard that increases production costs, consumers in other states may end up bearing a portion of the costs of more polluted states' preference for cleaner vehicles.

It is likely that the inherent economies of scale from the adoption of a single national standard for products sold in interstate commerce, such as automobiles, are less today than when Congress enacted Section 209. The costs of meeting variable state standards have declined with the development of customized manufacturing processes and just-in-time inventory.[41] Insofar as manufacturers are capable of tailoring production for different markets, state-specific product standards may not necessarily allow one state to externalize the costs of its environmental preferences on another. This does not mean that such concerns are wholly unwarranted, however. Nor does it alter the fundamental policy choice made by Congress that is reflected in CAA Section 209(b).

A Note on Timing & Policy Alternatives

California politicians have complained that the EPA has been particularly sluggish in responding to its waiver request. "If we don't see quick action from the federal government, we will sue the EPA," California Governor Arnold Schwarzenegger proclaimed in a recent speech.[42] In an April letter to the EPA, he warned that California would sue if a waiver were not granted within 180 days. Writing in the *Washington Post* with Connecticut Governor Jodi Rell, Schwarzenegger argued that the EPA's refusal to grant a waiver to date "borders on malfeasance."[43]

California first asked the EPA for a waiver in December 2005, but the federal agency stalled. With litigation concerning the CAA's applicability to greenhouse gases pending, the EPA was in no hurry to approve California's request. Indeed, prior to the Supreme Court's decision in *Massachusetts v. EPA*, the Agency had ample reason to defer action on California's waiver request. First, the EPA's legal position in that litigation suggested that the waiver provision was inapplicable. Equally important, it would have been perfectly reasonable for the EPA to decide to defer any action until the resolution of the litigation and a

judicial clarification of the scope of the EPA's authority and the applicability of the waiver provision.

After *Massachusetts v. EPA* was decided, the EPA finally began to review California's request and opened public comment on California's waiver request. If the past is any guide, it could take several months or more for the EPA to review the applicable comments, reach a final determination, and publish a final rule along with a reasoned explanation of its decision. EPA Administrator Stephen Johnson said that he expects an agency decision by the end of the year, but the EPA is hardly known for meeting deadlines.[44] Even if the EPA granted California's waiver forthwith, the decision could be litigated for years. In any event, a several-month delay is hardly unusual. It is the standard, deliberate (if regrettably slow) pace of federal administrative action—a pace that has not seemed to trouble California before. In the past, California has been more than willing to begin implementing new vehicle emission controls before obtaining an EPA waiver of preemption.

California Governor Arnold Schwarzenegger states that the EPA's failure to grant a waiver immediately will "result in California losing its right as a state to develop forward-thinking environmental policies."[45] This is not true. The lack of a waiver simply takes one policy option off of the table. Scores of alternatives remain available, and many states are exploring climate policy options that do not conflict with federal law. Nothing stops California from imposing emission controls on industrial and commercial facilities, and other sources of greenhouse gases, other than the political will to impose regulations on producers and consumers within the state. If reducing greenhouse gas emissions from automobiles and encouraging auto efficiency were really so important, California could increase its tax on gasoline. This would be a much more effective way to curb vehicular carbon emissions, particularly as it would reduce consumption by *all* car owners, not just those who purchase new cars. It also could be implemented far more quickly than vehicle emission standards that will not take effect for years. The problem, of course, is that the cost of such a policy would be readily identifiable, and clearly felt, by Californians. Thus it has less appeal to politicians than a vehicle emissions regulation, the costs of which will be buried in new car prices and dispersed to consumers and workers in other states.

CONCLUSION

The debate over California's request for a waiver of CAA preemption highlights the rigidity of federal environmental law. While many federal environmental statutes allegedly embody a "cooperative federalism" approach to environmental policy, in practice most federal environmental programs impose top-down regulatory requirements. The division of responsibility between the federal and state governments cannot be justified by reference to any coherent theory of their proper roles.[46] Federal environmental statutes are just as likely to impose rigid and uniform regulatory requirements on matters of local environmental concern as they are on those of truly national import. At the same time, states are increasingly likely to try their hand addressing trans-boundary or interstate environmental concerns, such as climate change. This creates a "jurisdictional mismatch" in environmental policy that is inefficient and, at times, even environmentally harmful.

California may well deserve a waiver of preemption, so that it may continue to experiment with potential greenhouse gas emission control policies. Yet, any serious policy argument for granting California a waiver in this instance would also justify authorizing waivers for other states to experiment in other areas in environmental law. The policy arguments for increased flexibility in the development of drinking water protection programs or local waste-site cleanup are far stronger than those for allowing an individual state to regulate products manufactured for national markets in order to address a globally dispersed pollution concern. The problem, however, is that CAA Section 209 authorizes potential waivers for California's air pollution control strategies, but waiver provisions in other environmental laws are few and far between. If policymakers wish to see California, and other states, experiment here, they should be willing to authorize broader experiments throughout much of the rest of environmental law.

Testifying before the Senate Environment Committee in May 2007, California Attorney General Edmund Brown, himself a former governor, suggested that if the EPA refuses to grant a waiver, the Congress should act: "Congress has to allow California to blaze its own trail with a minimum of federal oversight."[47] As he suggested, perhaps inadvertently, if Californians really want freedom from federal preemption, and a change in federal climate policy, they are better off getting Congress to act than seeking relief from the EPA. Congress may well have been too solicitous of the auto industry's desire for regulatory uniformity when it enacted Section 209, precluding California from taking some steps to address global warming on its own. If so, it is up to Congress, rather than the EPA, to fix it. Should Congress take this course, however, it should not stop with California's authority to regulate motor vehicle emissions. Instead, it should create broader opportunities for state experimentation and innovation across the board.

Endnotes

1 Statement of Edmund G. Brown, Jr., before the Senate Environment and Public Works Committee, Subcommittee on Clean Air and Nuclear Safety, May 22, 2007. Environmental law experts were more circumspect. Speaking to the Associated Press, Sean Hecht of the UCLA environmental law center commented that after *Massachusetts v. EPA*, "It's clear EPA has to consider California's waiver request now," but "that doesn't mean it's a foregone conclusions with respect to the waiver request." Quoted in Samantha Young, *EPA Revives California Emissions Rule*, AP Online Regional, Apr. 4, 2007, *abbreviated version available at* http://www.topix.com/sacramento/2007/ 04/ epa-revives-california-emissions-rule.

2 42 U.S.C. § 7543(a). This provision provides:

No State or any political subdivision thereof shall adopt or attempt to enforce any standard relating to the control of emissions from new motor vehicles or new motor vehicle engines subject to this part. No State shall require certification, inspection, or any other approval relating to the control of emissions from any new motor vehicle or new motor vehicle engine as condition precedent to the initial retail sale, titling (if any), or registration of such motor vehicle, motor vehicle engine, or equipment.

3 *See* Motor & Equipment Mfrs. Ass'n v. EPA (*MEMA I*), 627 F.2d 1095, 1109 (D.C.Cir. 1979), Congress' entry into the field and heightened state activity after 1965 raised the spectre of an anarchic patchwork of federal and state regulatory programs, a prospect which threatened to create nightmares for the manufacturers. Acting on this concern, Congress in 1967 expressed its

intent to occupy the regulatory role over emissions control to the exclusion of all the states all, that is, except California.

4 *See* E. Donald Elliott, Bruce A. Ackerman & John C. Millian, *Toward a Theory of Statutory Evolution: The Federalization of Environmental Law*, 1 J.L. Econ. & Org. 313, 330 (1985) (discussing automakers' role in the adoption of federal vehicle emission standards).

5 42 U.S.C. § 7543(b).

6 Section 209(b)(1) provides, among other things that:

The Administrator shall, after notice and opportunity for public hearing, waive application of this section to any State which has adopted standards... for the control of emissions from new motor vehicles or new motor vehicle engines prior to March 30, 1966, if the State determines that the State standards will be, in the aggregate, at least as protective of public health and welfare as applicable Federal standards.

As California was the only state to have adopted vehicle emission standards prior to March 30, 1966, this provision operates as a special provision for California only.

7 42 U.S.C. § 7507.

8 *See* 2002 Cal. Legis. Serv. Ch. 200, §3(a).

9 California Environmental Protection Agency, Air Resources Board, Request for a Clean Air Act Section 209(b) Waiver of Preemption for California's Adopted and Amended New Motor Vehicle Regulations and Incorporated Test Procedures to Control Greenhouse Gas Emissions: Support Document, Dec. 21, 2005 ("CARB Support Document"), at 4; *see also* Statement of Edmund G. Brown, Jr. (discussing why threat of global warming is "of particular concern to California.").

10 *See* Ann E. Carlson, *Federalism, Preemption, and Greenhouse Gas Emissions*, 37 U.C. Davis L. Rev. 281, 293 (2003). As Prof. Carlson notes, however, the EPA "has sometimes denied part of a waiver or delayed implementation of California emission standards." *Id.*

11 127 S.Ct. 1438 (2007).

12 *See* Steven Mufson & Michael A. Fletcher, *Bush Calls for Cuts in Vehicle Emissions*, Wash. Post, May 15, 2001, at D1 (reporting President Bush called upon federal agencies to draft regulations controlling greenhouse gas emissions from motor vehicles before the end of his term); *see also* Jonathan H. Adler, *Warming Up to Climate Change Litigation*, 93 Va. L. Rev. In Brief 61, 70-71 (2007) (discussing the legal implications of *Massachusetts v. EPA)*, *available at* http://www.virginialawreview.org/inbrief/2007/05/21/adler.pdf.

13 *See* California State Motor Vehicle Pollution Control Standards; Request for Waiver of Federal Preemption; Opportunity for Public Hearing, 72 Fed. Reg. 21260 (Apr. 30, 2007).

14 42 U.S.C. § 7543(b).

15 *MEMA I*, 627 F.2d at 1120-21.

16 *Id.* at 1121 (noting "the burden of proof lies with the parties favoring denial of the waiver").

17 42 U.S.C. § 7543(b)(1).

18 42 U.S.C. § 7543(b)(1)(A)-(C). *See also MEMA I*, 627 F.2d, at 1111 ("if the Administrator makes any one of these findings with respect to a waiver request involving California 'standards' he must deny the request").

19 CARB Support Document at 14.

20 Increasing automobile fuel economy reduces the cost-per-mile of driving, which leads, in turn, to an increase in vehicle miles traveled. This "rebound effect" offsets an estimated 10 to 20 percent of the fuel savings from increased fuel economy. *See* Paul R. Portney, et al., *The Economics of Fuel Economy Standards*, 17 J. Econ. Persp. 203 (2003); *see also* David L. Greene, James R. Kahn, & Robert C. Gibson, *Fuel Economy Rebound Effect for U.S. Household Vehicles*, 20 Energy J. 1 (1999) (discussing rebound effect from fuel economy standards).

21 42 U.S.C. §7521(a).

22 42 U.S.C. § 7543(b)(1)(B).

23 *See* Carlson, *supra* note 10, at 296-97 (noting how greenhouse gas emissions waiver may present different issues than prior waivers).

24 *See* CARB Support Document at 11.

25 42 U.S.C. §7521(a). The provision provides that:

> The Administrator shall by regulation prescribe (and from time to time revise) in accordance with the provisions of this section, standards applicable to the emission of any air pollutant from any class or classes of new motor vehicles or new motor vehicle engines, which in his judgment cause, or contribute to, air pollution which may reasonably be anticipated to endanger public health or welfare.

26 T.M. L. Wigley, *The Kyoto Protocol: CO2, CH4 and Climate Implications*, 25 Geophysical Research Letters 2285 (1998).

27 CARB Support Document at 16.

28 CARB Support Document at 15.

29 Section 209(b) only provides for the waiver of "application of this section." 42 U.S.C. § 7543(b); *see also* Central Valley Chrysler-Jeep v. Witherspoon, 456 F.Supp.2d 1160, 1173 (E.D. Cal. 2006) ("Section 209(b) does not provide that the regulations, once EPA grants a waiver, become federal law and are thereby rendered immune from preemption by other federal statutes.").

30 49 U.S.C. § 32902.

31 49 U.S.C. § 32919(a). This provision provides:

> When an average fuel economy standard prescribed under this chapter is in effect, a State or a political subdivision of a State may not adopt or enforce a law or regulation related to fuel economy standards or average fuel economy standards for automobiles covered by an average fuel economy standard under this chapter.

32 Average Fuel Economy Standards for Light Trucks Model Years 2008-2011; Final Rule; 71 Fed. Reg. 17654 (Apr. 6, 2006). NHTSA further explained:

> Given that the amount of CO_2, CO, and hydrocarbons emitted by a vehicle varies directly with the amount of fuel it consumes, EPA can reliably and accurately convert the amount of those gases emitted by that vehicle into the miles per gallon achieved by that vehicle.

Id. at 17656.

33 *Id.* at 17656.

34 127 S.Ct. at 1462.

35 *See generally*, Jonathan H. Adler, *Letting Fifty Flowers Bloom: Using Federalism to Spur Environmental Innovation*, *in* The Jurisdynamics of Environmental Protection: Change and the Pragmatic Voice in Environmental Law 263-64 (Jim Chen ed., 2004).

36 A proposal for ecological waivers is outlined in Adler, *Letting Fifty Flowers Bloom*. A similar proposal is put forth in Daniel A. Farber, Eco-Pragmatism: Making Sensible Environmental Decisions in an Uncertain World 194-98 (1999).

37 This argument is made in greater detail in Jonathan H. Adler, *Jurisdictional Mismatch in Environmental Federalism*, 14 N.Y.U. Envtl. L.J. 130 (2005).

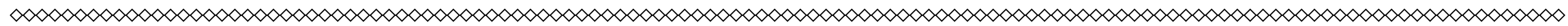
38 *See* Jonathan H. Wiener, *Think Globally, Act Globally: The Limits of Local Climate Policies*, 155 U. Pa. L. Rev. 101, 102 (2007) ("subnational state-level action is not the best way to combat global climate change"; "local action is not well suited to regulating mobile global conduct yielding a global externality"); *see also id.* at 103 ("state-level efforts could not only be ineffectual, but counterproductive, increasing net emissions and undercutting a wider effort to constrain global emissions.").

39 *See* Carlson, *supra* note 10, at 313-315 (suggesting that California's ability to obtain waivers of preemption may encourage greater policy experimentation).

40 *See* Pietro S. Nivola & Jon A. Shields, Managing Green Mandates: Local Rigors of U.S. Environmental Regulation 17 (2001) ("Business interests, not without justification, often prefer nationwide regulatory standards to a hodgepodge of local rules: broad scope and standardization may lower uncertainty and increase efficiency.").

41 Adler, *supra* note 37, at 149-50.

42 *See* Bernie Woodall, *Schwarzenegger Warns of Calif. Suit Against EPA*, Reuters, Apr. 25, 2007, *available at* http://www.reuters.com/article/bondsNews/idUSN2531693320070425.



43 Arnold Schwarzenegger & Jodi Rell, *Lead or Step Aside, EPA*, Wash. Post, May 21, 2007, at A13.

44 *See* Avery Palmer, *Panel Democrats Seek to Prod EPA on California Auto Emissions Rule*, CQ Today, July 31, 2007 (reporting "At a July 26 committee hearing, EPA Administrator Stephen L. Johnson testified that the waiver decision is complex, but that the agency has committed to issuing a decision before the end of the year.").

45 *See* Letter from Governor Arnold Schwarzenegger to President George W. Bush, Oct. 24, 2006, *available at* http://gov.ca.gov/index.php?/press-release/501/.

46 *See* Adler, *supra* note 37, at 157-177.

47 Statement of Edmund G. Brown, Jr., before the Senate Environment and Public Works Committee, Subcommittee on Clean Air and Nuclear Safety, May 22, 2007.

Do Inclusionary Zoning Laws Violate *Nollan, Dolan,* and the Doctrine of Unconstitutional Conditions?

*By James S. Burling**

In most parts of the country, there is a shortage of low-income housing. Local governments realize that more must be made available. And, for some reason, it must be new.[1] The problem with alleviating a shortage of new low-income housing is that someone has to build it. But since government does not have a legacy of success in building public housing, taxpayers are reluctant to give more money for the construction of government housing projects. Hence, government planners have turned to inclusionary zoning.

In a nutshell, inclusionary zoning says that if a developer builds somewhere between five and ten homes for sale at market rates the builder has to build one that is "affordable" to the "workforce."[2] Or pay the local housing commissariat for a subsidized housing fund.

Who picks up the tab for the subsidized housing? Proponents suggest the builders or landowners will eat the costs. Such burden shifting, they claim, is only fair because permission to build requires a quid pro quo. Alternatively, advocates argue that the construction of market rate housing creates a need for more "affordable" housing units—either because land used for market rate housing will no longer be available for "workforce" housing, or because the families who dwell in market rate housing create a need for low-income service sector jobs such as gardeners, beauticians, and the people who clean government buildings.

The idea seems unlikely to fare well economically. Whether courts will find it unconstitutional as well is an open question. The Supreme Court has on several occasions, first in *Nollan v. California Coastal Commission*, held that the Takings Clause prohibits government from demanding exactions from builders unless the exaction served to relieve some actual harm caused by the development.[3] Moreover, in *Dolan v. City of Tigard*, the Court said the government proposing the exaction had the duty to demonstrate on an individualized basis that the exaction is "roughly proportional" to burdens created by the development.[4] Property owners and planners, however, do not agree on whether *Nollan* and *Dolan* applies to inclusionary zoning. There are disagreements as to whether these cases apply to legislatively mandated formulae rather than individualized exactions, and arguments over the degree of scrutiny courts should apply in determining whether proffered justifications for inclusionary zoning meet constitutional standards. And in the post-*Lingle* world there are disputes about the confluence of takings doctrine, substantive due process, and the doctrine of unconstitutional conditions.[5] Finally, there are uncertainties over the existence in some inclusionary zoning jurisdiction of "sweeteners" that take the form of density bonuses and the like. Are these valid forms of quid pro quo that avoid the constitutional infirmities, or are they illusory in that they rob out of one pocket in order to compensate the other?

Despite the growth of inclusionary zoning laws, there has not been an accompanying rush by developers to bring constitutional challenges. Developers are, by and large, pragmatic business entrepreneurs. When confronted with what they perceive to be an obnoxious inclusionary zoning ordinance, they often see two stark choices. They may fight the exaction as being contrary to their constitutional rights. But such an endeavor may seem quixotic and insane. To bring such a challenge, a developer, in many jurisdictions, might have to forego the developing project while the dispute is being played out in the courts. The developer may suspect that courts likely to hear the dispute are not particularly hospitable to property rights claims. Moreover, most courts operate in a universe in which time means nothing; to developers, time is money. The developer may also harbor a well-grounded fear that there may be a direct relationship between the seriousness of the challenge and the time it will take to get other project permits approved for the foreseeable future. For these reasons, the second choice seen by developers may seem much more palatable: pay the freight, pass on the costs to the families who end up buying the homes (either through higher prices or reduced quality), and simply shut up. Developers almost universally choose the second alternative.

I. Exactions—From *Nollan* to *Lingle*

A. The Origins in Nollan and Dolan

The ability to make reasonable use of private property is a fundamentally important right that may not be burdened by unreasonable regulatory constraints. Two Supreme Court cases in particular illustrate this principle.

In *Nollan v. California Coastal Commission*, the Supreme Court was confronted with the California Coastal Commission's practice of demanding that landowners dedicate property for access to and along the coast in exchange for building permits.[6] In *Nollan*, the Commission demanded that the Nollans dedicate approximately one third of their property in order to be given a permit to replace a one-story home with a two-story home. The dedication would have given the public the right to cross back and forth along the dedicated property parallel to the ocean. The Commission justified the demand as a way of remedying the loss of ocean view of travelers along Highway 1, a loss the Commission characterized as a "psychological barrier."[7]

The Court was not persuaded and called this demand for an exaction an "out-an-out plan for extortion."[8] In doing so the Court established that a condition of this nature can be imposed only if certain conditions are present. First, the impact caused by the proposed development must be severe enough to justify a denial of the permit in the first place (a denial that would not itself rise to the level of a taking.) Second, a condition may instead be imposed so long as the exaction will actually serve to ameliorate the adverse impact which could have justified the denial of the permit in the first place.[9] Because forcing the Nollans to dedicate property parallel to the beach would do

---

* *James S. Burling is an attorney with the Pacific Legal Foundation. He presented a version of this article to the ALI-ABA Course on Regulatory Takings and Related Government Liability, April 12-14, 2007.*

nothing to ameliorate the lost view of drivers along Highway 1, the exaction was seen to lack the required nexus.

This standard was further refined in *Dolan v. City of Tigard*, where the owner of a hardware and plumbing store was told to dedicate and build a bicycle trail and give up land for public access in exchange for a permit to expand the business.[10] The justification for these demands was to ameliorate the impact of the expansion on traffic and flooding. The Supreme Court held that while there may be some nexus, Tigard had a duty to demonstrate that the exactions were "roughly proportional" to the impacts caused by the development.[11] The requirements of a nexus described in *Nollan* and the need for the government agency to prove rough proportionality in *Dolan* has been described as calling for "heightened scrutiny" of exactions.

B. Monetary and Legislative Exactions—
Conflicts in the State and Lower Federal Courts

After *Nollan* and *Dolan* it should be plain that government cannot condition the exercise of the right to put property to economically beneficial use upon the forced payment of land, money, or labor when those exactions are not closely related to impacts caused by the use of the underlying property. What remains in some debate is whether there is a distinction between exactions of land and money, and between exactions imposed as part of an adjudicatory permit process and those enacted pursuant to a legislative process. In *Ehrlich v. City of Culver City*, for example, the California Supreme Court held that *Nollan* and *Dolan* applied heightened scrutiny to monetary exactions, but not to legislatively adopted exactions.[12]

Legal scholars, however, have opined that legislative enactments restricting property rights deserve heightened scrutiny. For example, Professor Doug Kmiec suggests that heightened scrutiny is appropriate for legislative enactments imposing public burdens on specific private uses.[13] Other scholars have noted the conflicts that exist among the lower federal and state courts, with some courts treating exactions very differently according to whether they are legislatively adopted or not, and whether the exaction is for money or land.[14]

The holdings of several states and the First Circuit confirm that generally applicable legislative enactments are subject to the standards established in *Nollan* and *Dolan*.

New York: Before *Dolan* was decided, but after *Nollan*, the New York Court of Appeals struck down a permit fee imposed on owners of single room occupancy hotels who wished to change the use of their property. In *Seawall Associates v. City of New York*,[15] the fee was designed to subsidize the City's low-income housing program and had been imposed by legislation and, as such, was uniformly applied[16]. The court focused on the lack of a nexus between the complex problem of low-income housing in New York City and the landowners' use of their property for other purposes[17], stating that "[s]uch a tenuous connection between means and ends cannot justify singling out this group of property owners to bear the costs required by the law toward the cure of the homeless problem." The same court also struck down legislatively enacted rent control ordinances as being violative of the Takings Clause in *Manocherian v. Lenox Hill Hospital*.[18] That decision relied upon *Nollan* and *Dolan*.

Ohio: Home Builders Association of Dayton v. City of Beavercreek,[19] applied a "middle level of scrutiny" to legislatively adopted transportation impact fees imposed on subdivision approvals. The Court reasoned that a test that looks at both the need for the transportation project and the fairness of imposing the cost on developers is a test that will

> adequately balance the interests of local governments with those of property owners. The first prong of the test decides whether the ordinance is an appropriate method to address the city's stated interests, and the second prong assures that the city and developers are paying their proportionate share of the cost of new construction.[20]

The court upheld the fees in that case after noting that the trial court "reviewed volumes of evidence" and made the requisite factual findings regarding the necessity for the road projects and the impacts to be caused by subdivision development.[21]

Oregon: Citing to state supreme court precedent, the Oregon Court of Appeals has held in *J.C. Reeves Corp. v. Clackamas County* that " 'the character of the [condition] remains the type that is subject to the analysis in *Dolan*' whether it is legislatively required or a case-specific formulation.[22] The nature, not the source, of the imposition is what matters."[23]

Washington: In *Sparks v. Douglas County*,[24] the Washington Supreme Court applied *Dolan* to a legislatively adopted road dedication exaction. It upheld the exaction because there was evidence showing that the exaction was roughly proportional to the impacts of development. In *Trimen Development Co. v. King County*,[25] a developer had been given the choice of dedicating or reserving land for open space or paying a fee in lieu of such dedication. The court properly recognized the applicability of the *Dolan* test to these legislatively imposed fees.[26] The fees were upheld, incidentally, as the court found evidence to support the conclusion that the fees were "reasonably necessary as a direct result of Trimen's proposed development."[27]

Illinois: The Illinois Supreme Court in *Northern Illinois Home Builders Association, Inc. v. County of Du Page*,[28] applied *Dolan* to a legislatively enacted transportation district and impact fee enabling act. (The Court then proceeded to apply its even more exacting "specifically and uniquely attributable" rule in striking down part of the statute.) Similarly, an Illinois appellate court agreed in *Amoco Oil Co. v. Village of Schaumburg*[29] that

> [a]lthough not binding as precedent, we find Justice Thomas' comments [in *Parking Association of Georgia*] particularly persuasive and consonant with the rationale underlying *Dolan* and similar cases. Certainly, a municipality should not be able to insulate itself from a takings challenge merely by utilizing a different bureaucratic vehicle when expropriating its citizen's property.[30]

First Circuit: In *City of Portsmouth v. Schlesinger*,[31] the court reviewed a low-income housing fee that had been imposed when the owner of the site of some low-income apartments sought permission to replace the apartments with new condominiums.[32] While the challenge to the fee was saddled with a statute of limitations problem, the court noted that the facts were undisputed that "no rational nexus existed between the amount of the Developers'... [fee] and any burden imposed on the City due to the zoning change."[33] If not for a statute

of limitations problem, the court would have struck down the fee.[34]

New Jersey: The New Jersey Supreme Court has been particularly aggressive, not only in upholding inclusionary-zoning ordinances, but in actually mandating them. In *Holmdel Builders Association v. Township of Holmdel*, the New Jersey Supreme Court considered whether various municipal inclusionary zoning ordinances that required the payment of low-income housing fees were constitutional.[35] With respect to a challenge under the Takings Clause, the court merely observed that "[a]s long as the measures promulgated are not confiscatory and do not result in an inadequate return of investment, there would be no constitutional injury."[36] The court noted that it had earlier held in *Southern Burlington County NAACP v. Township of Mt. Laurel*,[37] that "municipalities [must] use affirmative inclusionary-zoning measures, including mandatory set-asides, to redress affordable-housing needs."[38] Interestingly, the court in *Mt. Laurel II* barely addressed a Takings Clause argument, dismissing with a single sentence that "mandatory set-asides keyed to the construction of lower income housing, are constitutional and within the zoning power of a municipality."[39] In *Holmdel Builders* the court ultimately concluded that "a residential developer could be required to set aside a percentage of units to be used for low- and moderate-income housing."[40] Since development fees perform an "identical function" they are permissible[41]. The difficulty with the court's analysis is that it completely ignores the "substantially advance" test of *Agins*, the nexus requirement of *Nollan*, and the rough proportionality standard of *Dolan*.[42]

North Dakota: In determining whether a drainage requirement could be applied to property of Burlington-Northern Railroad in light of *Dolan*, the Supreme Court of North Dakota held that it could because the duty "in this case arises not from a municipal 'adjudicative decision to condition,' but rather from an express and general legislated duty under a constitutional reservation of police power over a corporation." *Southeast Cass Water Resource District v. Burlington Northern Railroad Company*.[43]

Minnesota: In *Arcadia Development Corp. v. City of Bloomington*,[44] a Minnesota appellate court upheld the imposition of a tenant-relocation fee imposed on owners of mobile home parks in exchange for permission to go out of the mobile home park business. The appellate court rejected application of *Dolan* to tenant relocation fees because the fee was legislatively imposed. It found that "[o]nce legitimate governmental interests are identified, courts simply require a nexus between the local legislation and the legitimate governmental purposes."[45] Furthermore, "[b]ecause this case involves a challenge to a citywide, legislative land-use regulation, *Dolan*'s 'rough proportionality' test does not apply."[46]

Arizona: In Home Builders Association of Central Arizona v. City of Scottsdale,[47] the Arizona Supreme Court upheld the city's resource development fee that was imposed on land owners in order to build a new water supply infrastructure. The court found that *Dolan* did not affect the legality of the fee at issue and refused to apply heightened scrutiny because the fee was "legislatively" imposed: "Because the Scottsdale case involves a generally applicable *legislative* decision by the city, the court of appeals thought *Dolan* did not apply. We agree, *though the question has not been settled by the Supreme Court*."[48]

C. Exactions After Lingle

Up until *Lingle*, an oft-cited justification for the heightened scrutiny requirements found in *Nollan* and *Dolan* was a demand for an exaction that was insufficiently related to an adverse impact was a demand that "failed to substantially advance a legitimate governmental interest." This, of course, was the taking standard of *Agins*. Unlike most other takings, however, the usual remedy for a taking of this sort was not the payment of just compensation but invalidation of the exaction—as ultimately occurred in both *Nollan* and *Dolan*. With the rejection of the "substantially advance" standard as a *stand alone* takings standard in *Lingle,* some question could have been raised about the viability of these doctrines. However, the Court in *Lingle* also made it very clear that *Nollan* and *Dolan* retained their full vitality, repeating the formula of these cases, but noting instead that they fell under the rubric of "unconstitutional conditions" rather than "substantially advance" takings.

In *Lingle* the United States Supreme Court reaffirmed the vitality of the Doctrine of Unconstitutional Conditions. Citing to *Dolan v. City of Tigard*,[49] the unanimous Court wrote:

> As the Court explained in *Dolan*, these cases involve a special application of the "doctrine of 'unconstitutional conditions'" which provides that "the government may not require a person to give up a constitutional right—here the right to receive just compensation when property is taken for a public use—in exchange for a discretionary benefit conferred by the government where the benefit has little or no relationship to the property."[50]

An important point of similarity between the exactions in *Nollan* and *Dolan* and the condition at issue in this case, is that in these cases all involve more than mere governmental benefits. As the Court pointed out in *Nollan*:

> But the right to build on one's own property—even though its exercise can be subjected to legitimate permitting requirements—cannot remotely be described as a "governmental benefit." And thus the announcement that the application for (or granting of) the permit will entail the yielding of a property interest cannot be regarded as establishing the voluntary "exchange."[51]

II. The Meaning of the Doctrine of Unconstitutional Conditions, According to the Academy

The doctrine of unconstitutional conditions usually involves an exchange—the government gives a benefit to a person in exchange for something from the owner that the government would not ordinarily be entitled to. The lawfulness of the exchange is at the heart of the unconstitutional conditions doctrine. But to properly understand whether the exchange is lawful, it is important to understand the nature of the benefit and the condition being imposed.

However, many commentators do recognize the distinction between conditions placed upon the receipt of benefits and those placed upon the exercise of rights.[52] With respect to the doctrine of unconstitutional conditions, the analysis of "benefits" is pervasive. As Professor Kathleen Sullivan describes the doctrine, "unconstitutional conditions holds that

government may not grant a benefit on the condition that the beneficiary surrender a constitutional right, even if the government may withhold that benefit altogether."[53]

But there is a continuum of the degree to which a "benefit" is indeed a benefit, and when it is a right. Welfare benefits are an example of a benefit that the government is under no legal obligation to provide at all.[54] Of course, once the decision to provide such benefits is made, certain legal constraints must be employed so that benefits are not denied for unfair reasons and through unfair procedures. The right to develop property, on the other hand, involves the exercise of a right—albeit a right subject to regulation in order to prevent certain types of negative externalities. In *Nollan*, the Nollans had a right to put their property to an economically viable use. That right may be regulated in order to prevent external harms to neighbors and the public. In that case, the California Coastal Commission claimed that it could prevent the Nollans from adding a second story to their home in order to prevent a "psychological barrier" that would prevent people traveling on Highway 1 from realizing the presence of the ocean.[55] Assuming for the sake of argument the legitimacy of this justification,[56] the Court found that the condition sought by the Commission, the dedication of a lateral access path, had no relation to the alleged harm, and therefore could not be justified. Thus, while there were more than benefits at issue in *Nollan* the underlying leverage was on the exercise of a right, albeit a right subject to reasonable regulation.

Thus, when a court considers the theoretical justifications that have been put forth for the doctrine of unconstitutional conditions, it is important to bear in mind the nature of the "benefit" being sought from the government. The degree to which the "benefit" is a fundamental right, rather than an optional gift of the government, should influence the degree to which the government may condition the receipt of the benefit.

A. All Academic Theories, However Imperfect They May Be, Reject the Conditioning of Fundamental Rights

The doctrine of unconstitutional conditions is characterized by competing theoretical justifications, none of which is either universally accepted or in harmony with case law[57]. For example, Professor Kathleen Sullivan has described several such theories (e.g. coercion, corruption, and "commodification"), rejecting all of them in favor of her own systematic theory—even while noting that her theory is not consistent with recent Supreme Court precedent[58]. Likewise, Professor Richard Epstein's theory, based in part on law and economics, suggests that a robust doctrine of unconstitutional conditions is "second best" to a correct and limited view of the police power[59]. "But," as noted by Mitchell Berman, "these efforts, and those by other distinguished scholars, have left most observers unpersuaded."[60] However, Berman continues: "Very possibly, the Supreme Court has come closer to grasping the essential logic of coercion in its takings decisions than anywhere else."[61]

Theoretical justifications for the doctrine of unconstitutional conditions may involve debate on the margins and on questions of precisely when the doctrine should be invoked. This debate and uncertainty, however, is largely irrelevant in a case such as the present one. There is no exception to the requirement that government pays just compensation when property is taken.

To the extent that the ability to use private property is a fundamental right,[62] the requirement that a landowner seeking to develop property give up land or money in order to subsidize housing for unrelated third parties is a problematic proposition.

B. The Greater Power Includes the Lesser Power Theory

Sometimes, the doctrine of unconstitutional conditions is characterized as the right to exercise a "lesser power." That is, rather than denying a permit such as in *Nollan*, the government may instead grant the permit with conditions. Thus, Professor Sullivan describes the doctrine as:

> It reflects the triumph of the view that government may not do indirectly what it may not do directly over the view that the greater power to deny a benefit includes the lesser power to impose a condition on its receipt. Consensus that the better view won, however, has not put an end to confusion about its application.[63]

But as Sullivan notes, not all jurists and scholars have accepted this rather facile view of the doctrine.[64] Professor Epstein, for example, suggests that a law and economics approach reveals that this "greater power/lesser power" distinction in fact "gives political actors a greater opportunity to extract economic rents."[65]

More importantly, this justification for imposing conditions must disappear when there is no "greater power" in the first place. Here, the government has no right to force an ordinary citizen to subsidize the housing of unrelated third parties and has no right to arbitrarily deny a building permit that will not cause external harms—such as decreasing the supply of affordable housing. *A fortiori*, the lesser power—here the power of conditioning the receipt of a building permit on the condition that third parties receive housing subsidies—cannot be justified under the doctrine.

C. Coercion Theory

Another justification for the unconstitutional conditions doctrine is that it reduces certain forms of "coercion" by the government. In other words, government cannot "coerce" people into agreeing to conditions to which it is not otherwise entitled. The theory suggests that people would rather not enter into a particular bargain with the government, but they feel compelled to do so in order to gain a benefit. Some believe, for example, that this theory explains the "out-and-out plan of extortion" rhetoric of the Court in *Nollan*.[66] But Professor Sullivan is less than satisfied with this explanation:

> Neither the Court nor the commentary, however, has developed satisfying theory of what is coercive about unconstitutional conditions. Conclusory labels often take the place of analysis—for example, conditioned benefits are frequently deemed 'penalties' when struck down and 'nonsubsidies' when upheld.[67]

Sullivan explicates at length why the coercion justification for the doctrine is inconsistent with both case law and theory, concluding:

> While unconstitutional conditions doctrine thus is hardly unique in deeming some offers of benefit coercive, the concept of coercion will depend just as inescapably on independent conceptions of utility, autonomy, fairness, or desert in the unconstitutional

> conditions context as in other contexts. Coercion is a judgment, not a state of being.[68]

Professor Epstein articulates this problem in the coercion theory by saying, "the greatest difficulty with the coercion question is to identify the appropriate baseline against which the possibly coercive effects of government action must be evaluated."[69]

D. Maximization of Social Utility

Professor Epstein posits that a theory wherein only government conditions that maximize social utility can be justified under the doctrine[70]. Under this theory, a government bargain must be analyzed to determine whether it generates a net social utility. For example, when government condemns property for a highway and pays just compensation, there is net social utility: the government and the public gain a highway and its increased social and economic value, the landowner receives compensation *and* enjoys the social utility provided by the highway. However, in an unconstitutional conditions case, the government may try to put the receiver of a benefit in a worse position than before—which may result in a net social loss. In *Nollan*, for example, Epstein suggests that the Court's requirement of a nexus between the alleged public harm (the impact on the view) and the condition (the lateral easement), forces government to put some value on the two separate aspects of the bargain in order to prevent a net social loss[71]. In other words, if the value of the development is disproportionate to the value of the lateral easement, then the doctrine forces the government to make choices based on net social utility. It might not, therefore, require the lateral easement if that means the loss of the ability of development of beachfront lots. As will be shown later in this outline, there is serious doubt over whether inclusionary zoning mandates achieve any social utility—rather than providing more affordable housing they may be responsible for limiting the supply of such housing.

E. Government Overreaching and Other Theories

Other justifications for the doctrine of unconstitutional conditions abound. Professor Sullivan raises and, in turn, rejects theories of "germaneness,"[72] "corruption,"[73] and "commodification" (dealing with the supposed inalienability of certain rights)[74] before settling upon her own theory that concludes, "the doctrine guards against a characteristic form of government overreaching and thus serves a state-checking function."[75] And, more recently, Professor Berman raises and rejects all of these theories because none has garnered widespread academic support, and proceeds to posit his own theory that combines elements of each.[76] Notwithstanding his valiant attempt, this lack of settled understanding appears to be as true after Berman's attempt at a theory than it was before.

A hallmark of all of these academic treatments is that after positing a particular theory, they proceed to pick and choose amongst United States Supreme Court precedents, arguing that *this* case got it right, whereas *that* case got it wrong, or that *this* justice often gets it right and *that* justice rarely does, but most likely *these* justices are inconsistent. In other words, the utility of any of these grand unified theories is not particularly apparent in predicting real world outcomes.

Thus, the central difficulty with all of these recent attempts at a grand unified theory for the doctrine of unconstitutional conditions is that the Court has refused to play along. Academic uncertainty aside, it is clear that the Court in *Lingle* is untroubled by the assertion of the doctrine in the context of the Just Compensation and Takings Clauses. Moreover, while the government certainly had some discretion in *Dolan* to grant or deny the subject permits—depending on the degree and nature of the externalities caused by the development in those cases—that sort of discretion is lacking when it comes to questions of affordable housing and the development of new homes. That is because it is highly doubtful that a local government can ever prove that the externalities of new home construction include the exacerbation of a shortage of affordable housing.

III. Economic Analysis and Data Show that Regulatory Barriers are the Primary Cause of Shortages in Affordable Housing and That New Home Construction Reduces the Overall Price of Housing

To the extent that municipalities are required to justify the imposition of affordable housing mandates on the builders of market rate housing, then those municipalities ought to be able to demonstrate that the construction of market rate homes creates a need for more affordable housing. In many cases, local governments do little or nothing to justify the imposition of inclusionary zoning dictates. But where justifications are provided, they usually fall into two categories: (1) that the construction of market rate homes removes land from the inventory of land that might otherwise be used for affordable housing; and (2) the people who move into market rate homes often utilize the services of low-wage workers, such as gardeners, service-sector employees, and the like. If *Dolan* stands for the proposition that the government justifications for exactions are subject to heightened scrutiny, then in instances where local governments do provide one or both of these justifications, courts ought to consider what might be highly significant considerations.

Thus, where there is an alleged shortage of land inventory, a relevant consideration may be what caused the land shortage in the first place. It may be that the region is nearly all built-out. But it may also be that the government itself may be responsible for the shortage due to an aggressive campaign of *exclusionary* zoning, such as urban limit lines, large lot and setback requirements. Likewise, a court should consider whether the choice of developers to build market rate homes is in part due to government policies that force developers to build homes with higher prices. This could include the imposition of development fees and other exactions (paradoxically including workforce housing fees) that often add tens of thousands of dollars to the cost of housing. (In some parts of California, these fees exceed six figures).[77]

In situations where a municipality attempts to justify inclusionary zoning mandates with the explanation that purchasers of these homes employ lower wage workers, a court employing a heightened scrutiny analysis would be advised to consider whether the local government has examined whether these workers already exist in the community and, whether they are currently fully employed. By discouraging new families

from moving into an area and providing more employment opportunities, these lower-income workers would compete for fewer jobs and thus make it easier for employers to offer lower wages. In other words, rather than bringing more lower-income workers into a community, new growth might better employ those already in a community, making it easier for the employed to purchase or rent unsubsidized housing.

A court would also be well advised to consider where the families who move into the market rate housing otherwise live. In other words, it may well be that the construction of market rate housing does not itself create the families who move into these new homes; they might otherwise be competing with middle and lower income workers for the existing limited supply of housing, driving up costs even further.

The bottom line is that the law of supply and demand predicts that the greater supply of all housing, the lower will be the costs. Indeed, this is the near universal conclusion of economists who have studied the problem. For example, Edward L. Glaeser and Joseph Gyourko, in a comprehensive review of the zoning and housing affordability, lay the blame for higher home prices on zoning constraints.[78] Similar results were found in California in a study commissioned by the Reason Institute.[79]

The last major federal investigation into the causes of barriers to affordable housing concluded that

> government action is also a major contributing factor in denying housing opportunities, raising costs, and restricting supply. Exclusionary, discriminatory, and unnecessary regulations at all levels substantially restrict the ability of the private housing market to meet the demand for affordable housing, and also limit the efficacy of government housing assistance and subsidy programs.[80]

The Commission continued that large fees "have a regressive effect. Fees are generally fixed regardless of the cost of a new home. Thus, households that can only afford less expensive houses end up paying a higher proportion of the sales price to cover the cost of fees."[81] In examining the impacts of restrictive and exclusionary zoning in the suburbs the Commission found "[w]hen used in an exclusionary manner they have a notable impact on residential land costs, especially in preferred suburban locations.[82]

Another comprehensive collections of analyses was commissioned by the Pacific Institute for Public Policy Research.[83] In that study, Dr. Bernard Frieden, concluded:

> Freeing some land from development makes the remaining sites more expensive. Growth controls further increase the cost of land for each new home when they mandate large minimum-lot sizes.[84]

Reviewing the relationship between housing costs and growth controls in the Santa Barbara area, Lloyd J. Mercer and W. Douglas Morgan, authors of another report in the Pacific Institute study,[85] conclude: "To the extent that the 'housing crisis' on the South Coast is a lack of 'affordable housing'—that is, house prices have risen too much—that crisis has been exacerbated to a significant extent by exiting growth controls."[86]

There is an abundance of studies that demonstrate a causal connection between growth controls and a lack of affordable housing. For example, the Center for Urban and Regional Studies, of the University of North Carolina, concludes:

> As the cost of regulation drives up housing prices, demand falls and the bottom end of the housing market drops out, leaving mostly high-end houses in the building pipeline. Thus, lower- and middle-income households bear the burden, not simply through higher home ownership costs, but through the unavailability of homes in their price range."[87]

The lack of relationship between market rate housing and a shortage of affordable housing was described well in one study:

> [W]hile the development of market rate housing may generate a local need for new highway lanes or school rooms, it clearly does not create a need for more subsidized housing.[88]

Likewise, the contention that the construction of new homes creates a shortage of affordable housing has been said to contradict other economic studies that demonstrate that new home construction increases the supply of housing for lower-income persons through the "move-up" effect.[89] That is, when new housing is built and occupied by higher-income workers, there is a ripple effect whereby somewhat less well-off families move into the homes vacated by the higher-income workers, and the homes once occupied by the somewhat less well-off families are occupied by even less well-off families, and so on.

CONCLUSION

Because the economic theory and empirical economic data refute the connection between a shortage of affordable housing and the development of new market rate housing, communities wishing to impose inclusionary zoning mandates ought to have a high burden of persuasion when they try to justify such mandates with economic suppositions. Not only ought such a requirement be mandated by basic economic theory, but if the standards of *Nollan* and *Dolan* are to have any meaning, local governments must justify the imposition of inclusionary zoning mandates on developers with something more than mere pretense.

Endnotes

1 Of course, no one likes "public" or "low-income" housing, so planners renamed it "affordable housing" to make it more palatable to the middle class that ends up paying for and living next to it. More recently, planners, in a bow to the public's embrace of the value of something called "work," have renamed it yet again as "workforce housing."

2 The number varies by jurisdiction. In San Francisco, the City in early 2007 is considering lowering the threshold to two.

3 483 U.S. 825, 836-37 (1987).

4 512 U.S. 374, 385 (1994).

5 Lingle v. Chevron U.S.A. Inc., 544 U.S. 528 (2005).

6 483 U.S. 825 (1987).

7 *Id.* at 838.

8 *Id.* at 837.

9 *Id.* at 836-37.

10 512 U.S. 374 (1994).

11 *Id.* at 385.

12 911 P.2d 429 (Cal. 1996).

13 Douglas W. Kmiec, *The "Substantial Advance" Quandary: How Closely Should Courts Examine the Regulatory Means and Ends of Legislative Applications?*, 22 Zoning & Plan. L. Rep. 97, 102-04 (1999).

14 *See, e.g.*, James E. Holloway & Donald C. Guy, *A Limitation on Development Impact Exactions to Limit Social Policy-Making: Interpreting the Takings Clause to Limit Land Use Policy-Making for Social Welfare Goals of Urban Communities*, 9 Dick. J. Envtl. L. & Pol'y 1, 1 & 8 n.12 (2000); Brett Christopher Gerry, *Parity Revisited: An Empirical Comparison of State and Lower Federal Court Interpretations of Nollan v. California Coastal Commission*, 23 Harv. J.L. & Pub. Pol'y 233, 283 (1999) ("Even in the states, where a unified interpretation [of *Nollan*] might be expected to center around a leading state supreme court case, there are often multiple contradictory decisions; at the level of federal circuits, this inconsistency becomes chronic.").

15 542 N.E.2d 1059 (N.Y. 1989).

16 *Id.* at 1061-62.

17 *Id.* at 1069.

18 643 N.E.2d 479 (N.Y. 1994), *cert. denied*, 514 U.S. 1109 (1995).

19 729 N.E.2d 349 (Ohio 2000).

20 *Id.*

21 *Id.* at 358.

22 887 P.2d 360, 365 (Or. Ct. App. 1994).

23 (Citation omitted) (quoting Schultz v. City of Grants Pass, 884 P.2d 569, 573 (Or. 1994)).

24 904 P.2d 738, 746 (Wash. 1995).

25 877 P.2d 187 (Wash. 1994).

26 *Id.* at 194.

27 *Id.*

28 649 N.E.2d 384, 390 (Ill. 1995).

29 661 N.E.2d 380 (Ill. App. Ct. 1995), *cert. denied*, 519 U.S. 976 (1996).

30 *Id.* at 390.

31 57 F.3d 12 (1st Cir. 1995).

32 *Id.* at 13.

33 *Id.* at 14.

34 Similarly, the Virginia Supreme Court found a legislatively enacted inclusionary zoning law to violate the Virginia Constitution's Takings Clause in *Bd. of Sup'rs of Fairfax County v. DeGroof Enters.*, 198 S.E.2d 600, 602 (Va. 1971).

35 583 A.2d 277 (N.J. 1990).

36 *Id.* at 293.

37 456 A.2d 390 (N.J. 1983) (*Mt. Laurel II*).

38 *Holmdel Builders*, 583 A.2d at 294 (citing *Mt. Laurel II*, 456 A.2d at 436).

39 456 A.2d at 448.

40 583 A.2d at 294.

41 *Id.*

42 Indeed, a few courts have even found that development fees are not subject to Nolan's nexus test or Dolan's rough proportionality standard. *See, e.g.,* McCarthy v. City of Leawood, 894 P.2d 836, 845 (Kan. 1995). This is not only in conflict with the holdings from some of the states described in Part I(A)(3), *infra*, it is even in conflict with California. *See* Ehrlich v. City of Culver City, 911 P.2d 429, 450 (Cal. 1996), *cert. denied*, 519 U.S. 929 (1996) (finding that Dolan applies at least to individually applied fees). Furthermore, many of the state cases relied upon by the *Dolan* majority in formulating its "roughly proportional" test involved fees. As noted by Justice Stevens:

> All but one of the cases involve challenges to provisions... requiring developers to dedicate either a percentage of the entire parcel... or an equivalent value in cash... to help finance the construction of... parks, and playgrounds. In assessing the legality of the conditions, the courts gave no indication that the transfer of an interest in realty was any more objectionable than a cash payment.

Dolan, 512 U.S. at 400 (Stevens, J., dissenting).

43 Se. Cass Water Res. Dist. v. Burlington N. R.R. Co., 527 N.W.2d 884 (N.D. 1995).

44 552 N.W.2d 281 (Minn. Ct. App. 1996).

45 552 N.W.2d at 287.

46 *Id.* at 286.

47 930 P.2d 993 (Ariz. 1997), *cert. denied*, 521 U.S. 1120 (1997).

48 930 P.2d at 1000 (emphasis added).

49 512 U.S. 374 (1994).

50 544 U.S. 528, 547-48 (2005) (quoting *Dolan*, 512 U.S. at 385).

51 Nollan v. Cal. Coastal Comm'n, 483 U.S. 826, 833 n.2 (1987).

52 Some commentators have argued for an erosion between the distinction between rights and benefits, but this approach is best confined to the extension of procedural due process. *See, e.g.,* Charles Reich, *The New Property*, 73 Yale L.J. 733 (1964) (arguing for due process rights for welfare benefits, as later held in cases like *Goldberg v. Kelly*, 397 U.S. 254 (1970). *Cf.* Rodney A. Smolla, *The Reemergence of the Right-Privilege Distinction in Constitutional Law: The Price of Protesting Too Much*, 35 Stan. L. Rev. 69 (1982); *with* William W. Van Alstyne, *The Demise of the Right-Privilege Distinction in Constitutional Law*, 81 Harv. L. Rev. 1439 (1968).

53 *See* Kathleen M. Sullivan, *Unconstitutional Conditions*, 102 Harv. L. Rev. 1413, 1415 (1989).

54 Goldberg v. Kelley , 397 U.S. at 263 n.8.

55 483 U.S. at 835.

56 Epstein, in fact, posits that such cost-free regulatory restrictions that would impose the costs of a social benefit (the view) entirely on the permit seeker exceeds the scope of the police power. Richard A. Epstein, *Unconstitutional Conditions, State Power, and the Limits of Consent*, 102 Harv. L. Rev. 4, 63 (1988). This may not be the prevailing viewpoint today, but there must certainly be a point where the costs of a marginal social benefit cannot be justified, say, for example, if the Coastal Commission had placed restrictions on landowners based on impacts on the views *from* ocean-going vessels rather than Highway 1.

57 For those with an aversion to academic theory, the remainder of this section may be skipped after recognizing that the academics have provided the practitioners and the courts with no universal or practical theory of the doctrine of unconstitutional conditions.

58 *See* Sullivan, *Unconstitutional Conditions*, *supra* note 53, at 1413.

59 *See* Epstein, *Unconstitutional Conditions*, *supra* note 56, at 28. *See also* Richard A. Epstein, Bargaining with the State (1993).

60 Mitchell N. Berman, *Coercion Without Baselines: Unconstitutional Conditions in Three Dimensions*, 90 Geo. L.J. 1, 5 (2001).

61 *Id.* at 89.

62 Proponents of the modern progressive regulatory state may differ. *But see* David Thomas, *Why the Public Plundering of Private Property Rights is Still a Very Bad Idea*, 41 Real Prop. Prob. & Tr. J. 25 (2006); James S. Burling, *The Theory of Property and Why it Matters*, American Law Institute - American Bar Association Continuing Legal Education January 8-10, 2004 Eminent Domain and Land Valuation Litigation, SJ051 ALI-ABA 491; David A. Thomas, *Is the Right to Private Property a Fundamental or an Economic Right?*, American Law Institute – American Bar Association Continuing Legal Education January 4-6, 2007 Eminent Domain and Land Valuation Litigation, SM006 ALI-ABA 33.

63 Sullivan, *Unconstitutional Conditions*, *supra* note 53, at 1415.

64 *Id.* at *passim*.

65 Epstein, *Unconstitutional Conditions*, *supra* note 56, at 31. *See also,* Berman, *Coercion*, *supra* note 60, at 111 (calling the greater power/lesser power rationale "concededly thin").

66 483 U.S. at 837.

67 Sullivan, *Unconstitutional Conditions*, *supra* note 53, at 1420.

68 *Id*. at 1450.

69 Epstein, *Unconstitutional Conditions*, *supra* note 56, at 13.

70 *Id*. at passim.

71 *Id*. at 62-63.

72 The conditions in this case fail under the germaneness theory because there is no intrinsic condition between the right to just compensation and the right to challenge a governmental action.

73 It should be self-evident that if government can withhold the fundamental right of just compensation (or any other necessary right) in exchange for an agreement not to sue, then the government will obtain an unwarranted immunity for its bad acts.

74 A strong argument can be made for the case that these two bedrock rights (just compensation and petition) are too fundamental to be bargained away. Not only will the affected individuals lose, but all of society as well. The cliche that there is no price too high for freedom has some validity when one is confronted with forced bargains of fundamental rights.

75 Sullivan, *Unconstitutional Conditions*, *supra* note 53, at 1506.

76 Berman, *Coercion*, *supra* note 60, at 3 ("Regrettably, more than a century of judicial and scholarly attention to the problem has produced few settled understandings.").

77 Indeed, one study posits that inclusionary zoning mandates exacerbate housing affordability problems. *See* Robert C. Ellickson, "The Irony of 'Inclusionary' Zoning," *in* Pac. Inst. for Pub. Policy Res., Resolving the Housing Crisis (M. Bruce Johnson, ed.) 19-33 (1982) (*hereinafter* Resolving the housing Crisis).

78 Edward L. Glaeser & Joseph Gyourko, *The Impact of Zoning on Housing Affordability, Harvard Institute of Economic Research*, Discussion Paper No. 1948 (2002) (finding that zoning constraints are the cause of the lack of affordable housing in those regions with a housing crisis), *available at* http://post.economics.harvard.edu/hier/2002papers/HIER1948.pdf.

79 *See, e.g.*, Benjamin Powell & Edward Stringham, *Housing Supply and Affordability: Do Affordable Housing Mandates Work?*, Reason Policy Study No. 318, at 8-9 (Apr. 2004) (arguing that affordable housing mandates have not worked in practice and should not work in theory).

80 Advisory Commission on Regulatory Barriers to Affordable Housing, "Not in My Back Yard" Removing Barriers to Affordable Housing, 3 (1992).

81 *Id.* at 5.

82 *Id.* at 2-5. *See also* 2-1 through 2-18 (discusses in more detail the relationship between suburban land use controls and housing affordability).

83 Dr. Bernard Frieden, "The Exclusionary Effect of Growth Control," *in* Resolving the Housing Crisis, at 19-33.

84 *Id.* at 27-28.

85 Lloyd J. Mercer & W. Douglas Morgan, "An Estimate of Residential Growth Controls' Impact on House Prices," *in* Resolving the Housing Crisis, at 189-215.

86 *Id.* at 214.

87 Ctr. for Urban and Reg'l Studies, Univ. of N.C., Regulation and the Cost of New Residential Development 7 (1999). *Accord* John A. Landis, *Land Regulation and the Price of Housing: Lessons from Three California Cities*, 52 J. of Am. Planning Ass'n. 98 (1986) (examining relationship between land supply policies and price); Stephen Malpezzi, *Housing Prices, Externalities, and Regulations in the U.S. Metropolitan Areas*, 7 J. of Housing Research, 209, 236 (1996) ("regulation raises housing rents and values and lowers home ownership rates"); John Landis, et al., *No Vacancy: How to Increase the Supply and Reduce the Cost of Rental Housing in Silicon Valley*, Working Paper 96-251, at 13 (Fisher Ctr. for Real Estate and Urban Econs., U.C. Berkeley 1996) ("higher levels of apartment construction help to moderate rent increases").

88 Altshuler & Gomez-Ibanez, Regulation for Revenue: The Political Economy of Land Use Exactions 5 (1993).

89 *See also generally* John B. Lansing, et al., Inst. for Soc. Res., New Homes and Poor People, *passim* (1969).

Federalism and Separation of Powers

Federal Aid to the States

*By Chris Edwards**

In recent years, members of Congress have inserted thousands of pork-barrel spending projects into bills to reward interests in their home states. But such parochial pork is only a small part of a broader problem of rising federal spending on traditional state and local activities.

Federal spending on aid to the states increased from $286 billion in fiscal 2000 to an estimated $449 billion in fiscal 2007. The number of different aid programs for the states soared from 463 in 1990 to 653 in 2000, 814 by 2006.

The theory behind aid to the states is that federal policymakers can design and operate programs in the national interest to efficiently solve local problems. In practice, most federal politicians are not inclined to pursue broad, national goals, but are consumed by the competitive scramble to secure subsidies for their states. At the same time, federal aid stimulates overspending by the states, requires large bureaucracies to administer, and comes with a web of complex regulations that limit state flexibility.

At all levels of the aid system, the focus is on spending and regulations, not delivering quality services. And by involving all levels of government in just about every policy area, the aid system creates a lack of accountability. When every government is responsible for an activity, no government is responsible, as was evident in the aftermath of Hurricane Katrina.

The failings of federal aid have long been recognized, but reforms and cuts have not been pursued for years. Aid has spawned a web of interlocking interests that block reform, including elected officials at three levels of government, armies of government employees, and thousands of trade associations representing the recipients of aid.

Yet the system desperately needs to be scaled back, not least because the rising costs of federal programs for the elderly are putting a squeeze on the federal budget. This article examines the growth of the aid system and describes its failings.[1] Congress should begin terminating activities that could be better performed by state and local governments and the private sector.

Historical Development

Prior to the Civil War, proposals to provide federal subsidies for state and local activities were occasionally introduced in Congress, but they were routinely voted down or vetoed by presidents for being unconstitutional. In 1817, President James Madison vetoed a bill that would have provided federal aid to construct roads and canals.[2] In 1830, President Andrew Jackson vetoed a bill to provide aid for a road project in Kentucky, arguing that it was of "purely local character," and that funding would be a "subversion of the federal system."[3] In 1854, President Franklin Pierce vetoed a bill that would have provided aid to the states for the indigent insane, also citing federalism reasons.

The federal government approved grants of land to the states for schools, roads, and canal projects. However, there were no grant programs that disbursed cash to the states for ongoing activities. That started to change toward the end of the nineteenth century. The Morrill Act of 1862 provided grants of federal land to the states for the establishment of colleges. In 1879, Congress provided funds to a private, non-profit group in order to distribute to the states educational materials for the blind. In 1887, the Hatch Act provided subsidies to the states for agriculture research, and was the first cash grant program.[4] An 1888 act provided aid to the states for veterans' homes.

Federal aid activity increased substantially in the early twentieth century. The "dual federalism" of the nineteenth century was being replaced by what came to be called "cooperative federalism." When the income tax was introduced in 1913, it provided the means for policymakers to finance a large range of new federal aid programs. Here are some of the early aid laws:[5]

- The Weeks Act of 1911 provided aid to the states for forest fire prevention.
- The Smith-Lever Act of 1914 provided subsidies to land-grant colleges.
- The Federal Aid Roads Act of 1916 provided aid to the states to build highways.
- The Smith-Hughes Act of 1917 created grants for vocational education.
- The Chamberlain-Kahn Act of 1918 provided aid to the states for combating venereal disease.
- The Fess-Kenyon Act of 1920 provided aid to the states for vocational rehabilitation.
- The Sheppard-Towner Act of 1921 provided aid to the states "for the promotion of the welfare and hygiene of maternity and infancy."

These seven early aid programs had similar features and similar faults as today's aid programs. All seven programs required the states to match federal funds on a dollar for dollar basis—federal aid was called the "fifty-fifty system." Matching requirements induce excess spending and divert state-source funds from other, perhaps higher, priorities of each state. If states are induced to spend more of their funds on farm subsidies, for example, they may have less to spend on their justice systems.

The new aid programs usually mandated an expansion in state and local bureaucracies. Aid programs required the states to set up new boards and agencies to oversee government spending in the prescribed activities. The 1916 Act required states to create highway departments; the 1917 Act required the states to establish vocational education boards; the 1921

** Chris Edwards is the Director of Tax Policy Studies at the Cato Institute, and author of Downsizing the Federal Government (2005).*

Act required states to create children's bureaus; and so on. The states had to create detailed plans, file regular reports to Washington, and subject themselves to inspection by federal officials. In order to receive aid, states were often required to pass legislation that regulated state and local activities in ways sought by Congress.

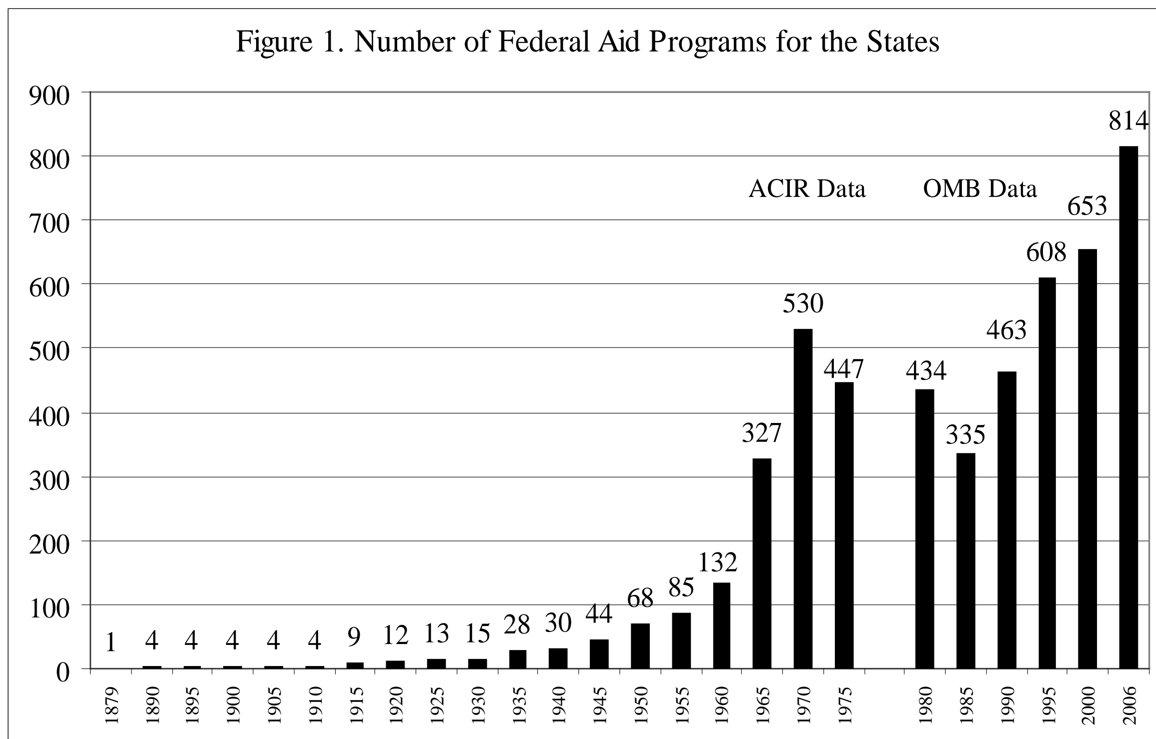
Various sleights-of-hand were used to get around constitutional concerns about expanded federal power. Funding for the 1890 Morrill Act was supposed to come from the proceeds of federal land sales in the states, but in practice the funds came from regular appropriations. The 1916 road subsidy law aimed to fund "post roads," or those that were used for federal mail delivery, but Congress defined it extremely broadly.[6] The 1911 aid bill was supposed to fund state forest fire prevention only near navigable rivers, providing a constitutional pretence for these activities as being related to interstate commerce.[7]

Figure 1 shows the number of aid programs for the states beginning with the education program of 1879. By 1930 there were fifteen federal aid programs. It was getting harder to hold the line on federalism as politicians became increasingly activist and new lobby groups were established. Labor unions pushed for federal funding of vocational education and succeeded with the passage of the Smith-Hughes Act in 1917. The passage of the 1916 road bill was preceded by the introduction of at least sixty-two different road subsidy bills in Congress.[8] State highway officials had formed a national organization in 1914 with an office in Washington to press for aid, and highway lobby groups helped draft the 1916 bill.[9]

Nonetheless, there was resistance to the growth in aid, and the shortcomings of aid programs were already becoming evident.[10] President Calvin Coolidge was a frequent and pointed critic of the aid system. In his budget message for 1926, Coolidge declared:

> I am convinced that the broadening of this field of activity is detrimental both to the federal and state governments. Efficiency of federal operations is impaired as their scope is unduly enlarged. Efficiency of state governments is impaired as they relinquish and turn over to the federal government responsibilities which are rightfully theirs. I am opposed to any expansion of these subsidies.[11]

Some leaders in the higher-income East Coast states strongly opposed expansions in aid. Governor Albert Ritchie of Maryland said that the "system ought to be abolished, root and branch" with the money "left in the states for the states to use for their own local needs and purposes."[12]



Sources: Advisory Commission on Intergovernmental Relations (1879-1975), Office of Management and Budget (1980-2000), and author's estimate based on OMB methodology (2006).

However, aid proponents were persistent, and as aid bills began to pass, new interest groups were formed and Congress was bombarded with requests for subsidies.[13] A few states initially refused to take part in some of the aid programs, but an observer at the time said that for most states "to get a few millions they shamelessly barter away their birthright" of reserved powers under the Constitution.[14]

By the time President Franklin Roosevelt came to office, many legal and political precedents had already been set for the large expansions in aid enacted under the New Deal. In the 1930s, aid programs were created for public housing, welfare, employment, and many other activities. The Federal Emergency Relief Act of 1933 provided more than $3 billion to the states over two years for work relief.

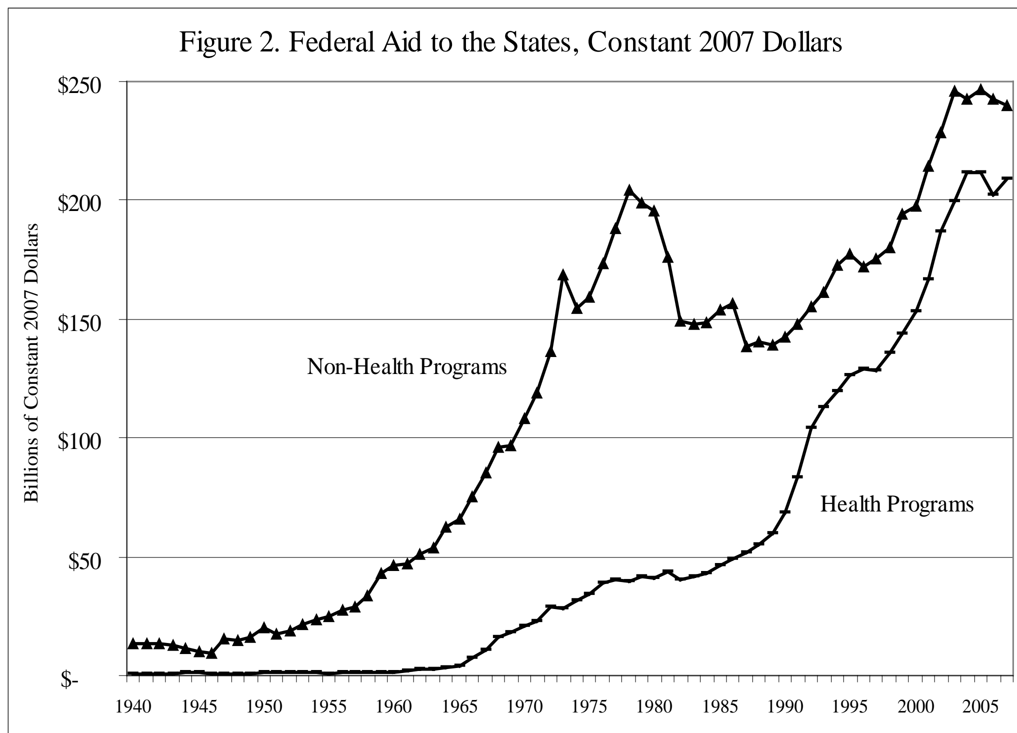
Federal aid expanded during the 1950s, with the number of aid programs almost doubling from 68 in 1950 to 132 by

1960. That expansion occurred despite President Dwight Eisenhower's expressed concerns about federalism. Eisenhower had established a commission in 1953 to identify federal activities that could be returned to the states, but unfortunately no reforms were enacted.[15] The largest new grant program during this era was the Federal-Aid Highway Act of 1956.

The Highway Act provides an illustration of how federal regulatory controls started coming as a package deal with federal dollars. The Act imposed Davis-Bacon rules on all state highway projects that received federal money, which mandated that construction workers be paid "prevailing wages." That usually had the effect of increasing labor costs on projects at taxpayer expense. Because many states were already constructing their own highway systems in the mid-1950s, one effect of the 1956 Act was to increase the costs of many highways that would have been built anyway.

revenues available to individual states. Of course, the ultimate source of federal money distributed to the states is the taxpayers who live in the fifty states, but policymakers have always ignored this inconvenient truth.

In addition, aid advocates have been infused with a belief in the "public interest theory of government," which is the idea that policymakers act with the best interests of the broad general public in mind. In past decades, conventional wisdom held that the federal government could be effective and efficient at solving local problems. Are there poor people and blighted buildings in your city? Let us use the seemingly unlimited resources of the federal government to hire experts, bulldoze the blight, and build modern high-rises to solve the problem. That type of top-down thinking was behind the creation of aid programs for urban renewal, housing, education, and many other activities.



Federal aid exploded during the 1960s. Figure 1 shows that the number of aid programs quadrupled from 132 in 1960 to 530 by 1970. Under President Lyndon Johnson, new aid programs were added for housing, urban renewal, education, health care, environmental protection, and many other activities. Aid spending rose from $47 billion in 1960 to $129 billion by 1970, measured in constant 2007 dollars, as shown in Figure 2.

President Johnson called his policies "creative federalism," but his activism dealt a severe blow to the federalism of the nation's Founders. By the end of the 1960s, many policymakers believed that the federal government should spend money on just about any activity that it wanted, and questions regarding constitutional propriety were seldom considered anymore.

The huge growth in federal aid in the 1960s occurred for many reasons. Policymakers were fooled by the mirage that "federal resources" appeared to be limitless compared to the tax

We know today that the federal government is not very good at solving local problems. Even casual observation of Congress reveals that policymakers put the various narrow interests of their states above all else most of the time. Aid is the perfect tool to satisfy parochial special interests, and that is why the aid empire prospers today—not because experts believe that it works well.

Growing Aid, Growing Problems

The unchallenged optimism of the 1960s about the federal government's ability to solve state and local problems did not last. By the late 1960s, budget analysts were becoming alarmed at the growing complexity and overlap of federal grants. Two of President Johnson's top economic advisors and other experts began to push for consolidation of narrow "categorical" grants into broader and more flexible "block" grants.[16]

Categorical grants fund a narrow range of eligible activities based on detailed federal rules that state governments are required to follow. Until the 1960s, all grants were categorical grants; the vast majority still are. Categorical grants are very complex. As far back as the Hoover Commission in 1949 experts had proposed replacing them with block grants.[17] Congress passed the first block grant in 1966, which converted sixteen extant health care grants into a single broader program. A block grant for law enforcement was enacted in 1968. But these small reforms were overwhelmed by the avalanche of new categorical grants enacted in the late 1960s. The number of programs seems to have reached a temporary peak in the early 1970s, and then declined for a few years as the pace of program creation slowed and some existing programs were consolidated. While fewer new programs were added during the 1970s, the cost of aid programs soared as spending on all the programs created during the 1960s kicked into overdrive. President Richard Nixon took some modest steps toward consolidating the burgeoning aid system into block grants, arguing that federal aid had become a "terrible tangle" of overlap and inefficiency.[18] In his 1971 State of the Union address, Nixon lambasted "the idea that a bureaucratic elite in Washington knows what is best for the people."[19]

Nixon's reforms fell far short of his rhetoric, and just a few of his "new federalism" initiatives were enacted. Nixon succeeded in creating three block grants.[20] In addition, an extreme form of block grant, "revenue sharing," was begun in 1972 to give funding to the states with almost no strings attached. The problem was that revenue sharing did not substitute for existing grants—it was added on top. The program was abolished in 1986.

Consolidation of narrow grants into broader block grants made sense to budget experts, but members of Congress usually favored categorical grants because they could be better targeted toward special interests. The expansion of categorical grants was in sync with the increasingly fragmented committee structure in Congress in the mid-twentieth century. That is, the number of aid programs grew as the number of committees and subcommittees grew. Each committee and subcommittee wanted its own realm of programs over which to preside.

During the 1970s, concern grew over the complexity of the mushrooming federal aid system. When President Jimmy Carter came into office he proposed a "concentrated attack on red tape and confusion in the federal grant-in-aid system." [21] Carter pursued a number of modest reforms, but, like Nixon, he did a better job describing the aid problem than enacting solutions.

The bipartisan Advisory Commission on Intergovernmental Relations (ACIR) criticized many aspects of the federal aid system through the 1960s, 1970s, and 1980s. In 1980, ACIR published an eleven-volume study on federalism, which concluded that Washington's power had become "more pervasive, more intrusive, more unmanageable, more ineffective, more costly, and, above all, more unaccountable."[22] In an ironic twist, the ACIR, a rare government agency that criticized government programs, was one of the few agencies abolished by the Republicans in the 1990s.

The Government Accountability Office also criticized the federal aid system. In 1979 it found:

> The federal assistance system is an array of often conflicting activities and initiatives which defy understanding to all but the most serious students of the system … During the 1960s, the explosion in the number of federal programs made shortcomings in the [aid] system apparent. Studies showed that red tape, delays, and vast amounts of paperwork were characteristics common to most federal [aid] programs. [23]

President Ronald Reagan came into office promising to respond to such concerns and to revive federalism. He tried to cut aid spending, convert existing aid programs to block grants, and transfer some activities back to the states. The Omnibus Budget Reconciliation Act of 1981 eliminated fifty-nine grant programs and consolidated eighty categorical grants into nine block grants.[24]

In 1982, under his "new federalism" agenda, Reagan sought to re-sort federal and state priorities, such that each level of government would have full responsibility for certain activities. For example, Reagan proposed that welfare and food stamps be both financed and operated by the states.[25] Reagan also proposed "turnback" legislation to end about forty federal programs. Thus, he proposed that the federal government end most highway programs, while canceling the federal gasoline taxes that supported them.

In his 1983 budget message, Reagan argued that "during the past 20 years, what had been a classic division of functions between the federal government and the states and localities has become a confused mess."[26] Reagan had some success in cutting federal aid. Both aid spending and the number of aid programs were cut substantially during the early 1980s. Data from the Office of Management and Budget show that the number of aid programs for the states was cut from 434 in 1980 to 303 in 1982, before beginning to rise again.[27]

Reagan's progress in cutting aid programs was reversed by President George H. W. Bush. Aid spending and the number of programs grew rapidly in the early 1990s. Then, in the mid-1990s, the new Republican congressional majority tried again to revive federalism. Their biggest success was welfare reform in 1996, which turned open-ended categorical welfare aid into a block grant.

Since the mid-1990s, there have been no serious efforts to reform or cut the federal aid system, even though it is larger and costlier than ever. The system's many failings, which were discussed often during the 1970s and 1980s, have only become more acute as hundreds of new programs have been added since then. The current Bush administration has expanded the aid system rather than trying to restrain it. Department of Education outlays have doubled since 2000, as President Bush has taken steps to further nationalize local public schools. In other areas, the Bush administration's faith-based and marriage initiatives have hooked thousands of private organizations on federal subsidies. Richard Nathan, an architect of Nixon's new federalism, opined that Bush's policies "have reflected a willingness to run roughshod over state governments that is out of character with previous Republican administrations."[28]

Federal Aid Today

The number of aid programs increased from 653 in 2000 to 814 in 2006.[29] Aid spending rose from $351 billion in fiscal

2000 to an estimated $449 billion in fiscal 2007, measured in constant 2007 dollars.[30] Federal aid to the states is the third-largest item in the federal budget after Social Security and national defense. The five largest federal aid programs are Medicaid, highway construction, Temporary Assistance for Needy Families, education for the disadvantaged, and Section 8 housing subsidies.[31]

The 814 federal aid programs for the states generally take the form of either "formula" or "project" grants.[32] While most aid programs are project grants, most aid spending is on formula grants. That is because many of the largest aid programs, including Medicaid, are formula grants.

Under formula grants, legislation spells out how much funding each state is to receive based on factors such as state income and population. Under project (or "discretionary") grants, federal agencies distribute thousands of individual grants on a competitive basis after an expert review of proposals, at least that is how it is supposed to work.

One form of "discretionary" aid is earmarking. That occurs when the grant process is hijacked by individual members of Congress seeking to divert funds to particular projects in their districts. Thus, while a federal agency might normally distribute cancer research grants based on an expert review of proposals, politicians can end-run around the agency and directly target funds to health facilities in their districts.

Earmarking has exploded in recent years, and numerous congressional scandals have stemmed from the practice. The number of earmarks in federal spending bills increased from under 2,000 per year in the mid-1990s to more than 15,000 per year recently.[33]

This article focuses on aid to state governments, but the federal government also has hundreds of programs that directly provide subsidies to businesses, non-profit groups, and individuals. A complete list of federal aid programs is the 2,437-page Catalog of Federal Domestic Assistance (CFDA).[34] In 2006, the CFDA listed 1,696 federal subsidy programs, of which 814 were for state and local governments, as noted.

Despite general agreement among experts that block grants are superior to narrow categorical grants, only about twenty of the more than 800 state aid programs are block grants.[35] The reason notes the GAO is that "legislation addressing a specific need holds far more political appeal than broader purpose block grant programs. Any effort to incorporate categorical programs into a broader purpose program is interpreted as an attack on the congressional committees who created the programs, the agencies who administer them, and the clientele groups who prosper."[36]

Thus, the federal aid system is not about financing and operating programs in the most efficient manner; it is about politics. Politicians, special interests, and aid recipients resist conversion of programs to block grants because that would reduce their control and make programs easier to cut. One can debate whether or not federal aid is a good idea in theory, but in practice the political system has locked the nation into the most complex and inefficient form of aid: categorical aid.

Both block and categorical grants involve top-down control of state and local activities from Washington. Both types of grant lead to the creation of large bureaucracies. Nevertheless, converting categorical grants to block grants would represent progress because it would make federal costs more controllable and aid spending easier to cut. A good first step toward restraining Medicaid's explosive spending growth, for example, would be to convert it to a block grant.

Today's federal aid structure is massive and complex. The three layers of government in the United States no longer resemble the tidy layer cake that existed in the 19th century, but a jumbled marble cake. Federal expansion into policy areas traditionally reserved to the states has proven to be a wasteful and bureaucratic way of governing the nation.

One problem is that Congress provides nowhere near enough oversight for aid programs. Members get a political payoff from setting up aid programs and pushing to increase spending, not from pruning those that do not work. The Bush administration has performed detailed reviews of 257 federal aid programs and found that 109 of them were "ineffective" or could not "demonstrate results."[37] Yet Congress has shown little interest in cutting or terminating those programs.

Federal Aid and Politicians

Over the decades, policymakers have argued that various state, local, and private activities needed federal intervention because they had become "national priorities." A fact sheet from the Secretary of Education, Margaret Spellings, begins:

> The responsibility for K-12 education rests with the states under the Constitution. There is also a compelling national interest in the quality of the nation's public schools. Therefore, the federal government… provides assistance to the states and schools in an effort to supplement, not supplant, state support.[38]

The flaw in logic here is that there are few activities the federal government performs that are not also priorities of individuals, businesses, and state and local governments. One can call education a "national" priority, but that does not mean that the federal government has to get involved. That is because education is also a high priority of local governments and families. Local governments are free to learn schooling techniques from each other, but there is no compelling interest for top-down control from Washington.

President Ronald Reagan made the following observation in his 1987 executive order on federalism:

> It is important to recognize the distinction between problems of national scope (which may justify federal action) and problems that are merely common to the states (which will not justify federal action because individual states, acting individually or together, can effectively deal with them).[39]

The confusion between problems that are truly national in scope and those that are merely common to the states even extends to homeland security. When you look at the details of federal aid to the states for homeland security, you find that much is going toward items that would be better funded locally, such as bulletproof vests and radio systems for first responders. When this sort of local spending is federalized, members of Congress play a game of tug-of-war over funding for their states, and put less emphasis on taxpayer value for money in their decisions.

The idea that aid to the states can be designed in the "national interest" is a theory that does not match political reality. The concern of members of Congress for their states and

districts almost always trumps any other policy considerations. Members may convince themselves that spending on aid projects in their hometowns is good for the country, but that is only because the resulting tax burdens are spread over the rest of the nation and invisible to them. Aid proponents say that it is in the national interest to help those people and regions of the country with the greatest needs. But in practice, the aid system has never operated in that fashion. A 1940 article in *Congressional Quarterly* lamented that "the grants-in-aid system in the United States has developed in a haphazard fashion. Particular services have been singled out for subsidy at the behest of pressure groups, and little attention has been given to national and state interests as a whole."[40] Forty years later, the ACIR concluded essentially the same thing: "Regarding *national purpose,* the record indicates that federal grant-in-aid programs have never reflected any consistent or coherent interpretation of national needs."[41]

In the operation of the aid system, political and parochial concerns are far more important than national priorities. The problem is not that members are not patriots, it is that they are also activists and—like most people—they have emotional and community ties to their hometowns. Of course, even before the modern grants-in-aid system, federal politicians championed spending activities that benefited their home states. Legislators with navy bases in their states have always supported navy spending, for example. But the expansion of the aid system in recent decades has magnified the age-old regional battles in Congress.

The recent explosion in the "earmarking" of federal aid and procurement has taken geographic political competition one step further.[42] Some earmarking misallocates resources for properly federal activities such as defense. But most earmarking is for federal spending on properly state, local, and private activities. By opening the floodgates to earmarking, Congress has encouraged a stampede of local interests to beat a path to Capitol Hill. Local governments and local organizations are increasingly making end-runs around state officials and going straight to Congress whenever they need a new parking lot, museum, or airport terminal.[43]

Earmarking is tied to recent corruption scandals. Disgraced lobbyist Jack Abramoff famously called the appropriations committees in Congress "favor factories." Indeed, they are. Politicians trade earmarks for campaign assistance, trips, sweetheart business deals, and general political support. Total fees paid to registered lobbyists in Washington have increased from $100 million in 1975 to $2.5 billion in 2006, with a substantial share of those fees related to earmark lobbying.[44] Recent scandals have shown that federal politicians cannot keep their hands out of the cookie jar, but the fundamental problem is that the federal cookie jar has grown so large. With 814 state aid programs and 1,696 federal subsidy programs overall, it is not surprising that the number of earmarks has soared because each program is a delivery vehicle for favors to home-state interests. The earmarking explosion was a scandal waiting to happen.

Parochial politics feeds on itself and has created a dynamic response from the states. The more aid programs and earmarking, the more federal lobbying state and local officials and interest groups will do. Highway contractors, school teachers, and policemen have learned that the payoff from the one-stop-subsidy-shop in Washington is higher than the payoff from lobbying each state separately.

Earmarks represent just a part of the regional skirmishing in Congress. The formulas used for distributing aid are a bigger battleground. Consider the ongoing fights over the formulas used to distribute homeland security aid. Homeland security aid has often gone to regions that don't need it in order to buy expensive items that are little used. Members of Congress also battle over health care grants. A *Washington Post* story profiled Senator Hillary Clinton's (D-NY) fight to tweak the formula that distributes federal grants for HIV/AIDS so that a little more flows to New York.[45] By engaging in such a fight, Clinton is signaling to her constituents that she is a champion for their interests. The efficiency of programs and their positive or negative effects are not politically important. It is spending that generates the favorable media coverage.

The States: America's Biggest Lobby?

Governors and other state leaders are putting increasing efforts into securing federal aid spending. The Republican Governor of Texas, Rick Perry, is considered to be a conservative, but his official webpage is chock full of press releases touting his hand-outs of federal subsidies such as "Gov. Perry Announces $1.6 Million in Grants to Juvenile Offender Accountability Programs."[46] As federal aid has increased, governors have become less like chief executives and more like regional deputies for the federal government. Since the explosion of federal aid in the 1960s, state and local governments have become major lobbyists in Washington. The ACIR reported in 1967 that "grantsmanship has become a popular new game in Washington."[47] The *Wall Street Journal* published a story in 1966 about the new profession of "grantsman."[48] Grantsmen were the high-paid middlemen who benefited from the maze of President Johnson's new state aid programs.

Many state and local interest groups were organized, or greatly expanded their Washington offices, during the 1960s. By 1967, thirteen states and twenty-four cities and counties had established Washington offices to lobby for aid.[49] Today there are 88,000 state and local government entities in the United States, including cities, counties, towns, school districts, and special districts.[51] Most receive—and many actively solicit—federal funding. All these governmental units, and their 16 million employees, represent a powerful lobby in support of aid programs and the vast federal welfare state.

As the number of aid programs has grown, state and local officials have put increasing efforts into federal lobbying. For example, it is routine for local groups across the country to organize "fly-ins" to Washington for personal arm-twisting on Capitol Hill. One recent news article profiled fly-ins by officials from California counties.[51] Local groups pay Washington lobbying firms to organize their meetings and strategies, and each group comes equipped with a wish list of local projects that they want funded.

Who can blame today's state and local officials putting so much effort into lobbying? There are winners and losers in the federal fiscal roulette. An analysis of federal aid to the

states for 2004 found large variations between jurisdictions.[52] The biggest recipients of aid on a per-capita basis were the District of Columbia ($7,445), Alaska ($4,972), and Wyoming ($3,268), while the smallest recipients were Nevada ($1,045), Virginia ($1,085), and Florida ($1,158). State governments treat federal aid like a goldmine, and they use a multi-pronged strategy to secure their share of aid nuggets. Texas, for example, has an Office of State-Federal Relations that provides news from Congress on aid programs and works with Texas agencies to maximize federal funding.[53] Maryland has a sophisticated grants agency that was created to tackle the "increasing competition with other states" for federal aid.[54] The agency seeks to increase Maryland's "market share" of aid through activities such as "relationship building" with federal aid decisionmakers. In California, a major performance review of state government under Governor Arnold Schwarzenegger found that the state "does not receive its fair share of federal grant funds."[55] The report examined the issue in detail and proposed that the state "develop aggressive strategies" for "maximizing" federal aid, including creating a new office to better coordinate aid efforts.[56]

Many states have created "think tanks" to research how to increase federal aid. California has the California Institute for Federal Policy Research in Washington D.C., which operates a sophisticated tracking system for federal legislation.[57] This organization has a corporate board stacked with California politicians and business leaders. California also has the Public Policy Institute of California based in San Francisco, which provides frequent reports regarding California's share of federal aid funding.[58] The Northeast-Midwest Institute represents a group of eighteen states stretching from Vermont to Minnesota. The Institute's website says that it publishes the "most detailed analysis of the flow of federal funds to the states, demonstrating the persistent federal disinvestment in Northeastern and Midwestern states."[59] The website notes that the Institute helped "protect Amtrak routes in the region," "altered the food stamp program's criteria to take into account higher costs of living in cold climates," "defeated persistent attempts by southern lawmakers to change the match rate for Medicaid and welfare payments to the detriment of the Northeast-Midwest," and "established a dual Community Development Block Grant (CDBG) funding formula that helps rebuild older communities."[60] The CDBG program illustrates how technical the battles over aid can be. One item in the formula that distributes CDBG funding is "housing built before 1940." How did this obscure item get into the CDBG formula? The Northeast-Midwest Institute got a member of Congress to insert it into legislation in 1977 in order to tilt aid toward older cities.

Or consider the Public Policy Forum of Southeastern Wisconsin. It argues that this region of the country is at a "competitive disadvantage" because of a "failure to take full advantage of federal grants."[61] The government and business leaders of this group are taking an aggressive strategy to fix the problem. A recent report says that "competitive federal dollars drive economic growth... federal funding is a diverse source of capital that fuels discovery and wealth creation."[62] The report urges that local leaders hire staff and raise money for efforts to maximize inflows of federal dollars. Wisconsin groups should hire grant experts, travel to Washington two to four times per year, and phone federal agencies weekly. And they should raise private money to hire the experts needed to grab federal grants. Clearly, federalism has descended into a highly professionalized competitive battle between the states—and against federal taxpayers. The number of state and local governments that have hired high-priced Washington lobbyists has doubled since 1998.[63] One lobbying firm, Alcade & Fay, has a dedicated "Municipalities Practice Group," which generates $4 million annually in fees. Such firms typically charge their state and local government clients $10,000 to $20,000 per month. Alcade & Fay boasts that it has "secured billions of dollars in earmarked appropriations and federal grants."[64]

Because the federal budget is a goldmine for the states, it is not surprising that state and local officials invest in high-priced prospectors. Perhaps the most successful prospector is Gerald Cassidy, co-founder of the Washington lobbying firm Cassidy and Associates. The firm has been the focus of a recent series of articles in the *Washington Post*.[65] Cassidy and his firm pioneered the now-common practice of earmarking money for state and local spending projects in the federal budget. Cassidy's efforts have enabled him to amass a personal fortune of $125 million. The *Washington Post* series reveals that the expansion of federal spending on state and local activities has not just been driven by activist politicians on Capitol Hill. Entrepreneurial lobbyists, such as Cassidy, have played a key role in advancing the process by pro-actively selling their services to universities and other local institutions across the country. These days, state and local officials know that Washington lobbyists are helping most other jurisdictions secure federal cash, so if they sit on their hands or are squeamish about paying for lobbyists, they will lose out.

The time has long passed when state policymakers would jealously guard the independence of state activities and resist federal encroachment. These days, the priority of the states is to use every means available to squeeze more money from federal taxpayers. State officials have complained about the onerous rules of the No Child Left Behind law of 2002, and thirty state legislatures passed resolutions attacking NCLB for undermining states' rights. But the states did not call for repeal of the education law, they simply demanded more federal aid money to spend on NCLB implementation.

CONCLUSION

Under the federal aid system, about $500 billion flows into Washington each year from taxpayers in the fifty states. The funds are allocated by power brokers in Congress and routed through the federal bureaucracies. Then, somewhat depleted, the funds are sent back down to the states coupled with thousands of pages of federal regulations to comply with. It is a roundabout funding system that serves no important economic purpose. If it was shut down, state governments and the private sector would step in and fund those activities they think are worthwhile. During the 1970s and 1980s, government auditors, official commissions, and many analysts determined that the aid system needed major reforms. Ronald Reagan put

the system on a diet for a few years, but the core pathologies were not addressed. Since then, hundreds more programs have been added to the system, the costs have grown higher, and the parochial battles over aid are bigger than ever.

The top-down micromanagement that comes with aid smothers policy diversity in the states. Aid mutes beneficial tax competition between the states. Aid destroys political accountability—when programs fail, politicians usually point fingers of blame at other levels of government. The federal aid system has been called a "the triumph of expenditure without responsibility." With the coming federal budget crunch from rising costs in Social Security and Medicare, the aid system is an ideal place to find budget savings. Initial reform steps should include converting Medicaid to a block grant to control costs and terminating hundreds of lower-priority aid programs. Cutting the aid system will require heavy political lifting, because the system is deeply entrenched. There are tens of thousands of state and local governments, unions, trade associations, and other groups addicted to the flows of dollars from Washington, and they will try to block any reforms. Ronald Reagan showed that aid can be cut, but it will take a fundamental challenge from another determined and reform-minded president.

Endnotes

1 The full version of this study on the Cato website is Chris Edwards, *Federal Aid to the States: Historical Cause of Government Growth and Bureaucracy*, *available at* www.cato.org/pub_display.php?pub_id=8246.

2 James Madison, Veto Message to Congress for the Bonus Bill, Mar. 13, 1817, *available at* www.constitution.org/jm/18170303_veto.htm.

3 Andrew Jackson, Veto Message for the Maysville Road Bill, May 27, 1830, *available at* www.gutenberg.org/files/10858/10858-h/10858-h.htm.

4 U.S. Gov't Accountability Office, Federal Grants: Design Improvements Could Help Federal Resources Go Further, GAO/AIMD-97-7, 4 (Dec. 1996).

5 For detailed early discussions of these grant programs, *see* G. B. Galloway, *Federal Subsidies to the States*, Cong. Q. (Dec. 13, 1924); *see also* Austin F. MacDonald, Federal Aid: A Study of the American Subsidy System (1928). Another good source for the history of grants is Advisory Comm'n on Intergov'tal Relations, Categorical Grants: Their Role and Design, Rep. A-52, May 1977.

6 MacDonald, *id.* at 90. *See generally also* Richard F. Weingroff, *For the Common Good: The 85th Anniversary of a Historic Partnership*, 64 Public Roads 5 (Fed. Highway Admin. Mar./Apr. 2001) *available at* http://www.fhwa.dot.gov/infrastructure/rw01a.htm.

7 MacDonald, *supra* note 5, at 30.

8 Galloway, *supra* note 5.

9 Weingroff, *supra* note 6.

10 *See* Galloway, *supra* note 5. *See also* James M. Beck, Our Wonderland of Bureaucracy (1932) ch. XVI.

11 Galloway, *supra* note 5.

12 MacDonald, *supra* note 5, at 238.

13 Galloway, *supra* note 5.

14 *Id.*

15 This was the Commission on Intergovernmental Relations. *See* Laurence J. O'Toole, "Overview," *in* American Intergovernmental Relations 9 (Laurence J. O'Toole, ed., CO Press, 1985).

16 Timothy J. Conlan, *The Politics of Federal Block Grants: From Nixon to Reagan*, 99 Pol. Sci. Q. 250-251.

17 *Id.* at 250.

18 Ben Canada, *Federal Grants to State and Local Governments: A Brief History*, Cong. Research Serv., RL-30705, 9 (Feb. 19, 2003).

19 Richard M. Nixon, Annual Message to the Congress on the State of the Union, (Jan. 22, 1971), *available at* www.nixonlibrary.org.

20 U.S. Advisory Comm'n on Intergov'tal Relations, Characteristics of Federal Grant-In-Aid Programs 8 (June 1995).

21 U.S. Advisory Comm'n on Intergov'tal Relations, The Intergovernmental Grant System: An Assessment and Proposed Policies 24 (Jan. 1978).

22 U.S. Advisory Comm'n on Intergov'tal Relations, The Federal Role in the Federal System: The Dynamics of Growth 13 (Dec. 1980).

23 U.S. Gov't Accountability Office, Perspectives On Intergovernmental Policy and Fiscal Relations, GGD-79-62, 3 (June 28, 1979).

24 U.S. Gov't Accountability Office, Federal-State-Local Relations: Trends of the Past Decade and Emerging Issues, GAO/HRD-90-34, 15 (Mar. 1990).

25 John L. Palmer & Isabel V. Sawhill, "Perspectives on the Reagan Experiment," *in* The Reagan Experiment 12 (John L. Palmer & Isabel V. Sawhill, eds., Urban Institute, 1982).

26 Ronald W. Reagan, Budget Message of the President, Budget of the U.S. Gov't, Fiscal Year 1983 M22 (U.S. Gov't Printing Office, Feb. 1982).

27 Office of Mgmt & Budget, Budget Systems and Analysis Division, Executive Office of the President, The Number of Federal Grant Programs to State and Local Governments: 1980 to 2004 (April 8, 2005) (unpublished).

28 Richard Nathan, There Will Always Be a New Federalism 7 (Rockefeller Inst. of Gov't, Dec. 14, 2005).

29 Author's update of data from the Office of Management and Budget. The OMB found 770 aid programs for state and local governments in 2004, based on their analysis of the *Catalog of Federal Domestic Assistance*. I updated the data based on the *CFDA* for November 2006. Because the OMB uses some discretionary judgment in its program count, my count might differ from theirs if they were to update to 2006. Thanks to Cato intern Emmanuel Caudillo for his research assistance.

30 Office of Mgmt & Budget, Executive Office of the President, Budget of the U.S. Government, Fiscal Year 2008, Analytical Perspectives 107 (Gov't Printing Office, Feb. 2007). The author converted the budget's constant 2000 dollars to constant 2007 dollars.

31 Office of Mgmt & Budget, Executive Office of the President, Budget of the U.S. Government, Fiscal Year 2008, Historical Tables, Table 12.3.

32 Aside from formula and discretionary grants, 41 of the 814 are "direct payments for specified use."

33 Chris Edwards, *Pork: A Microcosm of the Overspending Problem*, Cato Inst. Tax & Budget Bulletin No. 24 (Aug. 2005). The Congressional Research Service has a somewhat different earmark count. *See* Brody Mullins, *As Earmarked Funding Swells, Some Recipients Don't Want It*, Wall St. J., Dec. 26, 2006, at A1.

34 *Available at* www.cfda.gov.

35 A 1990 Urban Institute book called support for block grants a "consensus," but that consensus may have weakened in recent years. *See* Michael Fix & Daphne Kenyon, "Introduction," *in* Coping with Mandates 1 (Urban Inst. Press, 1990).

36 U.S. Gov't Accountability Office, Perspectives On Intergovernmental Policy And Fiscal Relations, GGD-79-62, 24 (June 28, 1979).

37 Office of Mgmt & Budget, *supra* note 30, at 105.

38 Margaret Spellings, 10 Facts About K-12 Education Funding, U.S. Dep't of Educ. (June 2005), *available at* www.ed.gov/about/overview/fed/10facts/10facts.pdf.

39 Exec. Order No. 12612, 52 Fed. Reg. 41, 685 (Oct. 26, 1987), *available at* www.reagan.utexas.edu/archives/speeches/1987/102687d.htm.

40 B. Putney, *Federal-State Relations Under Grants-In-Aid*, Cong. Q. (July 30, 1940).

41 U.S. Advisory Comm'n on Intergov'tal Relations, 10 The Federal Role in the Federal System: The Dynamics of Growth 94 (June 1981). Michael Greve's writing alerted me to this statement.

42 Dan Morgan & Helen Dewar, *GOP Dishes Out Pork in Growing Portions*, Wash. Post, Nov. 24, 2003, at A19.

43 The State of North Carolina is so tired of having its planning undermined by local officials that it has hired a lobbyist to ensure that the state's congressional delegation focuses on funding state requests only. *See* Pat Stith, *Road Planning Short-Circuited in Congress*, Raleigh News-Observer, Oct. 29, 2006, *available at* http://www.newsobserver.com/news/growth/traffic/story/504213.html.

44 *See generally* Robert G. Kaiser, *Citizen K Street*, Wash. Post Online, Mar. 4-Apr. 8, 2007, *available at* http://blog.washingtonpost.com/citizen-k-street. Broader measures of the size of the lobbying industry put total lobbying fees in the $20 billion range.

45 Jeffrey Birnbaum, *Sen. Clinton Delays AIDS Law's Renewal, Citing Cut in N.Y. Funds*, Wash. Post, Aug. 23, 2006, at A1.

46 *Available at* www.governor.state.tx.us. All releases appeared during August 2006.

47 U.S. Advisory Comm'n on Intergov'tal Relations, 1 Fiscal Balance in the American Federal System 151 (Oct. 1967).

48 Lewis Phelps, *Grantsmanship: New Specialists Help Communities*, Wall St. J., Nov. 22, 1966.

49 U.S. Advisory Comm'n on Intergov'tal Relations, *supra* note 48, at 151.

50 U.S. Dep't of Commerce, Bureau of the Census, Compendium of Public Employment: 2002, 14 (U.S. Gov't Printing Office, Sept. 2004).

51 Michael Doyle, *Stanislaus Officials Lobby Congress for Funds*, ScrippsNews, Mar. 14, 2007, *available at* www.scrippsnews.com/node/20193.

52 Curtis Dubay, *Federal Tax Burdens and Expenditures by State* (Tax Foundation, Mar. 2006).

53 *Available at* www.osfr.state.tx.us.

54 *Available at* www.gov.state.md.us/grants.

55 Governor of the State of California, "California Performance Review," August 3, 2004, http://cpr.ca.gov. *See* issue No. GG07, *at* http://cpr.ca.gov/report/cprrpt/issrec/gg/ir/gg07.htm.

56 Governor of the State of California, "California Performance Review," August 3, 2004, http://cpr.ca.gov. *See* issue No. GG07, *at* http://cpr.ca.gov/report/cprrpt/issrec/gg/ir/gg07.htm.

57 *Available at* www.calinst.org.

58 *Available at* www.ppic.org/main/series.asp?i=22.

59 *Available at* www.nemw.org.

60 *See* "What We've Done," *available at* www.nemw.org/past.htm.

61 Pub. Pol'y Forum, High-Stakes Game of Risk, Regional Report, vol. 3, no. 3 (Mar. 2006), *available at* http://www.publicpolicyforum.org.

62 *Id*. at 1.

63 Jodi Rudoren & Aron Pilhofer, *Hiring Federal Lobbyists, Towns Learn Money Talks*, N.Y. Times, July 2, 2006.

64 Cited in Gail Russell Chaddock, *Psst. K Street Delivers the Goods—For a Price*, Christian Science Monitor, Aug. 8, 2006.

65 *See generally,* Kaiser, *supra* note 44.

Financial Services and E-Commerce

Securities and Exchange Commission Issues Guidance on Management's Evaluation of Internal Control over Financial Reporting

*By Daniel Fisher**

The Securities and Exchange Commission (SEC) recently issued interpretative guidance ("Guidance") in connection with corporate management's evaluation of internal control over financial reporting ("ICFR"), which is required for under Section 404 of the Sarbanes-Oxley Act of 2002 ("SOX 404") for companies required to file reports under the Securities and Exchange Act of 1934.[1] The Guidance is part of a combined effort by the SEC and the Public Company Accounting Oversight Board (PCAOB)—which recently adopted a new auditing standard relating to audits of ICFR—to address the growing criticism that the costs of compliance with SOX 404 far exceed the benefits.[2] The Guidance is intended to make the ICFR evaluation process more efficient and cost-effective, and is designed to apply to companies of varying sizes and complexities.

In related actions, the SEC amended its rules to clarify that a management evaluation of ICFR conducted in accordance with the Guidance is one (but not the only) way to satisfy management's obligation to conduct the evaluation. The SEC also simplified the auditor opinion required under SOX 404. Finally, the SEC adopted a definition of "material weakness" and proposed a definition of "significant deficiency," in each case intending to provide greater clarity on those terms.[3]

The Guidance and related rulemaking do not require companies that already comply with SOX 404 to change processes and procedures that are currently in place and with which management, audit committees and auditors are comfortable. The Guidance may, however, provide such companies with greater flexibility to streamline their evaluation processes going forward and to reduce the costs of SOX 404 compliance. In addition, the Guidance should be helpful to the significant number of "non-accelerated filers"—generally, those companies with a market capitalization below $75 million—which have not yet been required to comply with SOX 404 and are preparing to come into compliance later this year, as well as those companies that are not yet reporting companies but may become reporting companies in the future.[4]

The Guidance — Principles

The Guidance is organized around two broad principles. The first is that management should evaluate whether it has implemented controls that adequately address the risks of failing to prevent, or not detecting in a timely manner, a material misstatement in the company's financial statements. The Guidance sets forth a top-down, risk-based approach to this principle. This approach contrasts with the practice of many companies that have attempted to identify and evaluate every possible control, regardless of the risk that a failure of that control would result in a material misstatement. And second, that management's evaluation of its controls should be based on its assessment of risk. This principle attempts to address concerns about rigid approaches to the level of documentation needed to substantiate management's evaluation of ICFR.

The Guidance — The Evaluation Process

As stated in the existing SEC rules implementing SOX 404, the objective of ICFR is to provide reasonable assurance regarding the reliability of financial reporting and regarding the preparation of financial statements for external purposes in accordance with GAAP. Those rules require that annual reports contain management's assessment of the effectiveness of ICFR; the purpose of the management-evaluation requirement in those rules is to provide management with a reasonable basis for that assessment. The Guidance describes the evaluation process as having two parts: first, management identifies risks to reliable financial reporting and the controls that address those risks (including evaluating whether the controls are designed to adequately address the identified risks); and second, management evaluates the operation of those controls to determine their effectiveness.

As described in the Guidance, the starting point is for management to consider "what could go wrong" with respect to a financial statement amount or related disclosure, in order to identify the sources and potential likelihood of misstatements and the risks that could result in a material misstatement in the financial statements. This risk-identification process should incorporate management's knowledge of the company's business and operations, including the vulnerability of the business to fraudulent activity. The size, complexity, and organizational structure of the company and its processes should be a factor in management's risk evaluation.

The Guidance describes the next step in the process as involving management's judgments about whether the controls in existence, if operating properly, can effectively prevent or detect misstatements that could result in material misstatements in the company's financial statements. The Guidance emphasizes that it is not necessary to identify all controls that may exist, and that, in a situation where multiple controls address the same risk, management may decide to focus its evaluation on the control for which evidence of operating effectiveness can be obtained more efficiently. For example, it may be more efficient to evaluate automated controls rather than manual controls where both address the same risk.

In addressing the process for evaluating evidence of the operating effectiveness of a control, the Guidance indicates that management should consider whether the control is operating as designed and whether the person performing the

** Daniel Fisher is an Associate in the Washington, D.C. office of Skadden, Arps, Slate, Meagher & Flom, LLP.*

control possesses the necessary authority and competence to perform the control effectively. The focus of this evaluation should be on those areas posing the highest risk to ICFR—and, to this end, factors that should be taken into account in determining risk include whether a particular financial statement amount or disclosure involves judgment in determining the recorded amounts, is susceptible to fraud, has complex accounting requirements, experiences change in the nature or volume of the underlying transactions, or is sensitive to changes in environmental factors such as economic developments. The Guidance states that the amount of evidence needed to support management's assessment of the effectiveness of ICFR—which assessment considers both the quantity and quality of the evidence—is a function of the materiality of the financial statement amount or disclosure in question and the susceptibility to material misstatement of the underlying account balances, transactions, or supporting information to material misstatement. The Guidance provides further clarification by stating that evaluation methods may be integrated with daily responsibilities or may be performed for other management reasons, and need not be limited to procedures implemented specifically for the ICFR evaluation.

The Guidance — Reporting Considerations

Under the SEC rules implementing SOX 404, management may not conclude that ICFR is effective if there are one or more "material weaknesses" as of the fiscal year end, and any such material weakness must be disclosed as part of management's report. The Guidance makes clear that, as part of management's evaluation, management must consider whether each deficiency, alone or in combination with other deficiencies, constitutes a material weakness, and that this evaluation involves both quantitative and qualitative factors. The Guidance sets forth a number of factors that could affect whether a deficiency, or a combination of deficiencies, will result in a misstatement of a financial-statement amount or disclosure. Among these factors are the nature of the financial reporting amounts or disclosures involved, the susceptibility of the asset or liability to loss or fraud, the complexity of any required judgment, interaction with other controls or deficiencies, and the possible future consequences of the deficiency. The Guidance also makes clear that the effect of compensating controls can be taken into account in determining whether a deficiency is a material weakness.

Current auditing standards list certain situations as "strong indicators" of a material weakness. As a practical matter, in most instances the presence of these factors has been viewed as mandating the conclusion that a material weakness exists. The Guidance departs from this position by emphasizing that whether a deficiency constitutes a material weakness is a matter of judgment based on all the relevant facts and circumstances. As a result, the Guidance calls for management to consider whether the following situations indicate a deficiency in ICFR and, if so, whether they represent a material weakness: fraud, whether or not material, by senior management; a restatement of financial statements to correct a material misstatement; detection of a material misstatement in current-period financial statements in circumstances that indicate the misstatement would not have been detected by ICFR; and ineffective oversight by the audit committee.

Consistent with prior, informal guidance by the SEC staff, the Guidance states that if a material weakness exists, companies should consider disclosing the nature of such weakness, its impact on the company's financial reporting and ICFR, and management's plans or ongoing actions to remedy the material weakness.

The Guidance also formalizes the SEC staff guidance that in the event of a restatement of previously issued financial statements management should consider whether the original disclosure—as to the effectiveness of ICFR and the effectiveness of disclosure controls and procedures for the period that is the subject of the restatement—is still accurate, and should amend those disclosures as necessary so as not to be misleading in light of the restatement.

Amendments to SEC Rules Implementing SOX 404

In connection with the Guidance, the SEC amended its rules implementing SOX 404 to make clear that a management evaluation of ICFR conducted in accordance with the Guidance will satisfy management's obligation under the rules. The amended rules also make clear that the Guidance is only one of many ways to conduct an evaluation of ICFR. Accordingly, there is no requirement for companies that are already SOX 404-compliant to alter their procedures to fit within the Guidance. In addition, the SEC amended its rules concerning the auditor's SOX 404 attestation. Instead of opining on whether management's assessment of the effectiveness of ICFR is fairly stated in all material respects, as well as whether ICFR is effective, auditors will opine only as to whether the company maintained effective ICFR.

Definitions of "Material Weakness" and "Significant Deficiency"

When first implementing SOX 404, the SEC referred to the definitions of "material weakness" and "significant deficiency" in the accounting literature in existence at that time and, later, as modified in PCAOB auditing standards. In connection with the Guidance, the SEC codified in its rules the definition of material weakness as "a deficiency, or a combination of deficiencies, in internal control over financial reporting such that there is a reasonable possibility that a material misstatement of the registrant's annual or interim financial statements will not be prevented or detected on a timely basis."[5]

In connection with its consideration of the Guidance and related rule amendments, the SEC has also proposed to codify in its rules a definition of "significant deficiency." Unlike a material weakness, a significant deficiency does not render ICFR ineffective and, accordingly, identifying significant deficiencies is not part of the purpose of management's evaluation or of the auditor's attestation regarding ICFR. Significant deficiencies are relevant in that the CEO and CFO certifications required by Sarbanes-Oxley must indicate that the CEO and CFO have disclosed any significant deficiencies in ICFR (as well as any material weaknesses) to the auditor and to the audit committee. Consistent with the underlying theme

of communication among management, audit committees and auditors, the SEC has proposed to define a significant deficiency as "a deficiency, or a combination of deficiencies, in internal control over financial reporting that is less severe than a material weakness, yet important enough to merit attention by those responsible for oversight of a registrant's financial reporting."[6]

Endnotes

1 The Guidance is *available at* http://www.sec.gov/rules/interp/2007/33-8810.pdf.

2 Following approval by the SEC, PCAOB Auditing Standard No. 5, An Audit of Internal Control Over Financial Reporting That Is Integrated With an Audit of Financial Statements, *available at* http://www.pcaobus.org/Rules/Docket_021/2007-06-12_Release_No_2007-005A.pdf, would replace the current Auditing Standard No. 2.

3 The adopting release amending the rules and defining "material weakness" is *available at* http://wwww.sec.gov/rules/final/2007/33-8809.pdf. The release proposing a definition of "significant deficiency" is *available at* http://www.sec.gov/rules/proposed/2007/33-8811.pdf.

4 Non-accelerated filers must include management's report on ICFR in their annual reports for fiscal years ending on or after December 15, 2007, although the related auditor attestation is not required until annual reports for fiscal years ending on or after December 15, 2008. Newly public companies may avail themselves of a transition period allowing them to omit management's report on ICFR and the related auditor attestation from their first annual report.

5 *See supr*a note 3.

6 *Id*.

Free Speech and Election Law

Violence ≠ Obscenity

*By Geoffrey W. Hymans**

On September 17, 2007, the United States District Court for the Western District of Oklahoma became the latest federal court to strike down a state law regulating the distribution of "violent" video games.[1] This followed hard on the heels of a similar decision from the Northern District of California.[2] In both cases the state crafted regulations that tracked the legal test for obscenity set forth in *Miller v. California*,[3] yet neither court accepted the argument that depictions of violence could constitute a separate category of unprotected "obscene" speech akin to currently unprotected sexual obscenity. An examination of the history of the "violence as obscenity" argument provides an opportunity to examine how and why sexual expression really is uniquely treated under the First Amendment.[4]

The recent United States District Court opinions are not the first to examine the regulation of violent video games. Two federal Circuit Courts have struck down local regulatory schemes. The Eighth Circuit stuck down a St. Louis County ordinance barring the sale, rental, or provision of "graphically violent" video games to minors.[5] In that case, the court dismissed comparisons of violence to sexual obscenity with a curt: "Simply put, depictions of violence cannot fall within the legal definition of obscenity for either minors or adults."[6]

Judge Richard Posner authored a Seventh Circuit opinion that engaged in significantly more analysis of the different regulatory rationales for censoring sexual obscenity and violence, but reached the same result.[7] In typical Posnerian style, the judge closely examined the distinct rationales for banning sexual obscenity and violence, and determined that insulating children from the "cartoon-like" violence presented in the video games at issue was not a compelling interest.

Five other district courts have invalidated direct state regulation of violent video games.[8] Direct regulation of violent movies and videos have fared no better,[9] and courts have also rejected secondary regulation through the tort liability mechanism.[10] Thus, those members of the Federalist Society with a libertarian bent might cheer the consistent protection the courts have provided when faced with the censorship of violent video games. Yet the consistent rejection of video game regulation using the same basis—and often the same language—as statutes regulating sexual obscenity encourages us to take a step up the abstraction ladder and inquire why violence is treated differently than sex when it comes to classification as obscenity?

The "Violence as Obscenity" approach can largely be traced back to a single source: Professor Kevin W. Saunders' book by that title published in 1996.[11] Although there appear to have been prior scattered legislative attempts to restrict access to violent media materials, particularly with respect to minors,[12] Professor Saunders provided the playbook when the development of interactive videogames and highly publicized school shootings coincided to spur state and local regulatory zeal.[13] The introduction to his book was straightforward:

> This work accepts the existence of the obscenity exception, but it will be argued that the exception is misfocused, or at least too finely focused, on depictions of sex and excretory activities. Violence is at least as obscene as sex. If sexual images may go sufficiently beyond community standards for candor and offensiveness, and hence be unprotected, there is no reason why the same should not be true of violence.[14]

Saunders work thoroughly tracked the development of obscenity law, including the sliding-scale obscenity standards which the Supreme Court applied to minors.[15] He cited studies from that time arguing that exposure to violent media causes aggressive behavior.[16] The book actually included a chapter specifying how to draft a statute regulating violent content using the *Miller* test adapted from sexual obscenity.[17]

That playbook was followed by the legislatures of Oklahoma, California, Louisiana, Michigan, and Illinois, and by local legislative bodies in the City of Indianapolis and St. Louis County.[18] All utilized the *Miller* standards adapted to violent content. Often the *Miller* standards were incorporated into the definition of a "violent video game," the distribution of which was then limited by the statute or ordinance. The California statute[19] is typical:

> (d) (1) "Violent video game" means a video game in which the range of options available to a player includes killing, maiming, dismembering, or sexually assaulting an image of a human being, if those acts are depicted in the game in a manner that does either of the following:
>
> (A) Comes within all of the following descriptions:
>
> (i) A reasonable person, considering the game as a whole, would find appeals to a deviant or morbid interest of minors.
>
> (ii) It is patently offensive to prevailing standards in the community as to what is suitable for minors.
>
> (iii) It causes the game, as a whole, to lack serious literary, artistic, political, or scientific value for minors.

Thus far, no court has adopted Professor Saunders view that portrayals of violent acts can be obscene in the same manner as portrayals of sexual acts. Obviously legislatures have differed in their conclusions. Commentators have even latched on to sexual descriptors, referring to very violent movies as "torture porn."[20] And scholars continue to debate in the pages of law journals.[21]

But since the courts seem set on rejecting the "violence as obscenity" trope, might we not inquire about the rationales for the distinction? Judge Posner came closest to enunciating such a rationale, asserting the violent content was being regulated for its potential effect upon the viewer, while obscene sexual

* *Geoffrey W. Hymans is a Washington State Assistant Attorney General.*

depictions were banned merely for their offensiveness.[22] But while that may have been true for the Indianapolis ordinance under review in *Kendrick*, that is not the focus of Professor Saunders' work, and presumably not the focus of the legislatures which used his road map. Saunders wants to regulate violence both for the effects it allegedly has on its viewers and because it is offensive.

In truth, as Posner seems to acknowledge, it is very hard to separate these rationales. Perhaps violent media is offensive to some both because they believe it causes violent behavior in viewers and because it makes their stomachs churn. Certainly, the effect on the viewer includes both any increased propensity to violence and any psychic—or gastrointestinal—impacts.

Is the redefinition of violence as obscenity any different than the attempted redefinition of pornographic media as actual acts of violence against women? Those arguments, made most prominently by Catherine MacKinnon and Andrea Dworkin in the 1980s,[23] were also rejected by courts.[24] In that instance, the inability to expand the *Miller* definition of obscenity to include "soft-core" pornography led to the attempt to change the "rules of the game" by shifting the terrain from message to action.

Although the courts have thus far rejected the social science studies cited to them regarding the correlation between violent media and aggressive behavior,[25] the question remains: why are the courts so much more vigilant when it comes to the effects of violence as compared to the effects of sex? Compare the treatment of the effects of violent video games with the treatment of "secondary effects" with regard to exotic dancing. One might wonder why the postulated "effects" (since courts have even eliminated the requirement to demonstrate such effects with regard to the actual establishment being regulated) of nude dancing on a surrounding neighborhood form the basis for significant regulation of that dancing (including regulations that experts have testified affect the "message"—such as distance regulations[26]) while similar correlations of game playing violence with "desensitivity" to actual violence are insufficient basis for regulation.[27] And why is the legislative body due deference in the fact-finding role of deciding what are the secondary effects of nude dancing while the legislature is not due any deference in determining the degree of correlation (or even causation, if that is a factual issue) with regard to violent video games?[28]

If the courts want to close the door on these repeated "redefinition" arguments to eliminate First Amendment protection for discrete categories of speech, the only effective way to halt such attacks would be for the courts to eliminate the comparison class. This is not because of actual similarity between materials deemed sexually obscene and other categories of expression regulators seek to restrict. Rather, it is because of the way lawyers think.

Lawyers reason and argue by analogy.[29] And lawyers representing those trying to regulate violent video games will reach for the nearest analogy. Categories of wholly unprotected speech related to violence include incitement[30] and "fighting words,"[31] both of which require immediacy of effect unlikely to ever be demonstrated in sufficient empirical certainty to support regulation of video games. But the obscenity category of wholly unprotected speech, because of its amorphous and inherently undefined character, presents the most attractive analogous target.

With seven decisions rejecting regulation of "violent" video games, only a quantum causational leap in the studies underlying the regulatory efforts, or an adverse Supreme Court decision, appears likely to permit future regulation. However, add a dose of sex and we will see if the courts take a different approach. [32]

Endnotes

1 Entm't Merchants Ass'n. v. Henry, 2007 WL 2743097 (W.D. Okla.).

2 Video Software Dealers Ass'n. v. Schwarzenegger, 2007 WL 2261546 (N.D. Cal.)

3 413 US 15 (1973). The *Miller* test is:

(a) whether the average person, applying contemporary community standards' would find that the work, taken as a whole, appeals to the prurient interest;

(b) whether the work depicts or describes, in a patently offensive way, sexual conduct specifically defined by the applicable state law; and

(c) whether the work, taken as a whole, lacks serious literary, artistic, political, or scientific value.

Miller v. California, 413 U.S. at 24.

4 Given time and space limitations, the scope of this article is necessarily narrow, addressing only the federal case law on violence as obscenity.

5 Interactive Digital Software Association v. St. Louis County, 329 F.3d 954 (8th Cir. 2003).

6 *Id.* at 958.

7 Am. Amusement Machine Ass'n. v. Kendrick, 244 F.3d 572 (7th Cir. 2001).

8 Entertainment Software Ass'n. v. Foti, 451 F. Supp. 2d 823 (M.D. La. 2006); Entm't Software Ass'n. v. Hatch, 442 F. Supp. 2d 1065 (D. Minn. 2006); Entm't Software Ass'n. v. Granholm, 426 F. Supp. 2d 646 (E.D. Mich. 2006); Entm't Software Ass'n. v. Blagojevich, 404 F. Supp. 2d 1051 (N.D. Ill. 2005); Video Software Dealer's Ass'n. v. Maleng, 325 F. Supp. 2d 1180 (W.D. Wash. 2004).

9 Video Software Dealers Ass'n. v. Webster, 968 F.2d 684 (8th Cir. 1992) (affirming Video Software Dealers Ass'n. v. Webster, 773 F. Supp. 1275 (W.D. Mo. 1991)); Davis-Kidd Booksellers v. McWherter, 866 S.W. 2d 520 (1993).

10 James v. Meow Media, Inc., 300 F.3d 683 (6th Cir.2002); Wilson v. Midway Games, Inc., 198 F.Supp.2d 167 (D.Conn.2002); Sanders v. Acclaim Entm't, Inc., 188 F.Supp.2d 1264 (D.Colo.2002); Byers v. Edmondson, 826 So.2d. 551 (2002).

11 Kevin W. Saunders, Violence as Obscenity: Limiting the Media's First Amendment Protection (1996) (*hereinafter* "Saunders").

12 *Supra,* note 10.

13 Highly publicized school shootings included the Columbine massacre in 1999. Perpetrators Eric Harris and Dylan Klebold were reportedly fans and players of "first-person shooter" video games such as Doom. *Available at* http://www.time.com/time/magazine/article/0,9171,1101991220-35870,00.html. The sponsor of the Washington State statute cited the Columbine shootings and violent video games such as Grand Theft Auto (in which characters can shoot police officers) as the motivation for her introduction of the legislation. *Available at* http://www.pbs.org/newshour/extra/features/july-dec03/video_7-16.html. And the St. Louis ordinance struck down by the Eighth Circuit specifically mentioned several school shootings in the legislative findings. *See* Interactive Digital Software Ass'n v. St. Louis County, Mo., 200 F.Supp.2d 1126 (E.D.Mo. 2002).

14 Saunders at 3.

15 Ginsburg v. New York, 390 U.S. 629 (1968).

16 *See, e.g.,* Saunders, ch. 2.

17 *Id.*, ch. 9.

18 Minnesota and Washington adopted more targeted statutes. Minnesota enacted a ban on rental or purchase of video games with an Entertainment Software Rating Board rating of AO (Adults Only) or M (Mature) by those under seventeen years of age. Minn. Stat. §325I.06. Washington enacted civil penalties for selling or renting to minors video games that contain "realistic or photographic-like depictions of aggressive conflict in which the player kills, injures, or otherwise causes physical harm to a human form in the game depicted, by dress or other recognizable symbols, as a public law enforcement officer." RCW 9.91.180. A bit of a disclaimer—the Washington statute was passed during the time in which I served as Senior Counsel for the Republican Caucus of the Washington State House of Representatives, however I did not work on the bill or the issue.

19 Cal. Civ.Code §§ 1746(d), cited in *Schwarzenegger*, *supra* note 2.

20 Commentator Julie Hilden has written presciently on these issues. *See* Julie Hilden, *The Attacks on "Violent" Video Games" and "Torture Porn" Films: Two Different Strategies to Try to Get Around First Amendment Protections*, Sept. 17, 2007, *available at* http://writ.news.findlaw.com/hilden/20070917.html; Julie Hilden, *Free Speech and the Concept of "Torture Porn": Why are Critics So Hostile to "Hostel II"?*, July 16, 2007, *available at* http://writ.news.findlaw.com/hilden/20070716.html . *See also, e.g.,* David Edelstein, *Now Playing at Your Local Multiplex: Torture Porn*; *available at* http://nymag.com/movies/features/15622/.

21 A simple Westlaw search for "video game" in the title of journals and law reviews produces thirty-eight citations.

22 *Kendrick*, 244 F.3d at 574-76.

23 *See, e.g.* Catherine McKinnon, Only Words (1993).

24 *See* American Booksellers Ass'n, Inc. v. Hudnut, 771 F.2d 323 (7th Cir. 1985).

25 *See, e.g.*, *Schwarzenegger*, *supra* note 2, at 6, 11 (discussing California legislature's reliance on studies by Dr. Craig Anderson, a Ph.D. psychologist).

26 Experts have testified, often without opposing testimony, that distance affects the *content* of exotic dance. *See, e.g.,* Colacurcio v. City of Kent, 163 F.3d 545, 559 (9th Cir. 1998) (Reinhardt, J., dissenting; discussing the expert opinions of Dr. Edward Donnerstein and Dr. Judith Hanna).

27 Compare *Interactive Digital Software Ass'n. v. St. Louis County*, *supra*, in which the County legislative body's conclusion – based on the testimony of an expert who appeared before them – that violent video games frequently result in more aggressive behavior, with the same federal circuits typical treatment of the evidentiary requirements for restricting nude dancing. SOB, Inc. v. County of Benton, 317 F.3d 856 (8th Cir. 2003)(holding that local government may rely on studies from other local governments to justify banning nude dancing). In many ways video games would appear to be rather close analogs of dance, since both are expressive activities rather than "pure" expression.

28 Perhaps the secondary effects test, like the obscenity doctrine itself, is merely a reflection of the "sex is different" approach of the Supreme Court. While we have started to see arguments regarding "secondary effects" made beyond the nude dancing/adult entertainment context, perhaps Justice Brennan's concurrence/dissent from *Boos v. Barry*, 485 U.S. 312, 334-338 (1988), Justice Souter's concurrence in *Barnes v. Glen Theatre, Inc.*, 501 US 560, 585-586 (1991), and footnote 7 from the majority opinion in *RAV v. St. Paul*, 505 U.S. 377 (1992), serve to limit the spread of this unique doctrine.

29 *See, e.g.,* Cass R. Sunstein, *On Analogical Reasoning,* 106 Harv. L. Rev. 741, 741 (1993); Scott Brewer, *Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy,* 109 Harv. L. Rev. 925, 926 (1996) (describing how the common law promotes reasoning by analogy to tie changing facts and circumstances to earlier judgments); Richard Posner, *Reasoning By Analogy*, 91 Cornell L. Rev. 761 (2006) (reviewing Lloyd L. Weinreb, *Legal Reason: The Use of Analogy in Legal Argument*, 2005).

30 Courts analyzing violent video games have noted that they don't spur imminent violence. *See*, *e.g., Schwarzenegger*, *supra* note 2 at 4, discussing Brandenburg v. Ohio, 395 U.S. 444 (1969)(*per curiam*).

31 *See, e.g.*, *Kendrick*, *supra* at 575*,* discussing Chaplinsky v. New Hampshire, 315 U.S. 568 (1942).

32 The "Grand Theft Auto: San Andreas" video game included code which, when unlocked with a software modification, allowed the player to view explicit sex scenes. Steve Lohr, *In Video Game, a Download Unlocks Hidden Sex Scenes*, N.Y. Times, July 11, 2005, *available at* http://www.nytimes.com/2005/07/11/technology/11game.html. This spurred the interest of at least one potential regulator. Raymond Hernandez, *Clinton Urges Inquiry Into Hidden Sex in Grand Theft Auto Game*, N.Y. Times, July 14, 2005, *available at* http://www.nytimes.com/2005/07/14/nyregion/14hillary.html.

Intellectual Property

How Patents Can Help Those Interested in the Environment And World Health

*By F. Scott Kieff**

It has become fashionable to argue that patents frustrate important social goals like protecting the environment and world health. This essay addresses some representative concerns relating to patents that appear to be prevalent in the environmental literature[1] and in public debates about world health, and shows how a strong patent system can be better seen as an important part of the solution, not part of the problem.

The impact of the U.S. patent system can be best understood by first exploring the system's central goals and effects, as well as its general context. The central goal of the U.S. patent system is to provide an economic tool for promoting public access to new technologies.[2] While the central effect of the system has been the achievement of this goal, the impact of such increased access is not an unmitigated good. For example, while some technologies when put to some uses may help causes like the environment and health, others may hurt. This is where an understanding of context becomes important, because the patent system does not operate in a legal vacuum.

The potential for harmful impact is well recognized and addressed by diverse parts of most national legal systems that regulate and in some cases prohibit the use of certain technologies, whether they happen to be patented or unpatented. Consider, for example, the extensive regulations administered by the U.S. Environmental Protection Agency (EPA) on the use of chemicals,[3] those of the Food and Drug Administration (FDA) on the use of drugs,[4] and those under many state laws on the use of firearms.[5] To the extent that environmental and health interests are in favor of such restrictions on use, the interests need not be troubled by the patent system because the system gives the patentee only an additional right to exclude use of whatever is covered by the patent claim.[6] Patents do not give patentees any right to use. Therefore, the patent system has no effect on other restrictions on use, whether the restrictions come from the environmental or health arenas, or elsewhere.

** F. Scott Kieff is a Professor at the Washington University School of Law, and a Research Fellow at Stanford University's Hoover Institution.*
He gratefully acknowledges contributions from the officers of the National Association of Environmental Law Societies Conference Planning Committee, the Washington University Environmental Law Society, and the Washington University Journal of Law and Policy for framing the issues addressed in this essay, as well as the particular insights provided by Christine Cochran, Alexis Gorton, Ron Kaufmann, Stephen Nickelsburg, Troy Paredes, and Susan Vitale. He especially appreciates the able research assistance, critical evaluation, and editing of Jessica E. Merrigan. This essay was first prepared for the 2002 National Association of Environmental Law Societies' Conference: "Sustainable Agriculture: Food for the Future," and originally published as F. Scott Kieff, Patents for Environmentalists, 9 Wash. U.J.L. & Pol'y 307 (2002). He gratefully acknowledges the initial publisher's permission to re-publish this updated version.

In contradistinction, to the extent that environmental, health, and other interests are in favor of increased use, the patent system can provide great benefit. A central concern about patents that is expressed both generally and specifically in the environmental literature, for example, is the fear that the patent right to exclude use will cause patented technologies to be underused.[7] But the patent literature teaches that the right to exclude use that is the core of the patent system's enforcement rules actually operates to increase use by facilitating ex ante investment in the complex, costly, and risky commercialization activities required to turn nascent inventions into new goods and services.[8] This right to exclude competitors who have not shared in bearing the initial costs of commercialization provides incentives for the holder of the invention and the other players in the market to come together in an organized way and incur the costs necessary to facilitate commercialization of the patented invention.[9] The drafters of our present patent system, the 1952 Patent Act, had precisely this concern for commercialization in mind when drafting the statute, and were motivated by the specific fear that, for example, the handicapped in need of a new wheelchair might not find one to buy if the patent system did not provide an incentive for it to be brought to market in the first instance.[10]

The patent system evolved a set of patentability rules such that the system can generate this increase in use while at the same time minimizing social costs, including those typically associated with information, administration, public choice, races for a common prize, and bargaining.[11] For example, patent law's requirements regarding the prior art—the § 102[12] and § 103[13] requirements that an invention be novel and non-obvious—operate to protect investments, including those by someone other than the patentee.[14] In addition, the § 112[15] disclosure requirements decrease social costs by giving clear notice about the property right, which both decreases the chance of inadvertent infringement and of duplicative efforts towards the same invention.[16]

The complex interactions in the patent system between the rules for enforcing and obtaining patents operate dynamically through the crux of the patent, the claim, to ensure that patents have a scope that is "just right."[17] As Judge Rich often said about patents, "*the name of the game is the claim*... [and] the function of claims is to enable everyone to know, without going through a lawsuit, what infringes the patent and what does not."[18] According to Judge Rich, claims present a fundamental dilemma for every patentee because "the stronger a patent the weaker it is and the weaker a patent the stronger it is."[19] By this dilemma, he meant that a broad patent claim is strong on offense because it covers more and therefore is more likely to be infringed, but it also is weak on defense because it may cover something in the prior art or fail to contain a sufficiently detailed disclosure, and therefore

is more likely to be invalid; while a narrow claim is weak on offense, because it covers less and therefore is less likely to be infringed, but it also is strong on defense because it is less likely to cover something in the prior art or fail to contain a sufficiently detailed disclosure, and therefore also is less likely to be invalid.[20]

Patents vetted through such a self-disciplining regime can form the basis of licensing transactions with others seeking permission from the patentee to practice whatever is claimed in the patent. These transactions allow those seeking use to obtain permission for use. For example, a patented technology that has beneficial environmental or health impact can be licensed to all those who wish to achieve that impact.

Importantly, the patent system has developed a set of rules about licensing that operate ex post to maximize the likelihood that all those wanting such use will get it. Putative licensees who place a high value on such use and those who place a low value on such use are both attractive targets to a patentee as long as the patentee is allowed to set a different price for different users. This practice is called price discrimination. Patent law allows patentees to price discriminate among such licensees because this gives patentees a strong financial incentive to ensure all those desiring use get use; even a monopolist who can price discriminate will push output to the full competitive output level.[21] Such beneficial price discrimination can take place because patent law, and contract law, allow for the enforcement of the restrictive licenses needed to prevent arbitrage between low value and high value users.[22] In the presence of such a system, a patentee is rationally motivated to avoid posting an excessive price because to do so would scare away would-be paying customers and this result would be a money-losing venture.

Even where the user is not able to pay any positive price, the patentee may be rationally motivated to grant a license for free. The granting of a free license may provide the patentee with an inexpensive way to preserve the legal force of the patent property right for use in other transactions with paying customers.[23] The patentee may also be able to derive advertising benefits from such uses as long as they are successful uses and their low price does not cause customer-relations harm with the high-paying customer base.[24] Thus, even very low value users are likely to be able to obtain licenses from the patentee.

Some argue that while patentees may be rationally motivated to sell permission to each user, and users may be rationally motivated to buy permission from patentees, such sales may not be consummated because of various market failures.[25] In response to these concerns, some commentators argue that patents should be protected by a liability rule[26] instead of a property rule.[27]

Indeed, there are already important liability rule provisions in patent law today. Otherwise infringing uses that are by or for the federal government enjoy sovereign immunity protection that effectively results in a compulsory licensing regime.[28] In addition, the high costs of litigation under the present rules of civil procedure and the ability for an infringer to be kept effectively judgment proof through corporate and bankruptcy laws may also operate as a form of liability rule gloss on the present property rule regime.[29]

Not only is the market power of the patent not as strong as it may seem,[30] it may have the beneficial effect of inducing even more new technologies. To the extent that some would-be licensees may not be able to obtain permission for use despite manifesting some willingness to pay some positive price,[31] the presence of such potential customers and the potential for an independent patent each provide incentives for others to bring to market some alternative non-infringing substitute.

Moreover, the political process provides several solutions for would-be licensees. They may prevail on the government simply to provide such use in particular cases.[32] They may alternatively prevail on the government to subsidize their ability to pay.[33]

Ensuring an environmental or health use through a switch in the patent system towards over-all liability rule treatment should be avoided because these other remedies are available and because such a shift will frustrate the important goals of the patent system, including those that are specifically pro-environment, such as the commercialization of beneficial technologies. The use of liability rules would lead to a net increase in social cost and frustrate the very efforts for ordering and bargaining around patents that are necessary to generate output of patented inventions in the first instance, thereby decreasing over-all social access to new technologies.[34] As recognized by Robert Merges, it is precisely because private parties have a comparative advantage over courts in valuing patents and patented inventions that a property rule is likely to work better than a liability rule according to the established test for choosing between the two types of regimes.[35]

The ability to exclude use through a patent not only facilitates increased use, it also provides individual actors with a legal alternative to self-help approaches that may have more pernicious impact on the ability to obtain use.[36] Consider, for example, the concern expressed in the environmental literature about a form of self-help in the agricultural sector called "terminator technologies" and the fear that these might cause environmentally important plant species to die out.[37] Terminator technology refers to seeds that were genetically altered so as to yield crops whose resulting seed will be sterile.[38] The technology prevents farmers from harvesting seeds from crops they have grown using genetically engineered seeds, thereby forcing farmers to buy more of the original seed each planting season.[39] Terminator technology can also be thought of as the agricultural equivalent of copy protection technology in the software industry.

Such terminator and copy protection technologies are each a form of self-help that can be used as an alternative to legal protection in a way that is likely to be more costly than legal protection. Consider a market for some modified form of seed that was altered so as to make it especially valuable compared to other seeds. Since seeds generate plants that in turn produce more seeds, the sale of a seed must take into account the potential of vast progeny seeds that are themselves potent for germination. The seller must consider the risk that the buyer will generate maximal progeny, maybe even returning to the market to sell some progeny seeds in competition with the original seller. The price needed to cover for this risk will

far exceed the price needed to cover a sale to a farmer who will only use the seed for production of a single crop and who will not generate progeny seed. Buyers seeking seed for the purpose of growing such a single crop will want to identify themselves convincingly to sellers. Sellers' willingness to sell to such buyers at the lower price will decrease to the extent the seller disbelieves that the buyer indeed intends to and will use the seed for a single crop. As a result, both pricing and consummation of that sale are frustrated. In contradistinction, terminator technology ensures that both sides of the sale will keep to its terms. Because both seller and buyer know the seed will only be of value for a single crop, pricing and consummation of that sale are facilitated.

But technological self-help is not needed if a legal device will have the same effect, especially if the legal device will be cheaper. One legal device may be a contract for sale having a restrictive term, such as a clause agreeing that the seed will only be used for a single crop. A problem with such a contract may be that it will have enforcement problems. The ordinary contract remedy of expectation damages is likely to under deter breach.[40] In addition, contract remedies will have difficulty reaching any third-party transferees of progeny seeds. Patent law offers a convenient aid because patents can be licensed with restrictive terms and patent remedies include the right to an injunction against any infringer, including both third parties and those in contract privity with the patentee. For this reason, courts uphold patent licenses that restrict buyers to a single use.[41] Indeed, restrictive patent licenses have the added advantage of avoiding the potential risk of some harmful biological consequences that are feared to be associated with self-help devices like the terminator technology, such as the potential for its accidental spread to other plants for which germination is otherwise desired.[42] Therefore, patents, especially when used with enforceable restrictive patent licenses, may be important tools for avoiding environmental concerns with terminator technologies.

Indeed, the patent system can also offer some help to those who are concerned about the need to ensure resources for custodians of biodiversity.[43] While developing nations are often the custodians of biodiversity, they are often excluded from sharing in the benefits of the patents that derive from such biodiversity.[44] But the enforcement of property rights should lead to an arrangement in which those benefits are shared with the custodians.[45] For example, intellectual property rights in the United States have been long recognized to be a critical factor in creating national wealth,[46] and this pool of financial wealth is available in at least several senses for use in helping the biodiversity custodians. Those having the pool of financial wealth may elect to share it through general international subsidies. They may also be encouraged to exchange some of that financial wealth for some continued access to the pool of biodiversity wealth. To the extent that those granting access to the biodiversity wealth have not had a fair shot when forging such deals, efforts to ensure legal representation during contract negotiations between indigenous cultures and bio-prospectors might provide one solution.[47] But it is important to realize that regardless of which method is used to allocate the wealth created by the patent system, a robust protection for patents must be maintained or the wealth itself will be sacrificed.

Not only are patents part of the solution to important health and environmental problems, calls to abrogate patents divert attention from many of the more serious problems. The United Nations' World Health Assembly and World Intellectual Property Organization say they will improve access to health care in the developing world by attacking the so-called "problem" of patents on crucial medicines like anti-malarials and anti-retrovirals.[48] They would have us believe that patents on such essential medicines are little more than vehicles for driving prices artificially high.

But patents do not cause drugs to be expensive as much as the high costs of research, development, regulatory approval, and distribution do. However, even these commercialization-related costs are not the real barrier for getting drugs delivered to patients in poverty-stricken regions like those in sub-Saharan Africa. Ridiculous taxes, import duties, and regulatory barriers are one set of important problems that must, and can, be eliminated immediately. In many cases, taxes and import duties reach well above 50%.[49]

Frequently, after drugs have been found safe and effective by careful regulatory review in the United States, Europe, and Japan, and used widely by citizens there, an additional regulatory review period involving thirty-odd months and high fees are imposed before importation is allowed to poverty-stricken regions. This is all in the name of the "public interest." Even drugs provided for free are often hit with tremendously high import duties; plus they face the same regulatory costs as expensive drugs.

All of these costs are borne by the poverty-stricken populations. And while some of these costs actually may be serving as important subsidies to local governments and economies, those can, and already are, being more efficiently provided directly by the U.S. and others. Of the new drugs that are essential and patented, many already are being provided to these regions at ultra-low cost or for free. They are getting there, but not to the patients. The problem is distribution, not price and not patents. Just imagine, if your newspaper delivery person were given free papers but not paid per delivery, and then ran the risk any profits they did make could simply be taken away. Would you really be surprised if the delivery service went out of business; or was never even started? Local patent enforcement would facilitate the business model for delivery by domestic operations in the first place and improving the general rule of law would protect livelihoods from expropriation. And what if instead of newspapers they were high-value drugs? Why would they not end up on the black markets in the U.S., Europe and Japan? Enforcing patents internationally helps block such black markets. The potential for such black markets discourages both initial supply and local distribution.

Strengthening the rule of law will empower local populations at the grassroots level to overcome these obstacles. Protecting patents and enforcing contracts serve as essential enabling devices for development of the businesses needed to get distribution done. And patents also help in an even more direct way, because the one natural resource that is uniformly distributed wherever there are people is the intellectual capital

stored in every person's mind. It is a tragic straw man to point out that inventors in these regions are ill-equipped to compete in solving the high technology and biotechnology problems being worked by top teams in wealthy U.S., European, and Japanese laboratories. Inventors in the developing world, like inventors everywhere, are smart people working to solve the problems they face; necessity is the mother of invention. These inventors can and do solve a range of practical problems in the developing world not solved by the top research teams in the most developed world precisely because of more keen contact and need.

Access to basic health care is a more critical problem that screams for a solution. When doctors, nurses, hospitals, and long-established, essential treatments and basic medical care are not even available, it surely is an imprudent triage to focus such attention on cutting-edge innovative medicine. The vast majority of the drugs on the World Health Organization's essential medicines list are not even covered by patents today. A host of important diseases including those most often discussed in addition to many others are well treatable by this large array of important, but unpatented, medicines. Indeed, the many wonderful private, foundation, nongovernmental organization, and public efforts already being made to improve access to basic health care in these regions are a proud testament to the effectiveness of this approach. Additional efforts in this direction can have high impact, quickly.

Real lives are being lost. Real action is needed. Patents are an important part of the solution. But it is even more important in the short run to continue to improve access to basic health care and to remove the barriers to distribution for drugs, whether patented or not.

We also should continue to help strengthen the rule of law over time. Botswana has long stood as an impressive example of efforts on this front, and new initiatives along these lines by President Bingu wa Mutharika of Malawi offer similar hope. The developed nations of the world interested in improving access to health care should continue to spend financial and political capital to bolster these efforts.

In conclusion, a well functioning patent system can help on matters of the environment and health because it increases public access to new technologies, decreases use of dangerous self-help approaches, and increases the wealth available for all purposes. To the extent new technologies are helpful to environmental or health goals, such as cleaner burning engines and better drugs, the patent system can be seen as helpful by facilitating their commercialization. To the extent new technologies are harmful to environmental and health goals, such as poisonous chemicals, the patent system can be seen as at least not causing damage because the patent right to exclude use would not interfere with a regulatory system's own effort to exclude use. Those concerned about health and the environment should look to patents as part of the solution, not part of the problem.

Endnotes

1 I am indebted to the NAELS Conference organizers for identification of some representative issues in the environmental literature. Although limited independent research confirmed a similar set of issues, I am confident that a great many other issues exist at this interface and that they will be aptly addressed in due course by others.

2 *See generally* F. Scott Kieff, *Coordination, Property & Intellectual Property: An Unconventional Approach to Anticompetitive Effects & Downstream Access*, 56 Emory L. J. 327 (2006).

3 *See generally* Major Environmental Laws, *available at* http://www.epa.gov/epahome/laws.htm (last visited Sept. 16, 2007) (index of representative statutory foundations for the work of the Food and Drug Administration, including active internet links).

4 *See generally* Compilation of Laws Enforced by the U.S. Food and Drug Administration and Related Statutes, *available at* http://www.fda.gov/opacom/laws/ (last visited Sept. 16, 2007) (index of representative statutory foundations for the work of the Food and Drug Administration, including active internet links).

5 *See generally* Department of the Treasury Bureau of Alcohol, Tobacco and Firearms State Laws and Published Ordinances-Firearms 22nd Edition, *available at* http://www.atf.treas.gov/firearms/statelaws/22edition.htm. (last visited Sept. 16, 2007) (index of representative state laws regulating firearms, including active internet links).

6 U.S. patents give to the patentee only a right to exclude others from using whatever is claimed in the patent. *See* 35 U.S.C. § 154 (a) (1994) ("Every patent shall contain... a grant to the patentee... of the right to exclude others.").

7 As discussed generally *supra* note 1, NAELS Conference meeting organizers suggested that I treat the work of Gollin and Derzko as representative of the environmental community's views on this particular issue. *See*, *e.g.*, Michael A. Gollin, *Using Intellectual Property to Improve Environmental Protection*, 4 Harv. J. L. & Tech. 193 (1991) (exploring the potential intellectual property protection has for promoting innovation in environmental technology); Michael A. Gollin, *Patent Law and the Environment/Technology Paradox*, 20 Envtl. L. Rep. 10171 (1990) (discussing the need to expand environmental law from merely controlling harmful technology to also encouraging beneficial technology and the model that patent law could provide for this); *see also, e.g.*, Natalie Derzko, *Using Intellectual Property Law and Regulatory Processes to Foster the Innovation and Diffusion of Environmental Technologies*, 20 Harv. Envtl. L. Rev. 3, 8 (1996) (exploring the problems for innovation in the current environmental regulatory scheme).

8 Kieff, *supra* note 3, (explaining how the right to exclude use promotes commercialization by facilitating the social ordering and bargaining around inventions that are necessary to get those inventions put to use).

9 *Id.*

10 F. Scott Kieff, *Property Rights and Property Rules for Commercializing Inventions*, 85 Minn. L. Rev. 697, 736-46 (2001) (showing how the drafters of the 1952 Patent Act were motivated by the commercialization theory and specifically contemplated such a wheelchair example).

11 F. Scott Kieff, *The Case for Registering Patents and the Law and Economics of Present Patent-Obtaining Rules*, 45 B.C. L. Rev. 55 (2003) (showing how positive patent law rules, particularly those for obtaining patents, operate to minimize social costs).

12 35 U.S.C. § 102 (1994) (novelty and statutory bars).

13 35 U.S.C. § 103 (1994) (nonobviousness).

14 Kieff, *supra* note 12, at 6. In this sense, the novelty requirement can be viewed as a tool for ensuring that patents do not issue on anything others are already doing and the non-obviousness requirement can be viewed as a tool for ensuring that patents do not issue on anything that others are about to do.

15 35 U.S.C. § 112 ¶¶ 1-2 (1994) (setting forth the disclosure requirements of patent law: (1) written description; (2) enablement; (3) best mode; and (4) definiteness, which is also stated as the requirement that the claims particularly point out and distinctly claim).

16 Kieff, *supra* note 12, at 6.

17 *Id.* at 9-10.

18 *See*, *e.g.*, Giles S. Rich, *The Extent of the Protection and Interpretation of Claims—American Perspectives*, 21 Int'l Rev. Indus. Prop. & Copyright L. 497, 499, 501 (1990) (quoted in Hilton Davis Chem. Co. v. Warner-Jenkinson Co., 62 F.3d 1512, 1539 (Plager, J., joined by Archer, J.; Rich, J; and Laurie, J., dissenting)) (emphasis in original).

19 *See*, *e.g.*, Giles S. Rich, *The Proposed Patent Legislation: Some Comments*, 35 Geo. Wash. L. Rev. 641, 644 (1967) (responding to proposed legislation S. 1042 and H.R. 5924, 90th Cong. (1967) and Report of the President's Commission on the Patent System (1966)).

20 *Id*. (explaining patentee's dilemma, or in his words, "puzzle").

21 Kieff, *supra* note 11, at 727-32 (showing how the patent system's facilitation of tie-ins and other forms of price discrimination where technological and economic factors alone might prevent price discrimination together provide incentives for the patentee to elect to keep output at competitive levels).

22 *Id*. The prevention of arbitrage is essential for price discrimination to work. For example, those obtaining senior citizen discounts could sell their low price tickets to patrons who would otherwise have to pay full price if movie theatres did not require some proof of age on admission, which may be as simple as looking at the ticket holder.

23 F. Scott Kieff, *Facilitating Scientific Research: Intellectual Property Rights and the Norms of Science—A Response to Rai & Eisenberg*, 95 NW. U. L. Rev. 691, 705 (2001) (discussing a property owner's rational decision to allow free users so as to avoid the cost of monitoring low value uses while preserving the full scope of the property right for other high value uses).

24 Giving away product to the poor will force the patentee to wrestle with a delicate customer-relations balance. On the one hand, paying customers may be offended to learn of the availability of a price that is lower, or even zero. On the other hand, paying customers may be motivated to buy when they learn of both the patented technology's success and the patentee's seemingly charitable contributions. Although it may seem crass to call such a contribution "charitable," since its purpose is the facilitation of some other objective (charging a higher price to some customers), presumably every donation willingly made to a charitable cause by a rational actor is done to further some objective of that actor and not to further only someone else's objective. While the net impact of these competing forces is uncertain, the patentee's desires to preserve value while avoiding transaction costs that are discussed *infra* note 24 will likely tip the net balance of incentives to be towards the use of such free licenses in certain cases.

25 This argument and its implications are explored in depth in the important works by Eisenberg et al. *See*, *e.g.*, Michael A. Heller & Rebecca S. Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 Science 698 (1998) (arguing that patents can deter innovation in the field of basic biological research); Rebecca S. Eisenberg, *Property Rights and the Norms of Science in Biotechnology Research*, 97 Yale L.J. 177 (1987) (exploring potential negative impact of patent rights on scientific norms in the field of basic biological research); Rebecca S. Eisenberg, *Patents and the Progress of Science: Exclusive Rights and Experimental Use*, 56 U. Chi. L. Rev. 1017 (1989) (exploring an experimental use exemption from patent infringement as a device for alleviating potential negative impact of patent rights on scientific norms in the field of basic biological research); Rebecca S. Eisenberg, *Public Research and Private Development: Patents & Technology Transfer in Government-Sponsored Research*, 82 Va. L. Rev. 1663 (1996) (offering preliminary observations about the empirical record of the use of patents in the field of basic biological research and recommending a retreat from present government policies of promoting patents in that field).

26 An entitlement enjoys the protection of a property rule if the law condones its surrender only through voluntary exchange. The holder of such an entitlement is allowed to enjoin infringement. An entitlement has the lesser protection of a liability rule if it can be lost lawfully to anyone willing to pay some court-determined compensation. The holder of such an entitlement is only entitled to damages caused by infringement. *See* Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 Harv. L. Rev. 1089 (1972); *see also* Jules L. Coleman & Jody Kraus, *Rethinking the Theory of Legal Rights*, 95 Yale L.J. 1335 (1986).

27 *See, e.g.,* Ian Ayres & Paul Klemperer, *Limiting Patentees' Market Power Without Reducing Innovation Incentives: The Perverse Benefits of Uncertainty and Non-Injunctive Remedies*, 97 Mich. L. Rev. 985 (1999) (criticizing recent increases in certainty in patent law and suggesting the use of liability rules instead of property rules for patent enforcement).

28 28 U.S.C. § 1498 (1994) (providing limited waiver of sovereign immunity for acts of infringement by or for the federal government and instead allowing suits against the government in the U.S. Court of Federal Claims for a reasonable royalty).

29 As explained more fully in Kieff, *supra* note 11:

> Concerning procedure, litigation costs may be high enough to prevent the patentee from seeking court intervention against an infringer. Concerning substance, the limitations on liability that are available to a would-be infringer through the use of the corporate form or bankruptcy laws, for example, may encourage acts of infringements that are essentially judgment proof.

Id. at 734 n.154.

30 In addition to the infringement threats discussed *supra* in the text accompanying notes 29-30, the patentee faces market threats from old technologies, alternative non-infringing technologies, and future technologies. *See* Kieff, *supra* note 11, at 729-31 (collecting sources).

31 *See id*. at 731 (discussing possibility that some licensees may not be able to obtain permission to use the patented invention).

32 *See supra* text accompanying note 29. The recent public demand for the patented drug Cipro® to treat anthrax infection provides an example from the healthcare arena of just such behavior. *See*, *e.g.*, Terence Chea, *Vaccines Are Hot Topic, But Not Hot Investment*, Wash. Post, Dec. 13, 2001, at E1. "At the height of the anthrax crisis, government officials considered overriding German drugmaker Bayer AG's Cipro patent to purchase pills at a better price. Under threat of losing its patent, Bayer agreed to sell the government the antibiotic at half price." *Id*.

33 *See* Douglas Gary Lichtman, *Pricing Prozac: Why the Government Should Subsidize the Purchase of Patented Pharmaceuticals*, 11 Harv. J.L. & Tech. 123, 124-25 (1997) (arguing that the government offer a cash subsidy to any consumer who values a patented good above marginal cost but is unwilling or unable to pay to such a price). *Cf.* Kieff, *supra* note 11, at 716 n.91 (noting that such proposals face the distortion and implementation concerns generally raised against subsidies).

34 Kieff, *supra* note 11, at 732-36 (showing how the potential infringements induced by a liability rule will discourage investment in the commercialization process *ex ante* and may even result in a net destruction of social wealth if the collective costs of entry and exit across infringers exceeds the social surplus otherwise created by the invention).

35 *Id*. at 734 n.152. (citing Robert P. Merges, *Of Property Rules, Coase, and Intellectual Property*, 94 Colum. L. Rev. 2655, 2664 (1994)).

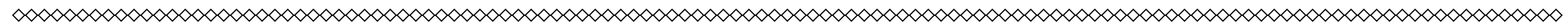
36 *See*, *e.g.*, Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470, 486-87 (1974) (Burger, C.J.) (highlighting, in the context of a discussion about the benefits of allowing even the lesser form of protection provided by trade secret laws, social costs of the self-help measures that would be used by individual actors if legal forms of protection were not available).

37 As discussed generally *supra* note 1, on this particular issue, NAELS Conference organizers suggested that I treat as representative of the environmental community's views the work of Ozcek and Evans. *See*, *e.g.*, Jeremy P. Ozcek, Note, *In the Aftermath of the "Terminator" Technology Controversy: Intellectual Property Protections for Genetically Engineered Seeds and the Right to Save and Replant Seed*, 41 B.C. L. Rev. 627 (2000) (exploring the problems associated with so-called "terminator technologies"); Laura E. Ewens, Note, *Seed Wars: Biotechnology, Intellectual Property, and the Quest for High Yield Seeds*, 23 B.C. Int'l & Comp. L. Rev. 285, 307 (2000) (same).

38 One example of such terminator technology is the Technology Protection System™ from Monsanto. Ozcek, *supra* note 39, at 628.

39 *Id.* at 629; Ewens, *supra* note 39, at 306-07.

40 *See*, *e.g.*, Fred S. McChesney, *Tortious Interference With Contract Versus "Efficient" Breach: Theory and Empirical Evidence*, 28 J. Legal Stud. 131 (1999) (arguing that so-called efficient breaches of contracts are often not efficient when viewed from the dynamic perspective).



41 *See*, *e.g.*, Mallinkcrodt v. Medipart, 976 F.2d 700 (Fed. Cir. 1992) (holding single-use restriction on patented catheter enforceable).

42 One feared mechanism by which such pernicious spreading of terminator technologies might take place is discussed in Kojo Yelpaala, *Owning The Secret Of Life: Biotechnology And Property Rights Revisited*, 32 McGeorge L. Rev. 111 (2000).

> Terminator Seed technology has the potential for serious environmental damage. Through cross-pollination, the Terminator Seed technology could spread from farm to farm and into other varieties of seeds. Given that the Terminator Seed technology can be combined with ordinary non-patented seeds and with other genetically engineered technologies such as the herbicide-resistant plant technology, the spread of the Terminator Seed technology would be virtually unstoppable. Imagine the thousands of different varieties of maize in Mexico being exposed to the Terminator Seed technology from a few farms. With time the technology could threaten the bio-diversity of seeds in Mexico.

Id. at 209.

43 As discussed generally *supra* note 1, on this particular issue, NAELS Conference organizers suggested that I treat as representative of the environmental community's views the work of my colleague, McManis. *See*, *e.g.*, Charles R. McManis, *The Interface Between International Intellectual Property and Environmental Protection: Biodiversity and Biotechnology*, 76 Wash. U. L.Q. 255 (1998) (exploring the influence of patent law on biodiversity).

44 *See* McManis, *supra* note 46, at 268-70 (arguing that Article 16 of the Biodiversity Treaty attempts to ensure that profits are shared with indigenous populations who are custodians of such biodiversity); *see also* Ewans, *supra* note 39, at 289 (commenting on the failure to share such patent profits with indigenous populations).

45 Gollin, *supra* note 8, at 216 (noting that grants of proprietary rights may permit such a system of sharing to develop through private agreements and international initiatives).

46 *See* Kieff, *supra* note 11, at 699 n.4 ("Economic research over the past sixty years has amply established a causal link between the development of intellectual property and the growth of our national economy, while also showing that intellectual property is an increasingly critical component of United States capital and foreign trade.").

47 Indeed, McManis and others devoted substantial personal effort towards this end, and we are indebted for the example they provide. Moreover, even where express individualized contracting cannot be achieved, an international treaty could be created to essentially impose a tax on gains from patents generated though access to such cultures and then transfer the proceeds from such a tax to the custodians. But each of these solutions raises yet other serious problems. For example, it may be quite difficult to identify who "merits" treatment as "custodian" of biodiversity. The national government, a local government, a tribal unit, a family unit, a political leader, a military leader, a spiritual leader, or some host of other individuals and organizations might each lay claim to that status or interfere with the claims by the others.

48 *See*, *e.g.*, *Treaties/Generic Drugs: WHO Urged To Aid Governments In Implementing Trips/Medicines Deal*, 74 Pat. Trademark Copyright J. (BNA) 157 (2007).

49 *See*, *e.g.*, Bryan Mercurio, *Resolving the Public Health Crisis in the Developing World: Problems and Barriers of Access to Essential Medicines*, 5 Nw. U. J. Int'l Hum. Rts. 1, 42 (2006) (collecting sources).

International Law & National Security

Public Committee Against Torture in Israel v. Government of Israel: U.S. Military Commissions Through the Lens of the Supreme Court of Israel

*By Michael A. Newton**

The legal dimension of the struggle to defeat terrorists has at times overshadowed the armed struggle, as the nature of the conflict itself has changed and the tides of politics and public opinion have swirled. In modern conflict, the overall mission is necessarily intertwined with political, legal, and strategic imperatives that are not accomplished in a legal vacuum or by undermining the threads of legality that bind diverse components of a complex operation together.[1] The struggle against transnational terrorists confronts civilized societies and the military commanders in their service with the challenge of implementing humanitarian restraints in an environment marked by utter disregard for the bounds of international law on the part of the adversary. Recognizing the inherent difficulty of a sustained "war against rampant terrorism," the President of the Israel Supreme Court, Aharon Barak, wrote that the "armed conflict is not undertaken in a normative vacuum. It is undertaken according to the rules of international law, which establish the principles and rules for armed conflicts. The armed conflict against terrorism is an armed conflict of the law against those who seek to destroy it."[2] Despite their own obligations to comply with the "law of war during all armed conflicts, however such conflicts are characterized,"[3] professional military forces are confronted with an adversary that intentionally targets civilians and participates in armed conflict without legal authority to do so.

In the context of the massive wave of terrorist acts directed at Israeli society that began in early 2000, termed the *intifada*, thousands of crimes have been committed against innocent civilians, resulting in the deaths of more than a thousand citizens and the injury of thousands more. The Government of Israel in turn adopted a strategy to disrupt terrorist acts, euphemistically termed "the policy of targeted frustration." The governmental policy permits military forces to intentionally target those who unlawfully seek to kill and injure Israeli citizens.[4] On December 11, 2005, the Israel Supreme Court, sitting as the High Court of Justice, addressed the following issue in a sweeping opinion on this policy, frequently referred to as the *Targeted Killings* case:

> The Government of Israel employs a policy of preventative strikes which cause the death of terrorists in Judea, Samaria, or the Gaza Strip. It fatally strikes these terrorists, who plan, launch, or commit terrorist attacks in Israel and in the area of Judea, Samaria, and the Gaza Strip, against both civilians and soldiers. These strikes at times also harm innocent civilians. Does the State thus act illegally? That is the question posed before us.

The *Targeted Killings* opinion provides an extensive analysis of the framework of humanitarian law with regard to terrorist acts as its rationale for upholding the policy. On behalf of the terrorists, the petitioners argued that "the targeted killings policy is totally illegal, and contradictory to international law, Israeli law, and basic principles of human morality. It violates the human rights recognized in Israel and international law, both of those targeted, and the rights of innocent passersby caught in the targeted killing zone."[5]

The Israeli policy, along with its judicial validation, rests on the premise that civilians illegally seeking to kill Israeli citizens have forfeited their otherwise non-derogable right to life. In its unanimous opinion, the bench concluded that "the State's struggle against terrorism is not conducted "outside" the law; it is conducted "inside" the law, with tools that the law places at the disposal of democratic states."[6] The Israeli approach to applying international law as a constraining system of principles in the fight against its enemies relies on the recognition that there is no moral or legal equivalency between terrorists and those who defend the state. The inequity between combatants operating under state authority and terrorists persists because the "state fights in the name of the law and in the name of upholding the law. The terrorists fight against the law, while violating it. The fight against terrorism is also law's war against those who rise up against it."[7] Thus, the court unanimously upheld the policy because the framework of existing humanitarian law warrants the conclusion

> not that such strikes are always permissible or that they are always forbidden. The approach of customary international law applying to armed conflicts of an international nature is that civilians are protected from attacks by the army. However, that protection does not exist regarding those civilians "for such time as they take a direct part in hostilities" (§51(3) of *The First Protocol*).... we cannot determine that a preventative strike is always legal, just as we cannot determine that it is always illegal. All depends upon the question whether the standards of customary international law regarding international armed conflict allow that preventative strike or not.

I. The Normative Framework of Armed Conflict

The legal regime applicable to armed conflict, termed *lex specialis* by the International Court of Justice,[8] is integral to the very notion of military professionalism, because it defines the class of persons against whom military forces can lawfully apply violence based on principles of military necessity and reciprocity.[9] States have historically sought to prescribe the conditions under which they owed particular classes of persons affirmative legal protections in the context of armed conflicts. The constant effort to be as precise as possible in describing the classes of persons entitled to those protections was an essential element of the diplomatic tussles that spawned positivist legal trends. The legal line between lawful and unlawful participants in conflict provided the intellectual impetus for the evolution of the entire field of law relevant to the conduct of hostilities.[10]

* *Michael A. Newton is Acting Associate Clinical Professor of Law at Vanderbilt University Law School.*

Since 1854, there have been over sixty international conventions regulating various aspects of armed conflicts, and a recognizable body of international humanitarian law has emerged from this complex mesh of conventions and custom.[11]

The *lex specialis* law of armed conflict establishes a bright line between those persons governed by its norms and those who derive rights and benefits from other bodies of law. The two essential strands that professional military forces must reexamine and apply in this new style of conflict are: How may we properly apply lethal force? If lawful means of conducting conflict are available, against whom may we properly apply military force? These two strands are the essential foundation of the professional military ethos, even against a lawless enemy. By extension, one of the cornerstones of humanitarian law is the mandate that respect for and protection of the civilian population and civilian objects requires that state parties must always "distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives."[12]

In fact, the detailed proscriptions of the laws and customs of war are shaped by the distinction between civilians and lawful combatants, and the correlative rights and duties that accrue from that status. This distinction forms the basis for the petitioners' arguments in the *Targeted Killings* case, that members of terrorist organizations must be treated as criminals, subject to the applications of relevant provisions of domestic criminal law, rather than as persons who may be intentionally targeted and killed using the instruments of state power.[13] In seeking to remove terrorists from the framework of the *lex specialis* regime applicable to armed conflicts,[14] the petitioners ignore the declarative norm that "the right of belligerents to adopt means of injuring the enemy is not unlimited."[15] The modern formulation of this foundational principle is captured in Article 35 of the 1977 Protocol I to the 1949 Geneva Conventions as follows: "In any armed conflict, the right of the Parties to the conflict to choose methods or means of warfare is not unlimited."[16] The imperative that logically follows from this core organizing principle is that the right of *non-belligerents* to adopt means of injuring the enemy is *nonexistent*. In embarking on the age of positivist legal development, states attempted to clarify the line between lawful combatants and unlawful criminals when the application of the law appeared unresponsive to changing military requirements. While the law evolved as a pragmatic response to the changing tactics of war, it is clear that the legal line remained fixed; the law of armed conflict has never accorded combatant immunity to every person who conducts hostilities. For example, from 1777 to 1782, the British Parliament passed an annual act declaring that privateers operating under the license of the Continental Congress were pirates, and as such could be prosecuted for their acts against the Crown.[17] As one European scholar noted, "unlawful combatants… though they are a legitimate target for any belligerent action, are not, if captured entitled to any prisoner of war status."[18] Persons outside the established framework of law who commit warlike acts do not enjoy combatant immunity and are therefore common criminals subject to prosecution for their actions.[19]

In other words, persons who employ violence amounting to the conduct of hostilities governed by the law of war do so unlawfully unless they can find affirmative legal authority for their acts under international law; those acting with the requisite legal authority have historically been termed belligerents or combatants. Conversely, persons who have no legal right to wage war or adopt means of inflicting injury upon their enemies have been described synonymously as non-belligerents, unprivileged belligerents, unlawful combatants, or unlawful belligerents. Lawful combatants become "war criminals" only when their actions transgress the established boundaries of the laws and customs of war. Prisoners of war, for example, are within the purview of the law of armed conflict and accordingly enjoy legal protection *vis-á-vis* their captors. However, because the legal regime protects them, prisoners of war in turn have no legal right to commit "violence against life and limb."[20] Two essential implications follow from the conceptual foundation of combatant immunity as an offshoot of state sovereignty. First, although the application of international humanitarian law has steadily expanded from international to non-international armed conflicts,[21] the concept of "combatancy" as a legal term of art has been strictly confined to international armed conflicts. Even as the law expanded to grant combatant immunity for irregular forces that do not line up in military uniforms on a parade field, participants in non-international armed conflicts remain completely subject to domestic criminal prosecution for their warlike acts. Protections found in domestic law are grounded not on the status of a person but on the basis of his or her actual activities, because no one has an international law "right to participate in hostilities" in a non-international armed conflict.[22]

Secondly, terrorists who participate in hostile activities in a non-international armed conflict cannot expect any immunity derived from international law for their actions. Therefore, the law applicable to non-international armed conflicts does not provide for combatant status, nor does it define combatants or specify a series of obligations inherent in combatant status. Thus, in order to reach the merits of particular cases, numerous domestic courts have rejected claims of combatant immunity that are unwarranted under existing international law.[23] Indeed, introducing the concept of "combatant immunity" in the context of non-international armed conflicts would grant immunity for acts which would be perfectly permissible in an international armed conflict, such as attacks directed at military personnel or property. This striking silence in the law applicable to non-international armed conflicts means that any effort to describe a "combatant engaged in a non-international armed conflict" is an oxymoron.

Taken together, these principles form the backbone of the law of armed conflict.

II. The Pending *Boumediene v. Bush* Case

The analysis of the issue stated above by the court in the *Targeted Killings* case offers instructive parallels for what will certainly be one of the most closely watched cases during the coming United States Supreme Court term. After President Bush used executive power to promulgate a system of military tribunals to try alien enemies for "offenses triable by military commission,"[24] the Court rejected that system. In *Hamdan v. Rumsfeld*,[25] the Court found that the President's executive order

violated the statutory mandate found in the Uniform Code of Military Justice that requires the President to promulgate "rules and regulations" for courts-martial, military commissions and other military tribunals that are "uniform insofar as practicable."[26] The UCMJ is clear that the statutory provisions for trial by courts-martial "do not deprive military commissions, provost courts, or other military tribunals of concurrent jurisdiction with respect to offenders or offenses that by statute or by the law of war may be tried by military commissions, provost courts, or other military tribunals."[27] Nevertheless, the Court concluded that the Presidential order failed to comply with the legal obligation imposed by Article 3 Common to the Geneva Conventions that prohibits "the passing of sentences and the carrying out of executions without previous judgment pronounced by a regularly constituted court affording all the judicial guarantees which are recognized as indispensable by civilized peoples."[28]

Repeating the pattern previously established following Supreme Court rejections of administration interpretations of Executive power regarding the treatment and punishment of terrorist suspects,[29] the Congress acted swiftly to ameliorate the legal lacunae identified by the Court. Enacting the Military Commissions Act of 2006,[30] Congress took heed of Justice Breyer's observation in his *Hamdan* concurrence that nothing "prevents the President from returning to Congress to seek the authority he believes necessary."[31] The Military Commissions Act specifically authorized the President to promulgate tribunals with jurisdiction over "alien unlawful enemy combatants."[32] The provisions of the 1949 Geneva Conventions form the backdrop for this new category of personal jurisdiction by requiring prisoners of war "can be validly sentenced only if the sentence has been pronounced by the same courts according to the same procedure as in the case of members of the armed forces of the Detaining Power."[33] Hence, the normal jurisdictional provisions of courts-martial extend personal jurisdiction over persons in U.S. custody who are legally entitled to the rights accorded prisoners of war.[34] The Military Commissions Act also expanded normal courts-martial jurisdiction to cover "lawful enemy combatants" who "violate the law of war."[35]

On the other hand, creating an express category of persons subject to the jurisdiction of military commissions, Congress defined the term unlawful enemy combatant to mean

> (i) a person who has engaged in hostilities or who has purposefully and materially supported hostilities against the United States or its co-belligerents who is not a lawful enemy combatant (including a person who is part of the Taliban, al Qaeda, or associated forces); or
>
> (ii) a person who, before, on, or after the date of the enactment of the Military Commissions Act of 2006, has been determined to be an unlawful enemy combatant by a Combatant Status Review Tribunal or another competent tribunal established under the authority of the President or the Secretary of Defense.

The Israeli Imprisonment of Unlawful Enemy combatants Law, in contrast, defines an unlawful combatant as "a person who took part in hostilities against the State of Israel, whether directly or indirectly, or is part of a force that commits hostilities against the State of Israel who does not fulfill the conditions granting prisoner of war status in international humanitarian law."[36] The statutory provisions of the Military Commissions Act and the attendant legal rights flowing from the conjunction of legislative and executive power in establishing jurisdiction over unlawful combatants will be foremost in one of the most anticipated cases scheduled for oral argument during the Supreme Court term beginning in October 2007. In the case *Boumediene et. al. v. Bush*, the D.C. Circuit squarely upheld the provisions of the Military Commissions Act on constitutional grounds.[37] In granting certiorari to reevaluate the scope of rights that accrue to enemy civilians who unlawfully take up arms to conduct hostilities against the United States, its citizens, and its property, the Supreme Court will of necessity address a different manifestation of the same issue confronted by the Israeli Court in the *Targeted Killings* case.

III. Terrorists are not Lawful Combatants

The Geneva Conventions incorporate the principle of lawful combatants only into international armed conflicts, and thus convey Prisoner of War status only to actors participating in an armed conflict between "two or more High Contracting Parties."[38] Because only states enjoyed the historical prerogative of conducting warfare, the principles of lawful combatancy developed from the premise that only states had the authority to sanction the lawful conduct of hostilities. Propelled by the classic view that "the contention must be between States" to give rise to the right to use military force,[39] the concept of "combatant status" developed to describe the class of persons operating under the authority of a sovereign state to wage war. Despite this clear premise of humanitarian law, the Israel Supreme Court determined that the conflict between Israel and terrorist fighters constitutes an international armed conflict because it "crosses the borders of the state."[40] This conclusion is unsupported by any example of state practice and contains no scintilla of legal support, as it represents the only example in which the legal character of a conflict has been made by reference to the geographic boundary rather than the identity of the participants.

The finding that the law of international armed conflicts applies to the *Targeted Killings* facts is unique in light of the historically incremental expansion of the concepts of combatancy. It was only the patriotic resistance of partisan fighters to the German occupations in World War II, and the horrific crimes committed by German forces against the civilian populace in response,[41] that convinced states to update the legal standards for obtaining combatant status during occupation. This is intellectually consistent because during the period of occupation, the sovereign power of the state has been displaced by the occupying power. In effect, fighters acting to repel the occupier are acting as proxies for the military of the defeated sovereign. Although a state of occupation does not "affect the legal status of the territory in question,"[42] the assumption of authority over the occupied territory implicitly means that the existing institutions of society have been swept aside. Because the foreign power has displaced the normal domestic offices, the cornerstone of the law of occupation is the broad obligation that the foreign power "take all the measures in his power to restore,

and ensure, as far as possible, public order and safety...."[43] The fact that the provisions for regulating the status of irregular fighters did not become part of the Fourth Geneva Convention[44] (designed to protect the civilian population) is itself legally significant. As one eminent commentator noted,

> The whole point about the lawful guerrilla fighter, so far as he could be identified and described was that he was not a civilian. The Civilians Convention was for protecting civilians who remained civilians and whose gestures of resistance, therefore, would be punished as crimes, just as would any acts of guerrilla warfare which lay outside whatever lawful scope could be defined.[45]

Accordingly, the 1949 Geneva Conventions restated and updated the law of combatant status within the Third Geneva Convention on Prisoners of War. In addition, the Conventions added a requirement for the detaining state to convene a "competent tribunal" to consider the facts relevant to a particular person's status, when that status is in doubt.[46] This provision was intended to "avoid arbitrary decisions by a local commander,"[47] and in the practice of the U.S. Army may be accomplished with regard to lawful combatants by three officers with expeditious thoroughness.[48] The important point is that, apart from adding language specifically referring to organized resistance movements, the 1949 modifications were taken directly from the Hague Regulations of 1907. Under the Geneva Conventions of 1949, the class of civilians entitled to prisoner of war status was described inter alia as:

> (1) Members of the armed forces of a Party to the conflict as well as members of militias or volunteer corps forming part of such armed forces.
>
> (2) Members of other militias and members of other volunteer corps, *including those of organized resistance movements, belonging to a Party to the conflict and operating in or outside their own territory, even if this territory is occupied,* provided that such militias or volunteer corps, *including such organized resistance movements*, fulfil the following conditions:
>
> (a) that of being commanded by a person responsible for his subordinates;
>
> (b) that of having a fixed distinctive sign recognizable at a distance;
>
> (c) that of carrying arms openly;
>
> (d) that of conducting their operations in accordance with the laws and customs of war."[49]

Despite its error in considering the conflict as one bound by the laws and customs applicable to international armed conflicts, the *Targeted Killings* court correctly applied these principles to conclude that the petitioners did not enjoy the protections granted to legal combatants, and as such "are not entitled to the status of prisoners of war; they may be put on trial for their membership in terrorist organizations and for their operations against the army."[50]

IV. Terrorists are not Civilians

The State of Israel requested the *Targeted Killings* court to accept the premise that terrorists are unlawful combatants in the sense that they "take active and continuous part in an armed conflict, and therefore should be treated as combatants in the sense that they are legitimate objects of attack."[51] This argument asked the court to employ the term 'combatant' in its common sense meaning rather than as a legal term of art derived from preexisting law. In light of its previous decision to apply the full body of the Geneva and Hague Conventions, the court was forced to address the specific mandate of Protocol I, which defines the term "civilian" purely in contradistinction to the opposing status of lawful combatant. Article 50 embodies a dualist view by defining a civilian as "any person who does not belong" to one of the specified categories of combatant.[52] The International Committee of the Red Cross (ICRC) has long held the position that because there "is no intermediate status" between combatant and civilian, every person in enemy hands "must have some status under international law."[53] Though the ICRC explicitly recognized that "[m]embers of resistance movements must fulfill certain stated conditions before they can be regarded as prisoners of war,"[54] its dualist preference led it to advocate that even those who fight unlawfully remain entitled to the benefits accorded to lawful combatants. Despite overturning several centuries of legal development and judicial practice,[55] the ICRC noted that this dualism would be "satisfying to the mind" in that it elevates the humanitarian point of view "above all."[56] The *Targeted Killings* court refused to accept the State's categorization of terrorists as unlawful combatants, and implicitly accepted the ICRC dualist view, because the "question before us is not one of desirable law, rather one of existing law. In our opinion, as far as existing law goes, the data before us are not sufficient to recognize this third category."[57]

In theory, the ICRC dualist view enshrined in Article 50 and adopted by the court protects unlawful combatants from the effects of their misconduct. Permitting terrorists to be legally classified as civilians puts the military forces opposing terrorist activities into the quandary of either supinely permitting the planning and conduct of terrorist activities or violating the clear legal norm that the "civilian population as such, as well as individual civilians shall not be the object of attack."[58] Article 50 seems to embody a system in which there is no theoretical gap; a person is either a combatant or a civilian. This leads ineluctably to the assumption that an unlawful combatant who fails to qualify as a prisoner of war must be a civilian entitled to protection. The plaintiffs in *Targeted Killings* argued precisely with this logic in advocating a revolving door immunity that would grant terrorists immunity from harm for the entire time they plan their attacks, suspend that immunity only for a very narrow window during which the actual attack is underway, then become protected again immediately after the attack even as they return to their homes intending to plan and execute their next attack. According to the plaintiffs, the state must demonstrate a high degree of imminency to justify any attack on terrorists, who continue to wear the mantle of "civilian," thus losing protections "*only* during such time that he is taking a direct and active part in hostilities, and *only* for such time as said direct participation continues"(emphasis added).[59]

Nevertheless, Protocol I sustained the preexisting principle of unlawful combatancy by specifying that civilians enjoy the protections embodied in the Protocol "unless and for such

time as they take a direct part in hostilities."[60] By choosing to participate in hostilities, the unlawful combatant has separated himself legally from the shield afforded by international law. As Professor Yoram Dinstein has observed, "one cannot fight the enemy and remain a civilian."[61] Just as it is possible to lose combatant status (by becoming a prisoner of war, for example), and the immunity that goes with it (by failure to comply with the law of war), the unlawful belligerent cannot properly be termed a civilian in the same sense as those innocents who huddle in their homes while combat rages round them. In using this legal rationale to uphold the Israeli policy of intentionally directing attacks against terrorists, The *Targeted Killings* court correctly characterized the plaintiffs' arguments as "unacceptable."[62] Permitting terrorists to "change their hat at will, between the hat of a civilian and the hat of a combatant" would reduce the legal structure regulating conflicts to obsolescence. In his concurrence, Judge Rivlin termed this class "international law breaking civilians, whom I would call "uncivilized civilians." The *Targeted Killings* case reinforces the laws and customs of war by recognizing that granting legal protection to any civilian who takes up arms in violation of the rules specified in international law would create a perverse incentive to defy the conventions regulating conduct and deliberately conduct hostilities outside the bounds of the law all the while relying on the goodwill of the enemy to apply that same body of law.

Though it makes no reference to the underlying negotiations in 1949, the conclusion of the *Targeted Killings* court also preserves the intentions behind the Geneva Conventions. At the time, some delegates pushed for a broader interpretation of the textual provisions that would have protected illegal combatants based on the understanding that the "categories named in Article 3 [the present Article 4 of the Third Geneva Convention] cannot be regarded as exhaustive, and it should not be inferred that other persons would not also have the right to be treated as prisoners of war."[63] This position was politely but firmly rejected by the delegates in favor of the view that the text itself is the exclusive source to "define what persons are to have the protection of the Convention."[64] In fact, a number of other delegations explicitly confirmed that the textual provisions of the Geneva Conventions did not foreclose the traditional category of unlawful combatants. The Dutch delegate pointed out that a summary conclusion to the effect that combatants who did not meet the criteria for prisoner of war status "are automatically protected by other conventions is certainly untrue."[65] He further clarified that the Fourth Geneva Convention (the civilians convention) "deals only with civilians under certain circumstances; such as civilians in an occupied country or civilians who are living in a belligerent country, *but it certainly does not protect civilians who are in the battlefield, taking up arms against the adverse party*."[66] Furthermore, the United Kingdom delegate observed that "the whole conception of the Civilians Convention was the protection of civilian victims of war and not the protection of illegitimate bearers of arms, who could not expect full protection under rules of war to which they did not conform."[67] Thus, the *Targeted Killings* case used a semantic distinction to reach an outcome that is consistent with the underlying fabric of the Geneva Conventions.

V. Unlawful Combatants in United States Law

In *Boumediene v. Bush*, the Supreme Court will address the legal categorization of private persons who take up arms and their correlative rights against the United States in light of the MCA provision that terms them "unlawful combatants." The MCA definition of "unlawful combatants" is substantively accurate despite the semantic imprecision that could have been avoided by using the labels "unlawful belligerents" or "unlawful participants." The United States refused to accept Protocol I in part on the basis that Article 44(3) contributed to the "essence of terrorist criminality" by its "obliteration of the distinction between combatants and non-combatants."[68] The United States rejected Protocol I on precisely the same intellectual grounds used by the *Targeted Killings* court to uphold the Israeli government policy. Protocol I appears to pave the way for any civilian to take up arms when the fancy strikes, engage in hostilities, put down his or her weapons, hide among the innocent civilian populace, strike at will, and yet claim combatant status (with an accompanying combatant immunity from prosecution for those warlike acts). This on/off combatant status would effectively erode the law of unlawful combatancy to its vanishing point. Article 44, paragraph 3 of Protocol I undermines the traditional qualifications for achieving combatant status by accepting the notion that there may be some circumstances "owing to the nature of the hostilities" in which combatants cannot distinguish themselves from the civilian population.[69] In such circumstances, the duty to distinguish may be watered down to the point that the combatant need only carry his arms openly "during each military engagement" or when "visible to the enemy."[70] In the real world of state practice, the protections ostensibly provided to terrorists under Protocol I have remained inapplicable and have by no means been universally accepted as a matter of law, nor do they obviate the other requirements for lawful combatant status set forth in the Hague Regulations and Article 4 of the Geneva Conventions of 1949. Thus, the explicit recognition of a category of "unlawful combatants" is valid under both United States and international law.

The Military Commissions Act further conforms United States practice to international law by establishing a mechanism for the prosecution of persons who take part in hostilities but are not legally entitled to prisoner of war status. The official ICRC Commentary to Protocol I specified that "anyone who participates directly in hostilities without being subordinate to an organized movement under a Party to the conflict, and enforcing compliance with these rules, *is a civilian who can be punished for the sole fact that he has taken up arms"* (emphasis added).[71] The ICRC text further restates the long-established principle that "anyone who takes up arms without being able to claim this status [of a "lawful combatant"] will be left to be dealt with by the enemy and its military tribunals in the event that he is captured."[72] Accepting the reality that such persons are unlawful belligerents who may be prosecuted for their warlike acts, Protocol I describes a minimum set of due process obligations applicable to such prosecutions. Article 45 provides that "[a]ny person who has taken part in hostilities, who is not entitled to prisoner-of-war status and who does not benefit

from more favourable treatment in accordance with the Fourth Convention shall have the right at all times to the protection of Article 75 of this Protocol."[73] The MCA Provisions meet and exceed the relevant international standards.

Finally, international law is clear that no accused should face punishment unless convicted pursuant to a fair trial affording all of the essential guarantees embodied in widespread state practice.[74] Common Article 3 states this principle with particularity by requiring that only a "regularly constituted court" may pass judgment on an accused person.[75] Interpreting this provision in light of state practice, the ICRC concluded that a judicial forum is "regularly constituted if it has been established and organized in accordance with the laws and procedures already in force in a country."[76] Accepting the ICRC benchmark of legitimacy, the MCA provisions meet the criteria derived from the law of war because they apply the general principles of criminal law drawn from the existing Manual for Courts-Martial with only slight (but important) modifications. Similarly, the MCA provisions are "established by law" in accordance with human rights principles.[77] The United Nations Human Rights Commission adopted a functional test that the tribunal should "genuinely afford the accused the full guarantees" in its procedural protections.[78] Noting that the ICCPR drafters rejected language specifying that only "pre-established" forums would provide sufficient human rights protections, the ICTY Appeals Chamber concluded:

> The important consideration in determining whether a tribunal has been "established by law" is not whether it was pre-established or established for a specific purpose or situation; what is important is that it be set up by a competent organ in keeping with the relevant legal procedures, and should that it observes the requirements of procedural fairness.[79]

The MCA provisions for external review of the ground for detention and jurisdiction over an unlawful combatant pending trial meet these international standards. Indeed, the Geneva Conventions specifically permit the deprivation of liberty even for innocent civilians who are fully protected by their provisions if 'the security of the Detaining Power makes it absolutely necessary.'[80] Even large-scale internment may be permissible in situations where there are 'serious and legitimate reasons' to believe that the detained persons threaten the safety and security of the occupying power.[81] The Geneva Conventions nowhere reference any right of such detained civilians to review by an external judicial body pursuant to a grant of habeas corpus. It would be unprecedented for the Supreme Court to convey greater procedural rights to unlawful combatants than those enjoyed by persons fully protected by the laws and customs of war.

CONCLUSION

September 11th reopened international discussion over the basic division between those persons and incidents within the proper purview of the law of war, as opposed to those persons and incidents within the purview of criminal conduct. In describing terrorists under the armed conflict rubric, the phrases unlawful combatants, unlawful participants, unlawful belligerents, illegal saboteurs, and unprivileged belligerents are merely semantic distinctions to describe the same legal relationship between the person taking up arms and the other persons caught up in the conflict. Despite the holding of the *Targeted Killings* case, in international law today there is no legal category that might be described as an unprotected civilian. The proper characterization of such individuals is the historically accepted determination that they are unlawful combatants (or unprivileged belligerents). Persons who take part in hostilities without meeting the legal criteria as prisoners-of war are neither protected civilians nor lawful combatants. Therefore, though the historical evolution of the law of armed conflict has steadily shrunk the range of conduct deemed to be outside the boundaries of international law, the principle of unlawful combatancy has survived intact.

In theory, the modern law of armed conflict is nothing more than a web of interlocking protections and specific legal obligations held together by the thread of respect for humankind and a reciprocal expectation that other participants in armed conflict are bound by the same normative constraints. This foundational premise is at the heart of the challenge posed by the modern manifestations of transnational terrorism. Recent history shows that the expectations of reciprocity and professionalism that make the law of armed conflict a coherent whole are increasingly challenged in practice. The European Parliament opined that the law of armed conflict must "be revised to respond to the new situations created by the development of international terrorism."[82] The *Targeted Killings* opinion postulates that legal norms "developed against the background of a reality which has changed must take on a dynamic interpretation which adapts them, in the framework of accepted international rules, to the new reality."[83] The MCA provisions as enacted by the United States Congress accomplished precisely this goal.

Endnotes

1 The United States doctrine for Counterinsurgency that was rewritten and repromulgated in response to operations in Iraq makes this principle plain by referring to legal considerations in more than 100 places throughout its text. Dep't of The Army, Field Manual No. 3-24, Marine Corps Warfighting Publication No. 3-33.5, Counterinsurgency (Dec. 15 2006), *available at* http://usacac.army.mil/CAC/Repository/Materials/COIN-FM3-24.pdf.

2 H.C. 3451/02, *Almadani v. IDF Commander in Judea and Samaria*, 56 (3) P.D. 30, 30-35, *quoted in* Judgments of the Israel Supreme Court: Fighting Terrorism within the Law 13 (2005).

3 *See, e.g.*, DOD. Dir. 2311.01E, para. 4.1.

4 *Pub. Comm. Against Torture in Israel v. Gov't of Israel*, HCJ 769/02, para. 2, *available at* http://elyon1.court.gov.il/Files_ENG/02/690/007/a34/02007690.a34.pdf . The majority opinion was written by President Emeritus A. Barak, with President D. Beinisch and Vice President E. Rivlin concurring and attaching separate comments [hereinafter *Targeted Killings Judgment*].

5 *Id.* para. 3.

6 *Id.* para. 61.

7 *Almadani, supra* note 1, at 34, *quoted in Targeted Killings Judgment, supra* note 4, para. 62.

8 *See* Michael J. Matheson, *The Opinion of the International Court of Justice on the Threat or Use of Nuclear Weapons*, 91 Am. J. Int'l. L. 417, 422 (1997) ("the test of what is an arbitrary deprivation of life, however, falls to be determined

by the applicable lex specialis, namely, the law applicable in armed conflict which is designed to regulate the conduct of hostilities.").

9 *See generally* Leslie C. Green, *What is – Why is There – The Law of War, in* Essaya on the Modern Law of War 1 (2d. ed. 1999).

10 The field is frequently described as international humanitarian law. This vague rubric is increasingly used as shorthand to refer to the body of treaty norms that apply in the context of armed conflict as well as the less distinct internationally accepted customs related to the treatment of persons. The core of the international law of war includes the Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field, *opened for signature* Aug. 12, 1949, 75 U.N.T.S. 31, 6 U.S.T. 3114 (replacing previous Geneva Wounded and Sick Conventions of 22 August 1864, 6 July 1906, and 27 July 1929 by virtue of Article 59); Geneva Convention for the Amelioration of the Condition of Wounded, Sick, and Shipwrecked Members of Armed Forces at Sea, *opened for signature* Aug. 12, 1949, 75 U.N.T.S. 85, 6 U.S.T. 3217 (replacing Hague Convention No. X of 18 October 1907, 36 Stat. 2371); Geneva Convention Relative to the Treatment of Prisoners of War, *opened for signature* Aug. 12, 1949, 75 U.N.T.S. 287, 6 U.S.T. 3316 (replacing the Geneva Convention Relative to the Protection of Prisoners of War of 27 July 1929, 47 Stat. 2021) [*hereinafter* Geneva Convention on Prisoners of War]; Geneva Convention Relative to the Protection of Civilians in Time of War, *opened for signature* Aug. 12, 1949, 75 U.N.T.S. 287, 6 U.S.T. 3516 [*hereinafter* Civilians Convention].

11 M. Cherif Bassiouni, Crimes Against Humanity in International Criminal Law 56 (2d ed. 1999).

12 *See* 1977 Protocol Additional to the Geneva Conventions of 12 August 1949 and Relating to the Protection of Victims Of International Armed Conflicts, Dec. 12, 1977, 1125 U.N.T.S. art 483, 23 [*hereinafter* Protocol I].

13 *Targeted Killings Judgment, supra* note 4, para. 4.

14 For example, beginning in 1874, states accepted the principle that "the laws of war do not recognize in belligerents an unlimited power in the adoption of means of injuring the enemy." *See* The Brussels Project of an International Declaration concerning the Laws and Customs of War, Art. 12, *reprinted in* Dietrich Schindler and Jiri Toman, The Laws of Armed Conflicts 25-34 (2d ed. 1981) [*hereinafter* Brussels Declaration]. By 1907, this concept morphed into the phrase "the right of belligerents to adopt means of injuring the enemy is not unlimited."1907 Hague Regulations, *supra* note 12, art 22.

15 Regulations annexed to Hague Convention IV Respecting the Laws and Customs of War on Land, 1907, *entered into force* Jan. 26, 1910, *reprinted in* Documentation on the Laws of War 73 (3d ed., eds. Adam Roberts & Richard Guelff 2000)[*hereinafter* 1907 Hague Regulations].

16 Protocol I, *supra* note 12, art. 48.

17 Alfred P. Rubin, The Law of Piracy, 63 Naval War Col. Int'l. L. Stud. 154 (1988), *citing* 16 Geo. III c. ix (1777).

18 Ingrid Detter de Lupis, The Law of War 148 (2000) (noting that unlawful combatants are "also personally responsible for any action they have taken and may thus be prosecuted and convicted for murder if they have killed an enemy soldier").

19 In a classic treatise, Professor Julius Stone described the line between lawful participants in conflict and unprivileged or "unprotected" combatants as follows, Julius Stone, Legal Controls on International Conflict 549 (1954):

> The distinction draws the line between personnel who, on capture, are entitled under international law to certain minimal treatment as prisoners of war, and those not entitled to such protection. 'Non-combatants' who engaged in hostilities are one of the classes deprived of such protection… Such unprivileged belligerents, though not condemned by international law, are not protected by it, but are left to the discretion of the belligerents threatened by their activities.

20 Geneva Convention on Prisoners of War, *supra* note 10, art. 93.

21 *See, e.g,. Rome Statute of the International Criminal Court*, United Nations Diplomatic Conference of Plenipotentiaries on the Establishment of an International Criminal Court, July 17, 1998, U.N. Doc. A/CONF.183/9, art. 8(2)(e).

22 Marco Sassoli & Antoine Bouvier, How Does Law Protect in War? 208 (International Committee of the Red Cross 1999).

23 In one of the most widely publicized cases in the current struggle, a*vailable at* http://www.globalspecialoperations.com/reid.html, United States District Court Judge William Young announcing a life sentence against Richard Reid, the so-called shoe bomber, with the emphatic pronouncement that

> Here in this court, where we deal with individuals as individuals, and care for individuals as individuals, as human beings we reach out for justice, you are not an enemy combatant. You are a terrorist. You are not a soldier in any war. You are a terrorist. To give you that reference, to call you a soldier gives you far too much stature. Whether it is the officers of government who do it or your attorney who does it, or that happens to be your view, you are a terrorist. And we do not negotiate with terrorists. We do not treat with terrorists. We do not sign documents with terrorists. We hunt them down one by one and bring them to justice. So war talk is way out of line in this court. You are a big fellow. But you are not that big. You're no warrior. I know warriors. You are a terrorist. A species of criminal guilty of multiple attempted murders."

See also The Military Prosecutor v. Omar Mahmud Kassem and Others, Israeli Military Court, Ramallah, April 13, 1969, 41 I.L.R. 470 (1971), *reprinted in* Howard Levie, Documents on Prisoners of War, 60 Naval War Col. Int. L. Stud. 771 (1979) (rejecting the claim of combatant immunity raised by a member of the "Organization of the Popular Front for the Liberation of Palestine"); House of Lords (Privy Council), Osman Bin Haji Mohamed Ali and Another Appellant and the Public Prosecutor Respondent On Appeal From the Federal Court of Malaysia, 1 Law Rep. 430 (1969), *reprinted in* Marco Sassoli & Antoine Bouvier, How Does Law Protect in War? 767 (International Committee of the Red Cross 1999) (rejecting combatant status for members of the Indonesian armed forces who failed to comply with the provisions of Article 4 of the Geneva Conventions).

24 *See* Military Order, Detention, Treatment, and Trial of Certain Non-Citizens in the War Against Terrorism, 66 Fed. Reg. 57,831 (Nov. 16, 2001).

25 548 U.S. ____, 126 S.Ct. 2749, 165 L.Ed. 2d 723 (2006).

26 10 U.S.C. § 836 (also permitting the President latitude implement procedural rules that "so far as he considers practicable, apply the principles of law and the rules of evidence generally recognized in the trial of criminal cases in the United States district courts, but which may not be contrary to or inconsistent with this chapter.).

27 10 U.S.C. § 821. *See also* Michael A. Newton, *Continuum Crimes: Jurisdiction over Foreign Nationals Who Commit International Crimes,* 153 Mil. L. Rev. 1 (1996)(an early call for expansion of UCMJ authority under article 21 to expand the jurisdiction of military commissions).

28 *Hamdan*, *supra* note 25, at 2790, *quoting* Geneva Convention on Prisoners of War, *supra* note 10, art. 3(1)(d).

29 Following the decision by Justice O'Connor in *Hamdi v. Rumsfeld*, 542 U.S. 507 (2004) in which a plurality of Justices required the military to hold status hearings to verify the United States does not mistakenly detain "the errant tourist, embedded journalist, or local aid worker." *Id.* at 534. The military quickly created a system of Combatant Status Review Tribunals, *See* http://www.defenselink.mil/news/Combatant_Tribunals.html Congress responded with the Detainee Treatment Act of 2005, Pub. L. No. 109-148, 119 Stat. 2680 (2005) (DTA), which the President signed into law on December 30, 2005.

30 P.L. 109-366, October 17, 2006 (passing the House and Senate by a combined vote of 313-204-12).

31 126 S.Ct. at 2799.

32 10 U.S.C. § 948(c).

33 Geneva Convention on Prisoners of War, *supra* note 10, art. 102.

34 10 U.S.C. § 802(a)(9).

35 *See* 10 U.S.C. § 802(a)(13)[enacted by the Military Commissions Act, P.L. 109-366, sec. 4].

36 *Targeted Killings Judgment, supra* note 4, para. 25.

37 476 F.3d 981 (D.C. Cir. 2007), *cert. granted on rehearing* 2007 LEXIS

US 8757 (June 29, 2007), consolidated with *Al Odah, et. al. v United States*, 2007 LEXIS 8810.

38 Geneva Convention on Prisoners of War, *supra* note 10, art. 102.

39 L. Oppenheim, International Law: A Treatise, § 56 at 203 (H. Lauterpacht ed., 7th ed., 1952).

40 *Targeted Killings Judgment, supra* note 4, para. 18.

41 *See, e.g.,* U.S. v. Otto Ohlendorf et al. (The Einsatzgruppen Case), (Case No. 9), 4 L. Rep. of Trials of War Criminals 411 (U.N. War Crimes Comm. 1948), *reprinted in* Howard Levie, Documents on Prisoners of War, 60 Naval War Col. Int'l. L. Stud. 408 (1979); Italy, Sansoli and Others v. Bentivegna and Others, Court of Cassation, 24 Int. L. Rep. 986 (1958), *reprinted in* Marco Sassoli & Antoine Bouvier, How Does Law Protect in War? 701 (International Committee of the Red Cross 1999) (upholding the legal status of Italian irregulars in fighting against German occupation).

42 Protocol I, *supra* note 12, art. 8. The United States policy in this regard is clear that occupation confers only the "means of exercising control for the period of occupation. It does not transfer the sovereignty to the occupant, but simply the authority or power to exercise some of the rights of sovereignty." Department of the Army, Field Manual 27-10: The Law of Land Warfare, para. 358 (1956) [*hereinafter* FM 27-10] (devoting the whole of Chapter 6 of the United States Army Field Manual related to the law of armed conflict to explicating the text of the law related to occupation as well as the United States policy related to occupation), *available at* http://www.globalsecurity.org/military/library/policy/army/fm/27-10/.

43 1907 Hague Regulations, *supra* note 15, art. 42 (emphasis added). UK Ministry of Defence, The Manual of the Law of Armed Conflict 275, P 11.3 (2004).

44 Civilians Convention, *supra* note 10.

45 Geoffrey Best, War and Law Since 1945 127-29 (1994).

46 *Id.*, art. 5.

47 IIb Final Record of the Diplomatic Conference of Geneva of 1949, at 270.

48 FM 27-10, *supra* note 41, para. 71(c).

49 Geneva Convention on Prisoners of War , *supra* note 10, art. 4 (emphasis added). This is only a partial listing of the provisions most relevant for determining the status of terrorists and their supporters.

50 *Targeted Killings Judgment, supra* note 4, para. 25.

51 *Targeted Killings Judgment, supra* note 4, para. 27.

52 Protocol I, *supra* note 12, art. 50.

53 The Geneva Conventions of 12 August 1949: IV Geneva Convention Relative to the Protection of Civilian Persons in Time of War, Geneva 51(1958).

54 *Id.* at 50 ("There are certain cases about which some hesitation may be felt. We may mention, first, the case of partisans, to which Article 4, A (2) , of the Third Convention refers. Members of resistance movements must fulfil certain stated conditions before they can be regarded as prisoners of war. If members of a resistance movement who have fallen into enemy hands do not fulfil those conditions, they must be considered to be protected persons within the meaning of the present Convention. That does not mean that they cannot be punished for their acts, but the trial and sentence must take place in accordance with the provisions of Article 64 and the Articles which follow it").

55 *See, e.g.,* The Hostages Trial: Trial of Wilhelm List and Others (Case No. 47), 8 L. Rpts. of Trials of War Criminals 34, 58 (U.N. War Crimes Comm. 1948) ("Guerilla warfare is said to exist where, after the capitulation of the main part of the armed forces, the surrender of the government and the occupation of its territory, the remnant of the defeated army or the inhabitants themselves continue hostilities by harassing the enemy with unorganised forces ordinarily not strong enough to meet the enemy in pitched battle. They are placed much in the same position as a spy. By the law of war it is lawful to use spies. Nevertheless, a spy when captured, may be shot because the belligerent has the right, by means of an effective deterrent punishment, to defend against the grave dangers of enemy spying. The principle therein involved applied to guerrillas who are not lawful belligerents. Just as the spy may act lawfully for his country and at the same time be a war criminal to the enemy, so guerrillas may render great service to their country and, in the event of success, become heroes even, still they remain war criminals in the eyes of the enemy and may be treated as such. In no other way can an army guard and protect itself from the gadfly tactics of such armed resistance. And, on the other hand, members of such resistance forces must accept the increased risks involved in this mode of fighting. Such forces are technically not lawful belligerents and are not entitled to protection as prisoners of war when captured.").

56 The Geneva Conventions of 12 August 1949: IV Geneva Convention Relative to the Protection of Civilian Persons in Time of War, Geneva 51(1958) ("In short, all the particular cases we have just been considering confirm a general principle which is embodied in all four Geneva Conventions of 1949. Every person in enemy hands must have some status under international law: he is either a prisoner of war and, as such, covered by the Third Convention, a civilian covered by the Fourth Convention, or again, a member of the medical personnel of the armed forces who is covered by the First Convention. ' There is no ' intermediate status; nobody in enemy hands can be outside the law. We feel that that is a satisfactory solution—not only satisfying to the mind, but also, and above all, satisfactory from the humanitarian point of view").

57 *Targeted Killings Judgment, supra* note 4, para. 27.

58 Protocol I, *supra* note 12, art. 51(2).

59 *Targeted Killings Judgment, supra* note 4, para. 6.

60 Protocol I, *supra* note 12, art. 51(3).

61 Yoram Dinstein, "*Unlawful Combatants*", 32 Isr. Y.b. on Hum. Rts. 247, 248 (2002).

62 *Targeted Killings Judgment, supra* note 4, para. 13.

63 IIb Final Record of the Diplomatic Conference of Geneva of 1949, at 268 (quoting the Dutch delegate who argued that "the cases provided for by Article 3 must be treated separately and in accordance with present-day international law").

64 *Id.*

65 *Id.* at 271.

66 *Id.* (emphasis added).

67 *Id.* at 621. Brigadier Page, the United Kingdom delegate, noted that illegal combatants "should no doubt be accorded certain standards of treatment, but should not be entitled to all the benefits of the convention." *Id*. The Swiss delegate expressed an almost identical view that "in regard to the legal status of those who violated the laws of war, the Convention could not of course cover criminals or saboteurs." *Id.*

68 A Message from the President of the United States Regarding Protocol II Additional to the 1949 Geneva Conventions, and Relating to the Protection of Victims of Non-International Armed Conflicts, Treaty Doc. 100-2, Dec. 13. 1986, *reprinted in* Marco Sassoli & Antoine Bouvier, How Does Law Protect in War? 603, 604 (International Committee of the Red Cross 1999).

69 Protocol I, *supra* note 12, art. 44(3). The text does contain some qualifiers by which Protocol I proponents sought to legitimize this erosion of traditional principles. For example, Article 44(7) specifies that "[t]his Article is not intended to change the generally accepted practice of States with respect to the wearing of uniforms by combatants assigned to the regular, uniformed armed units of a Party to the conflict."

70 *Id.*

71 Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949 at 514 (Y Sandoz, Ch.Swinarski, & B. Zimmermann eds., 1987) [*hereinafter* ICRC Commentary on Protocol I] (emphasis added).

72 *Id.*

73 Protocol I, *supra* note 12, art. 45(3).

74 For a summary of state practice and its implementation in treaty norms

and military manuals around the world, see 1 Jean-Marie Henckaerts & Louise Doswald-Beck, Customary International Humanitarian Law 352-75 (2005) [hereinafter ICRC Study].

75 Geneva Convention on Prisoners of War, *supra* note 10, Art. 3.

76 ICRC Study, *supra* note 74, at 355.

77 International Covenant on Civil and Political Rights art. 14, opened for signature Dec. 19, 1966, 999 U.N.T.S. 171, 6 I.L.M. 368 (entered into force Mar. 23, 1976) [*hereinafter* ICCPR] (describing analogous provisions derived from international human rights law).

78 *See* Human Rights Comm., General Comment on Article 14, P 4, U.N. Doc. A/43/40 (1988); Human Rights Comm., Cariboni v. Uruguay, U.N. Doc. A/39/40 (Oct. 27, 1987). The Inter-American Commission has taken a similar approach. *See, e.g.,* Inter-Am C.H.R., Annual Report 1972, OEA/Ser. P, AG/doc. 305/73 rev. 1, 14 March 1973, at 1; Inter-Am. C.H.R., Annual Report 1973, 2-4, OEA/Ser. P, AG/doc. 409/174 (March 5, 1974).

79 Prosecutor v. Tadic Case No. IT-94-1-AR72, 1995 WL 17205280, Decision on the Defense Motion for Interlocutory Appeal on Jurisdiction, P 45 (Oct. 2, 1995), *available at* http://www.un.org/icty/tadic/appeal/decision-e/51002.htm.

80 Civilians Convention, *supra* note 10, art 42.

81 *Prosecutor v. Delalic, Mucic, Delic, and Landzo (Celibici)*, Case No. IT-96-21-T, Judgment, 20 Feb 2001.

82 European Parliament Resolution B5-0066/2002 (Feb. 7, 2002).

83 *Targeted Killings Judgment, supra* note 4*,* para. 28.

Labor and Employment Law

Sarbanes-Oxley Whistleblower Provisions: A Study in Statutory Construction

*By Jay P. Lechner**

"When I use a word," Humpty Dumpty said, in rather a scornful tone, "it means just what I choose it to mean—neither more nor less." "The question is," said Alice, "whether you can make words mean so many different things." "The question is," said Humpty Dumpty, "which is to be master—that's all."[1]

In the February 2007 issue of *Engage*, J. Gregory Grisham and James H. Stock, Jr. expertly explored the civil whistleblower provision of the Sarbanes-Oxley Act ("SOX") and addressed criticisms that the scope of SOX's civil whistleblower provision does not provide sufficient protection to employees. The authors concluded that plaintiffs have not fared well under this provision and that judges have "been reluctant to stray from the specific statutory language set out in Section 1514A…."[2] The purpose of the instant article is not to recover the same ground already thoroughly addressed by Grisham and Stock, but rather to argue that judges have been inconsistent in their interpretation of Section 1514A, and in fact have often strayed from its specific statutory language to the detriment of employers. Furthermore, the ambiguous and incomplete language of both SOX's civil and criminal whistleblower provisions invites further unintended anti-employer construction, unless judges are consistent in strictly construing the statutory text.

In the rush to push SOX through Congress, nearly no discussion took place regarding its whistleblower provisions. Thus, statutory gaps, omissions, and ambiguities inevitably slipped through the cracks.[3] As a result, it has been left to judges to fill those holes on a case-by-case basis through statutory construction. As is often the case with matters of statutory construction, judges have interpreted the pertinent statutory provisions inconsistently. Exacerbating these problems, SOX governs fields such as corporate governance mandates, accounting standards, and corporate whistleblower protections that traditionally were reserved for the states, and for which federal agencies such as the Department of Labor lack any meaningful experience.[4]

Not surprisingly, SOX's ambiguities have given rise to unanticipated legal issues. For example, courts and administrative law judges (ALJs) have come to conflicting conclusions as to whether SOX's whistleblower provisions extend to subsidiaries of publicly traded companies, whether their jurisdiction extends to parent companies of non-publicly traded employers, and what type of complaints may constitute "protected activity." Moreover, judges, and even foreign governments, have struggled with issues arising from application of SOX's whistleblower protections to overseas companies and employees. Finally, ambiguities in SOX's Section 1107 criminal whistleblower provisions have presented complex issues which have yet to be resolved by the courts.

** Jay P. Lechner is an attorney with Greenberg Traurig, P.A., in Tampa, Florida, where he represents employers in the areas of labor and employment litigation, including whistleblower, discrimination, and wage and hour.*

I. Whether Employees of Subsidiaries Are Covered

Section 806, SOX's civil whistleblower provision, clearly covers publicly traded companies. It does not expressly cover subsidiaries of publicly traded companies. Yet, many publicly traded companies own and operate through subsidiaries; indeed, Enron had over 2,500. In light of SOX's statutory purpose, should employees of these subsidiaries be covered under SOX's whistleblower provisions? In interpreting the relevant statutory language (with varying degrees of liberality), judges have addressed three distinct inquiries: (1) whether the employee of the subsidiary is a covered "employee" under SOX; (2) if so, if the employee names the subsidiary as a respondent, whether the subsidiary is a covered entity subject to suit; and (3) if the employee names the parent as a respondent, whether the existence of separate corporate identities insulates the parent from liability.

A. Is an Employee of a Subsidiary a Covered "[E]mployee"?

Section 806 prohibits publicly traded companies from retaliating against whistleblowing "employees."[5] The term "employee" is not defined, but employees of subsidiaries of publicly traded companies are not expressly covered. Despite this lack of statutory authority, judges and ALJs consistently have found that, based primarily on perceived legislative intent and common law principles, employees of subsidiaries are covered under SOX.

In some cases, judges have looked to the interrelatedness of the corporate structures in ultimately concluding the subsidiary's employee was covered. For example, in *Collins v. Beazer Homes USA, Inc.*, a district court in Georgia held that a subsidiary's employee was covered because officers of the publicly traded parent had authority to affect the employment of the subsidiary's employees.[6] Similarly, in *Platone v. Atlantic Coast Airlines Holdings Inc.*, an ALJ held that a subsidiary's employee was covered where the company's publicly traded holding company was the alter ego of the subsidiary and had the ability to affect the employee's employment.[7]

Other judges, looking to SOX's legislative intent and purpose of SOX, have found that employees of subsidiaries are almost automatically covered, regardless of the subsidiary's relationship with its parent. The First Circuit, in *Carnero v. Boston Scientific Corp.*, suggested in dicta that an employee of a subsidiary of a publicly traded company could be a covered employee because the subsidiary could be considered an "agent" of the parent.[8] The court opined that "the fact that [complainant] was employed by [the parent's] subsidiaries may be enough to make him a[n] 'employee' [of the parent] for purposes of seeking relief under the whistleblower statute." Likewise, in *Morefield v. Exelon Servs. Inc.*, an ALJ concluded that employees of subsidiaries of publicly traded companies are covered regardless of the parent company's role in affecting the employment of the subsidiary's employees.[9] An ALJ, in *Gonzalez v. Colonial Bank,* agreed with *Morefield* that based on legislative intent a subsidiary's employee was covered.[10]

As these divergent results reflect, in the absence of unambiguous statutory language, a judge's approach to statutory construction plays a significant role in the development of the law.

B. Is a Non-Publicly Traded Subsidiary a Covered Entity?

An inquiry that has resulted in even more mixed results is whether a subsidiary of a publicly traded parent company, standing alone, is a covered entity subject to suit. Section 806 prohibits any publicly traded company or "any officer, employee, contractor, subcontractor or agent of such company" from retaliating against whistle-blowing employees. These terms are not defined in the statute. Whether these terms incorporate subsidiaries appears primarily to depend upon whether the judge applies a strict or liberal construction of the statute.

There appears to be a developing trend of federal courts construing the pertinent statutory provision more strictly than ALJs. For instance, in *Rao v. Daimler Chrysler Corp.*, a district court in Michigan narrowly interpreted Section 806 as not providing a cause of action directly against the subsidiary alone.[11] The court reasoned that "Congress could have specifically included subsidiaries within the purview of § 1514A if they wanted to," and, because they did not, "the general corporate law principle would govern and employees of non-public subsidiaries are not covered under § 1514A." The judge appropriately concluded that "it is not the job of the Court to rewrite clear statutory text."

The strict interpretation of *Rao* is consistent with a series of earlier ALJ decisions holding that Section 806 does not provide a cause of action directly against the subsidiary. For example, in *Ambrose v. U.S. Foodservice, Inc.*, the ALJ dismissed a SOX complaint on the basis that the complainant, an employee of a wholly owned subsidiary of a publicly traded company, was not protected under Section 806.[12] The ALJ reasoned that Section 806's caption, "Whistleblower Protection For Employees Of Publicly Traded Companies," clearly reflected Congress' intent to not extend coverage to employees of non-publicly traded subsidiaries. Similarly, in *Grant v. Dominion East Ohio Gas,* the ALJ concluded that the plain language of Section 806 provides no cause of action against a non-public subsidiary standing alone, regardless of whether complainant could produce evidence to justify piercing the corporate veil.[13] The ALJ reasoned that, even if the complainant could establish that the parent company was liable for the acts of its subsidiary, this "does not cure the deficiency of not naming a company covered by the Act as Respondent. In other words, neither the doctrine of piercing the corporate veil, nor agency law principles generally operate to pull a parent company into litigation if the parent company is not named as a party in the first place."

This strict construction line of ALJ decisions may have reached its conclusion with the Administrative Review Board's ("ARB") 2006 adoption of a more liberal construction in *Klopfenstein v. PCC Flow Technologies Holdings, Inc.*[14] In *Klopfenstein*, the ARB found that a whistleblower claim may proceed directly against a non-publicly traded subsidiary under the theory that the subsidiary is an "agent" of the parent company. The ARB explained that whether a subsidiary is an agent of a publicly traded parent "should be determined according to principles of the general common law of agency." The ARB explained that an agency relationship may be found where there is a manifestation by the principal that the agent shall act for it, the agent's acceptance of the undertaking, and the understanding of the parties that the principal is to be in control. The ARB concluded that commonality of management and involvement by the principal in decisions relating to the complainant's employment were factors weighing in favor of finding existence of an agency relationship.

Unlike the courts, ALJs are bound by *Klopfenstein*. For instance, in *Savastano v. WPP Group, PLC*, an ALJ followed *Klopfenstein*, but narrowed it somewhat by holding that the agency relationship must pertain to employment matters.[15] The ALJ explained that "for an employee of a non-public subsidiary to be covered under Section 806, the non-public subsidiary must act as an agent of its publicly held parent, and the agency must relate to employment matters." In other words, the fact that the companies share an agency relationship for other purposes, such as collecting and reporting financial data, is insufficient to establish subsidiary coverage under SOX.

To further complicate the picture, the Solicitor of Labor has argued for an alternate approach to subsidiary coverage—the four-part "integrated enterprise" test.[16] The "integrated employer" test focuses on (a) interrelation of operations; (b) common management; (c) centralized control of employment decisions; and (d) common ownership or financial control. The integrated enterprise theory finds some support in the case law interpreting Section 806. Likewise, in *Hughart v. Raymond James & Associates, Inc.*,[17] the ALJ (in a pre-*Klopfenstein* decision) suggested that a case under Section 806 may proceed solely against a subsidiary if the parent company and its wholly owned subsidiary are "so intertwined as to represent one entity."[18]

In sum, courts, ALJs and the DOL have all adopted widely differing views regarding whether and under what conditions a subsidiary is covered under SOX. With the courts and the ALJs headed down divergent paths, the unintended practical effect will be that an employee's choice of forum may have a material impact upon whether he/she ultimately can recover under the statute.

C. Does the Existence of Separate Corporate Identities Insulate the Parent from Liability?

The final inquiry—whether the existence of separate corporate identities insulates the parent from liability for acts of the subsidiary—focuses on whether piercing the corporate veil or some other basis for ignoring corporate separateness is warranted so that the parent may be subject to suit. In cases following the ARB's *Klopfenstein* reasoning, this inquiry may now be moot because the complainant usually will be permitted to proceed directly against the subsidiary alone. However, in cases applying the narrower view that non-publicly traded subsidiaries are not directly subject to suit, addressing corporate separateness may often become important.

Federal courts have not yet addressed this issue, but the pre-*Klopfenstein* ALJs who have done so have not agreed upon the proper standard for holding a parent company liable. In *Powers v. Pinnacle Airlines Corp.*, an ALJ dismissed a subsidiary employee's complaint because his attempt to hold the parent liable "ignore[d] the general principle of corporate law that a parent corporation is not liable for the acts of its subsidiaries."[19]

The ALJ, noting that the subsidiary's impact on the parent was "questionable at best," found no basis for piercing the corporate veil. In *Hasan v. J.A. Jones-Lockwood*, an ALJ held that a parent company was not an "employer" under the ERA merely because it was the parent of the employer where no evidence showed that it had the power to hire, promote, discipline or give raises or had input in those decisions.[20]

By contrast, the ALJ in *Platone v. Atlantic Coast Airlines Holdings Inc.* held that the parent/holding company *was* a proper respondent in an action by an employee of a non-publicly traded subsidiary where the subsidiary was a "mere instrumentality" of the holding company.[21] The ALJ reasoned that the holding company had no employees; the companies disregarded the separate identity of the subsidiary in its dealings with the public, the SEC, and its employees; there was a great degree of commonality between the senior management of the two corporate entities, including those responsible for labor relations within the subsidiary; and the holding company had the ability to affect the complainant's employment, including making the ultimate termination decision.

In addition, constitutional and jurisdictional principles may limit courts' efforts to hold parent companies liable for the acts of its subsidiaries. For example, if courts continue to hold that there is no direct cause of action against non-publicly traded subsidiaries, they inevitably will be required to address whether, and to what extent, they have personal jurisdiction over distant publicly traded parent corporations. In *Personalized Brokerage Servs., LLC v. Lucius*, the parent company was a German company with no contacts with the forum state other than the fact that its subsidiary operated there.[22] The court, finding that "there is no case authority that the Act permits a court to dispense with jurisdictional prerequisites in this context," dismissed the plaintiff's Section 806 claim on the ground of lack of personal jurisdiction. The court noted that there was no evidence the subsidiary was an "alter ego" of the parent.[23]

Whether ALJs interpret similar restrictions into the administrative process could impact whether some complainants opt to proceed through the ALJ process rather than run the risk of dismissal by a federal court.

II. Extraterritoriality

Foreign corporations doing business in the United States are subject to Section 806's whistleblower provisions; however, SOX does not express whether its whistleblower protections are intended to extend to employees working outside the United States.[24] Courts and ALJs generally have refused to afford such protections; at least where the complainant is not a U.S. citizen and the employment relationship lacks a substantial nexus with the U.S. To date, it is unclear what degree of nexus will be required to establish coverage for overseas workers.

In *Carnero*, the First Circuit refused to apply Section 806 to a foreign national who was directly employed by Argentinian and Brazilian subsidiaries of a corporation covered by SOX.[25] The court reasoned that Congress was silent as to its intent to apply Section 806 abroad, and it is presumed that federal statutes do not apply extraterritorially absent clear language to extend the statute's protections abroad. However, the court left open the possibility that Section 806 may apply to conduct occurring overseas where a complainant's employment relationship had a more substantial nexus to the U.S.

Similarly, in *Beck v. Citigroup, Inc.*, the ALJ refused to apply Section 806 to a foreign national working exclusively in Germany for respondent's German division.[26] The ALJ reasoned that, although the complainant reported the alleged misconduct to U.S.-based company officials, and although U.S.-based company officials may have participated in the decision to terminate his employment, "the essential nature of the employment relationship" was foreign.[27]

Despite this reluctance to apply SOX's whistleblower provisions overseas, several decisions have acknowledged the potential application of Section 806 to foreign conduct where the complainant's employment relationship has a significant connection to the United States. In *Penesso v. LLC International, Inc.*, although the complainant worked in Italy, the ALJ denied summary decision for the employer, finding "this case has a substantial nexus to the United States, and it is appropriate for the complainant to bring this claim under §1514A of the Sarbanes-Oxley Act."[28] The ALJ reasoned that the complainant was a U.S. citizen, much of the protected activity took place in the U.S. when the complainant came to the company's U.S. headquarters to inform corporate officers of alleged financial improprieties, and at least one of the alleged retaliatory actions took place in the U.S.

In *Neuer v. Bessellieu*, the complainant was not a U.S. citizen, the alleged SOX protected activity occurred in Israel, and the parent company was located in Israel.[29] Nonetheless, the ALJ refused to dismiss the complaint based on lack of extraterritorial effect because the complainant alleged he was employed by a U.S. subsidiary, spent most of his time working in the U.S., that his termination occurred in the U.S., and that there was significant intermingling of the business activities of the Israel-based parent company and its U.S. subsidiary.

The overriding theme of these cases is that, although Section 806 presumably will not apply extraterritorially, it may apply overseas if there is a substantial nexus with the United States. If so, there is a serious concern that SOX's whistleblower protections may conflict with foreign law and cultural norms.[30] For example, the French Data Protection Authority (CNIL) has found that whistleblower hotlines proposed by French subsidiaries of McDonald's and CEAC to comply with SOX Section 301 were in violation of French privacy laws.[31] The CNIL later issued a Guidance explaining that SOX whistleblowing systems are not necessarily prohibited by French law, provided that eleven conditions are satisfied that ensure that "the rights of individuals directly or indirectly incriminated through them are guaranteed with regard to personal data protection rules."[32]

Likewise, a European Union organization recently issued an opinion setting forth a number of conditions which must be met to ensure that internal whistleblowing systems comply with EU data protection rules.[33] In Germany, a Düsseldorf court invalidated Wal-Mart's whistleblowing policy on the ground that it was not first presented to the works council for approval. The court did not address whether the policy violated German privacy laws.[34]

Although these European decisions do not directly address Section 806, they do highlight some of the concerns that might arise if SOX whistleblower provisions are broadly applied to overseas employees.

III. Protected Activity

The issue of what constitutes SOX protected activity has involved difficult questions, such as when does an employee have a "reasonable belief" that conduct violates the predicate fraud or securities provisions, how particularly must the employee articulate a violation of those provisions, and must the alleged wrongdoing have an impact on shareholders? Understandably, OSHA investigators (and employment lawyers) with little background in the vagaries of securities law have been wary of tackling these issues head-on.

A. Is an Allegation of Fraud Against Shareholders Required?

To constitute protected activity, the subject matter of a SOX complaint must involve a purported violation of federal mail, wire, bank or securities fraud laws, any SEC rule or regulation, or "any provision of Federal law relating to fraud against shareholders." A conflict exists among the various courts and ALJs that have addressed the issue as to whether this statutory provision limits protected activity under SOX to complaints alleging fraud "against shareholders." Some judges have adopted a narrow view that "[t]o be protected under Sarbanes-Oxley, an employee's disclosures must be related to illegal activity that, at its core, involves shareholder fraud."[35]

Other judges have applied a broader interpretation which appears more consistent with the statute's plain meaning, that complaints of mail, wire, bank or securities fraud, or violation of any SEC rule or regulation need not relate to fraud against shareholders.[36] This broader approach appears to be developing into the prevailing view. For instance, in *Reyna v. ConAgra Foods, Inc.*, a district court in Georgia found that "[t]he statute clearly protects an employee against retaliation based upon that employee's reporting of mail fraud or wire fraud regardless of whether that fraud involves a shareholder of the company."[37] The court concluded that when a complaint involves noncompliance with internal accounting controls promulgated in compliance with Sarbanes-Oxley mandates or SEC rules and regulations, the non-compliance itself would constitute a violation of one of the laws enumerated in section 806.[38]

Another area of disagreement has been whether a complaint must allege intentional deceit as an element of protected activity. In *Hopkins v. ATK Tactical Systems*, an ALJ found that a complaint was not protected where it did not allege that the company's activities involved intentional deceit.[39] The ALJ reasoned that "an element of intentional deceit that would impact shareholders or investors is implicit" under the SOX whistleblower provision. Similarly, in *Getman v. Southwest Securities, Inc.*, an ALJ noted that the scienter requirement under the "any rule or regulation of the SEC" provision, specifically Rule 10b-5, requires "a mental state embracing intent to deceive, manipulate, or defraud," and is established by showing that the respondent acted intentionally or with severe recklessness.[40]

In contrast, other judges have determined that a complaint need not implicate intentional deceit as long as a violation of SEC rule or regulation is raised. For example, in *Smith v. Corning, Inc.*, a federal district court found that the submission of quarterly reports that simply were not prepared in accordance with General Accepted Accounting Principles (GAAP) would violate a "rule or regulation of the Securities and Exchange Commission," and therefore an employee who alleged that the defendants refused to address a problem that was resulting in incorrect financial information being reported to the company's general ledger sufficiently alleged protected activity under SOX.[41] Likewise, in *Morefield v. Exelon Servs. Inc.*, an ALJ concluded that the catchall, "any provision of Federal law relating to fraud against shareholders" provision "may provide ample latitude to include rules governing the application of accounting principles and the adequacy of internal accounting controls implemented by the publicly traded company in compliance with such rules and regulations."[42]

In short, in light of the purposes of SOX, one might reasonably conclude that some level of fraud and intentional deceit relating to shareholders is required for a complaint to be protected. However, the plain statutory language does not appear to support such an interpretation.

B. Must the alleged wrongdoing be material?

Taken to a logical extreme, almost any wrongdoing by a publicly traded company, no matter how insignificant, and regardless of the existence of fraud, may have some degree of impact on its shareholders, and therefore could implicate SOX. To date, judges have struggled with the issue of whether, and to what degree, employee complaints must implicate material wrongdoing.

Materiality is an element of the predicate fraud provisions.[43] In addition, ALJs have applied a materiality element under SOX's "any rule or regulation of the SEC" and "any provision of Federal law relating to fraud against shareholders" provisions. However, some ALJs have not applied a materiality requirement. For example, in *Morefield*, the ALJ placed little emphasis on the materiality requirement under the catchall "any provision of Federal law relating to fraud against shareholders," and denied respondent's motion to dismiss despite the fact that the amounts involved totaled less than .0001% of the annual revenues of the parent company.[44] The ALJ reasoned that whether or not "materiality" is a required element of a criminal fraud conviction "we need be mindful that Sarbanes-Oxley is largely a prophylactic, not a punitive measure."

In contrast, judges appear to apply a heightened materiality requirement when the employee does not complain directly about accounting or security-related matters. For instance, several complainants have argued that their complaints about violations of other employment laws were protected under SOX because the potential liability could have an adverse impact on shareholders. Judges have accepted this argument in theory, but to date have not found a sufficient level of materiality or nexus to fraud against shareholders to warrant protection.

One such case is *Harvey v. Home Depot, Inc.*, in which an ALJ concluded that an employee complaint about alleged race discrimination that had "a very marginal connection with" a corporation's accurate accounting and financial condition did not constitute activity protected under SOX.[45] The ALJ recognized that a company's discriminatory practices can

implicate federal law relating to fraud against shareholders in that it is acting contrary to the best interests of its shareholders, but rejected this argument because its nexus to fraud against shareholders was extremely tenuous. The only applicable federal law that could have possibly been implicated was SOX itself (which requires certification that a financial disclosure is accurate and does not contain any untrue statement of material fact). However, "the connection between the incident of discrimination and the accuracy of corporate disclosures was "tenuous upon close examination of SOX." The ALJ noted that the discrimination complaints at issue centered on the alleged *existence* of discrimination, not the company's *failure to report* such discrimination to the public. Nevertheless, the ALJ suggested that perhaps the failure to disclose a class action discrimination lawsuit might become the subject of a SOX protected activity "if an individual complained about the failure to disclose that situation."

Similarly, in *Harvey v. Safeway, Inc*., 2004-SOX-21 (ALJ Feb. 11, 2005), the complainant complained that discrepancies in his weekly paychecks violated the FLSA. In his subsequent SOX complaint, the complainant argued his reports of FLSA violations constituted protected activity. The ALJ rejected this argument based on a lack of materiality, explaining that the employee's "personal experience over the course of a couple of weeks with Safeway and an anecdotal report of one other employee's wage concerns did not provide an objectively reasonable factual foundation for a... complaint about systematic wage underpayment." Implicit in this reasoning is that had the company actually engaged in systemic wage underpayment and not reported it to its shareholders, the employee's complaint may have been protected.

Likewise, in *Smith v. Hewlett Packard*, an employee alleged that his threats to take allegations of a potential race discrimination class action to the EEOC constituted protected activity under SOX.[46] The ALJ rejected this argument, reasoning that "[m]ere knowledge that an employee-evaluation process adversely affected minorities (without knowing whether this result was intentional), coupled with an insider's access to disgruntled employees' conversations about 'external' resolutions, is not enough." The ALJ noted that a rumor of a class-action lawsuit, absent such litigation, is not something the company must disclose to its shareholders. The ALJ did state, however, that disclosure of company-wide discrimination could form the basis of SOX whistleblower claim if "such a suit actually been filed, and if HP had prevented that information from reaching its shareholders, and if the Complainant learned of this omission and if he had reported it."

C. How Have Judges Interpreted "Reasonable Belief"?

An employee must "reasonably believe" the conduct at issue constitutes a violation of the predicate fraud or securities provisions. SOX does not define "reasonable belief," and it has not been clear how the DOL and courts would interpret this phrase under SOX. However, the prevailing view now appears to be that this term should be interpreted consistent with the "reasonable belief" standard commonly applied under other employment-related statutes, which contains both a subjective and objective component.

An interesting effect of the reasonable belief analysis in the SOX context is that judges have evaluated reasonableness based in part on the level of the complaining employee's knowledge or expertise regarding the matters about which he/she complains. Therefore, employees with little or no expertise in accounting, for example, have been held to a very low standard, while employees with extensive pertinent knowledge or expertise are held to a higher standard. For example, in *Grove v. EMC Corp*., an ALJ found that, although the company did not engage in any fraudulent or wrongful activities enumerated in Section 806, because the complainant was a salesman with no specialized training or expertise in the area of corporate acquisitions, it was not unreasonable for a person in the complainant's position to believe that the company's actions impacted the company's financial condition.[47]

In contrast, in *Welch v. Cardinal Bankshares Corp*., the company's CFO, an experienced CPA, complained about figures contained in the company's SEC financial reports.[48] The ARB found that because the complainant was an experienced CPA/CFO, he could not have reasonably believed that errors in the company's quarterly SEC report presented a misleading picture of the company's financial condition.

A corollary to this trend is that employers may argue, understandably, that an employee with little or no experience in accounting or corporate fraud matters may not have a *subjective* belief of a violation of SOX's predicate fraud provisions because the employee lacks familiarity with the requisite elements under those provisions. To date, the response to this argument appears to have been that employees with less pertinent expertise or knowledge will be held to a lower standard.

D. How Particular Must the Employee's Complaint Be?

Another issue upon which the statute is silent is with what level of particularity must an employee articulate in the complaint a violation of the predicate fraud or securities provisions? Thus far, judges have reached mixed results on this issue.

Some judges have interpreted the statute as requiring very specific allegations of fraud of securities violations. For example, in *Van Asdale v. International Game Technology*, a federal district court explained that, to constitute protected activity, "an employee's act must implicate securities fraud definitively and specifically.... Thus, the whistleblowing cannot be vague...."[49] The court found that plaintiff's recommendation that the company "investigate these issues, the potential for fraud" did not rise to the level of protected activity because the court could not "infer that [plaintiff] implied that shareholder fraud had occurred and that [defendant] understood the implication."

Likewise, in *Lerbs v. Buca Di Beppo, Inc*., an ALJ found that the complainant failed to show that he engaged in protected activity because his alleged complaints did not state a particular concern about the company's practices.[50] Rather, he simply asked about certain entries in a general ledger and on another occasion allegedly told an officer that he thought an entry was misleading. The ALJ found that these remarks were more like general inquiries which are not protected under SOX.

Rejecting such an exacting standard, a federal district court in *Collins v. Beazer Homes USA, Inc*., found a genuine

issue of material fact as to whether the plaintiff engaged in protected activity where she contended that her complaints to management disclosed attempts to circumvent the company's system of internal accounting controls in violation of Section 13 of the Exchange Act.[51] The court rejected that the complaints were too vague, noting that the company took the allegations seriously enough to investigate. Moreover, the court concluded that "the mere fact that the severity or specificity of her complaints does not rise to the level of action that would spur Congress to draft legislation does not mean that the legislation it did draft was not meant to protect her."

IV. Procedural Issues

Under the SOX administrative scheme, a complainant must exhaust certain administrative requirements prior to initiating adjudicatory proceedings. However, SOX contains a "kick out" provision that, in most cases, allows the complainant to bring a *de novo* action in district court if the DOL does not issue a final decision within 180 days. In deciding between federal court or remaining before the DOL OALJ, a complainant must take into consideration a number of factors, including the possibility of fewer evidentiary restrictions, less formal pleading requirements in agency adjudications, and the likelihood that one forum may interpret ambiguous statutory language more favorably than another. The trend appears to be that federal courts are more likely than ALJs to strictly construe administrative exhaustion and pleading requirements.

A. May a Complainant Add Claims After the Initial OSHA Determination?

One procedural issue that has led to divergent construction is whether a complainant is permitted to add claims after OSHA issues its initial determination. The federal courts have held that a SOX complaint filed in federal court after the expiration of 180 days generally must be limited to the claims identified in the initial OSHA complaint. For example, in *Willis v. Vie Financial Group, Inc.*, the court held that the SOX administrative exhaustion requirement precluded recovery for a discrete act of retaliation which was never presented to OSHA for investigation.[52] The court reasoned that the SOX administrative scheme, unlike that of Title VII, "is judicial in nature and is designed to resolve the controversy on its merits…."

One issue that was not presented to the *Willis* court, however, was whether a complainant should be permitted to add claims, which were not before OSHA during the investigation, in *an ALJ proceeding* after OSHA issued its initial determination. ALJs have been inconsistent on this issue. Reading the statute narrowly, the ALJ in *Ford v. Northwest Airlines, Inc.*, reasoned that although the substance of the new claims was based on the same core of operative facts, OSHA was not given the opportunity to investigate the allegations "under the two-tiered scheme Congress provided for handling whistleblower claims."[53] The ALJ concluded:

> I will not arbitrarily usurp the system established by Congress and determine the legitimacy of this allegation in the first instance. A better procedure is to make the initial complaint to OSHA and then move to consolidate the complaint with litigation pending before the OALJ.[54]

In contrast, in *Hooker v. Westinghouse Savannah River Co.*, the complainant failed to allege his refusal to rehire claim in his initial ERA discrimination complaint.[55] The ALJ sua sponte amended the complaint to include the refusal to rehire allegation. On review, the ARB did not contest the sua sponte amendment, but explained that the proper procedure for amending complaints is found at 29 C.F.R. § 18.5(e).

B. May a Complainant Add Parties After the Initial OSHA Determination?

Another procedural issue that has led to divergent construction is whether a complainant is permitted to add parties after OSHA issues its initial determination. Federal courts generally have rejected plaintiffs' efforts to add new defendants who were not named in the initial OSHA complaint.

For example, in *Smith v. Corning, Inc.*, the employee named the company but not his supervisor as a respondent in his OSHA complaint.[56] He then filed a complaint in federal court naming both the company and his supervisor as defendants. The court dismissed the claims against the supervisor, finding that plaintiff failed to exhaust his administrative remedies as against the supervisor. Likewise, in *Bozeman v. Per-Se Techs., Inc.*, the court dismissed claims against individual defendants who were not named in the OSHA proceedings.[57] The court reasoned that, "[w]hile the regulations implementing SOX may provide for individual liability, that does not obviate the need for the Plaintiff to exhaust his administrative remedies for each claim he seeks to assert against each defendant."[58]

Complainants' attempts to add new respondents during an *ALJ proceeding* subsequent to an initial determination by OSHA have met with much friendlier results. For instance, in *Gallagher v. Granada Entertainment USA*, the ALJ, citing no authority, stated that "[i]ndividuals and entities may be added as parties when they were not joined below through error."[59] The ALJ permitted the complainant to add as respondents the individual executives of the named corporate respondent, named as those who terminated the complainant's employment. Although the ALJ observed that the initial OSHA complaint is "not a pleading under Rule 8(a), Fed.R.Civ.P., but a complaint in the ordinary sense," the ALJ did not reconcile this observation with 29 CFR § 18.5(e), which only grants the ALJ discretion to permit amendments to "complaints, answers and other pleadings, as defined by the Rules."

Likewise, in *Gonzalez v. Colonial Bank*, the ALJ, citing 29 C.F.R. § 18.5(e), permitted complainant to amend his initial OSHA complaint to include the parent as a respondent.[60] Further, the ALJ permitted the amendment to relate back to the date of the initial OSHA complaint, thereby rendering the claims against the parent timely. The ALJ reasoned that, although the complainant was aware of the identity and role of the parent from the outset, "amending the complaint filed before OSHA by adding... the parent company... as a respondent comports with the purpose of Rule 15(c) and the purpose of the Act." This decision illustrates why a complainant might choose to pursue agency adjudication rather than removing to federal district court. For example, if the complainant in *Gonzalez* had removed to federal court, the court, consistent with the reasoning in *Willis* and *Smith*, likely would have held that SOX's administrative exhaustion requirement precluded addition of the parent as a defendant. Moreover, in federal court, the OSHA administrative complaint would not have been

subject to amendment under Rule 15(a), Fed.R.Civ.P. Although the OALJ Rules of Practice would appear consistent with the federal rules on this issue, the ALJ was willing to interpret the rules more liberally than a federal judge likely would have, in light of the statutory purpose.

V. Section 1107 Criminal Provision

Section 1107 makes it a crime to knowingly and intentionally retaliate against any person who provides truthful information to a law enforcement officer relating to the commission or possible commission of any federal offense.[61] Criminal sanctions include, for individuals, fines up to $250,000 and/or imprisonment up to ten years and, for organizations, fines up to $500,000.[62]

A. Potentially Expansive Criminal Liability under Section 1107

To date, Section 1107 has not been extensively litigated. However, there are several aspects of Section 1107 that could be construed liberally to result in its extremely broad application. Section 1107 applies not only to publicly traded companies, but to *any* "person," meaning employers, supervisors and other employees may be criminally liable for retaliatory conduct. Employers are covered regardless of their corporate status or number of employees. Moreover, Section 1107 coverage is not limited to the employment relationship, therefore third parties, regardless of their agency relationship with the employer, may be liable.

In addition, Section 1107 is not limited to employees reporting fraud or securities violations. In fact, Section 1107 covers disclosures to *any* federal law enforcement officer relating to commission or possible commission of any federal offense. Because the term "law enforcement officer" is broadly defined as any federal officer or employee "authorized under law to engage in or supervise the prevention, detection, investigation, or prosecution of an offense," this provision reasonably could be interpreted as encompassing complaints to the EEOC under federal employment discrimination statutes such as Title VII, ADA or ADEA or complaints under various other labor statutes such as NLRA, OSHA, FLSA, etc.[63] Whether such an interpretation is adopted hinges largely on the meaning of the term "federal offense," which is not defined in SOX, but which has been applied in both civil and criminal contexts.[64]

Finally, the conduct prohibited by Section 1107 is extremely broad, covering any action "harmful" to a person, including "interference with the lawful employment or livelihood" of any person. Moreover, an employee need not report an actual violation, as long as his/her disclosure is "truthful" and relates to the "possible commission" of a federal offense. Because there is no statutory language limiting the terms "harmful" or "interference" to injuries involving economic harm, or even to retaliation occurring within the scope of the employment relationship, the scope of prohibited conduct under Section 1107 is probably at least as broad as conduct prohibited under the hostile work environment theory under other employment statutes.

Due to Section 1107's broad scope, it would not be unreasonable to interpret Section 1107 as criminalizing retaliatory conduct that previously only would have given rise to civil liability under Title VII or other federal employment statutes. In *MacArthur v. San Juan County*, plaintiffs contended they suffered retaliation in violation of Section 1107 for having informed their employer/hospital's governance board of ethnic remarks made by hospital administration concerning another employee.[65] The court noted that Section 1107 "simply cannot be read to reach the reporting of ethnic remarks to a local hospital's governance board." The court did not address, however, whether such reports would have been covered if they instead had been made to the EEOC. The court also did not address whether a private cause of action even exists under Section 1107, although every court to date to address this issue has held that it does not.[66]

B. SOX Violations Could Give Rise to Civil RICO Claims

Retaliation against corporate whistleblowers may give rise to a cause of action under the civil RICO statute. Section 1107 amends 18 U.S.C. § 1513(e) and, under RICO, "racketeering" includes "any act which is indictable under... 18 U.S.C. § 1513."[67] Therefore, by engaging in retaliation prohibited by Section 1107 (*e.g.,* conceivably by engaging in ongoing retaliatory acts or creating a hostile work environment), a company or person commits a predicate act of racketeering under RICO.

With the availability of treble damages, plaintiffs have an incentive to pursue civil RICO claims. Prior to the enactment of Section 1107, civil RICO claims arising from retaliatory discharge rarely succeeded because retaliatory discharge did not fall within the definition of "racketeering."[68] Even if an employee could legitimately allege that the employer committed a predicate act under RICO, the employee rarely could assert a viable RICO claim because the employee's injury was almost never proximately caused by the predicate act, but rather by a separate adverse employment action.[69] Section 1107, by including retaliatory conduct as a predicate act, significantly expands the likelihood of establishing the necessary causal link between the predicate act and the injury.

Of course, courts have been reluctant to expand civil RICO coverage, and a plaintiff must still establish the other civil RICO elements, such as existence of an enterprise and a "pattern of racketeering." For example, in *Compact Disc Minimum Advertised Price Antitrust Litigation*, plaintiff alleged he was fired in retaliation for conveying truthful information during a federal investigation, and that his firing constituted a predicate act for purposes of RICO under 18 U.S.C. § 1513.[70] The court dismissed plaintiff's RICO claims because he failed to establish a "pattern of racketeering." The court reasoned that the single act of termination, combined with alleged instructions by the employer to withhold information during "a single federal investigation," did not constitute a "pattern" sufficient to support a RICO claim. It remains to be seen whether an ongoing hostile work environment or systematic retaliation (as opposed to a single adverse employment action) may give rise to a "pattern of racketeering" under RICO.

Therefore, although a plaintiff alleging a civil RICO claim arising for a Section 1107 violation faces significant hurdles, a broad interpretation of the statute could provide a basis for such claims.

C. Potential SEC Criminal Issues

Beyond Section 1107, Section 3(b) of SOX states that "a violation by any person of th[e Sarbanes-Oxley] Act... shall be treated for all purposes in the same manner as a violation of the Securities Exchange Act of 1934 (15 U.S.C. 78a et seq.)... and any such person shall be subject to the same penalties, and to the same extent, as for a violation of that Act or such rules or regulations." Due to this broad and ambiguous language, Section 3(b) could reasonably be interpreted as expanding *criminal* liability for any retaliatory action prohibited by Section 806, regardless of whether the retaliation was related to the disclosure of truthful information to a law enforcement officer.

Indeed, some in Congress have pushed for such a broad interpretation. In November 2004, Senators Grassley and Leahy sent a letter to SEC Chairman William Donaldson stating that they wanted "aggressive enforcement to deter retaliation against corporate whistleblowers," asking: "[w]hat is your position on whether or not a violation of the Section 806 whistleblower prohibitions can generate criminal liability under Section 3(d) [sic] of the Act?" Chairman Donaldson responded to the effect that, while Section 3(b) is a useful provision allowing the SEC to enforce SOX, the SEC would leave it to the Labor Department to investigate and prosecute potential Section 806 whistleblower violations.[71]

Although to date the SEC has not pursued Section 3(b) actions in this context, the broad statutory language could allow future administrations to adopt a broader interpretation of this provision. Moreover, even if Section 3(b) is not interpreted as criminalizing retaliation prohibited by Section 806, all Section 806 complaints are referred to the SEC and could give rise to prosecution for substantive violations of the securities laws. For example, in *Matter of Ashland Inc.*, an employee filed a complaint with DOL alleging retaliation for raising concerns about understatement of the company's environmental reserves.[72] Upon referral from the DOL, the SEC instituted proceedings against the company, ultimately finding the company violated the reporting, books and records, and internal controls provisions of the Exchange Act. The company and the SEC settled the matter.

CONCLUSION

Ambiguities in the statutory language of SOX's whistleblower provisions have forced judges to resolve important issues under the law on a case-by-case basis through statutory construction. Judges' contrasting approaches to statutory construction have led to divergent results, which has exacerbated the lack of clarity in the law and has created a situation where parties' choice of forum may materially impact their substantive rights under the law. Additionally, ambiguities in Section 1107's criminal whistleblower provision have created uncertainty regarding its scope. A broad interpretation of Section 1107 could result in serious criminal and civil liability well beyond its intended scope. Finally, the expansive reach of SOX's whistleblower protections has created extraterritoriality and jurisdictional concerns. A consistent, strict construction of SOX's whistleblower provisions would minimize conflicts and create a greater degree of uniformity in this evolving area of law.

Endnotes

1 Lewis Carroll, Through the Looking-Glass, Ch. 6 (1872).

2 J. Gregory Grisham & James H. Stock, Jr. 8 Engage 1, 29, 33 (Feb. 2007).

3 It has been said that SOX was "rushed into law in the hysterical atmosphere surrounding the Enron and WorldCom bankruptcies." *See* 151 Cong. Rec. E657 (daily ed. Apr. 15, 2005) (statement by Rep. Ron Paul). The bill, H.R. 3763, passed the House by a vote of 423-3 and passed the Senate unanimously. *See* Final Vote Results For Roll Call 348, *available at* <http://clerk.house.gov/evs/2002/roll348.xml> and U.S. Senate Roll Call Votes 107th Congress, *available at* <http://www.senate.gov/legislative/LIS/roll_call_lists/roll_call_vote_cfm.cfm?congress=107&session=2&vote=00192>.

4 SOX's whistleblower provisions were expressly intended to correct what some in Congress perceived as a "patchwork and vagaries" of state laws relating to corporate whistleblowers. *See* 148 Cong. Rec. S7420 (daily ed. July 26, 2002) (statement of Senator Leahy). *See also* Roberta Romano, *The Sarbanes-Oxley Act and the Making of Quack Corporate Governance*, 114 Yale L.J. 1521, 1523 (May 2005).

5 18 U.S.C. § 1514A(a).

6 334 F. Supp. 2d 1365 (N.D. Ga. 2004).

7 2003-SOX-27 (ALJ Apr. 30, 2004).

8 433 F.3d 1 (1st Cir. 2006).

9 2004-SOX-2 (ALJ Jan. 28, 2004).

10 2004-SOX-39 (ALJ Aug. 20, 2004).

11 2007 U.S. Dist. LEXIS 34922 (E.D. Mich. May 14, 2007).

12 2005-SOX-105 (ALJ Apr. 17, 2006).

13 2004-SOX-63 (ALJ Mar. 10, 2005).

14 ARB 04-149, 2004-SOX-11 (ARB May 31, 2006).

15 2007-SOX-34 (ALJ July 18, 2007).

16 *See* Brief of the Assistant Secretary of Labor for Occupational Safety and Health, Ambrose v. U.S. Foodservice, Inc., ARB 06-096, 2005-SOX-105 (brief filed Sept. 1, 2006).

17 2004-SOX-9 (ALJ Dec. 17, 2004).

18 *See also* O'Keefe v. TIAA-CREF, 2005-SOX-86 (ALJ Oct. 19, 2005) (denying respondent's motion for summary decision because a question of fact existed as to whether "the high degree of intermingling between TIAA and its subsidiaries, between TIAA and CREF and the presentation of TIAA-CREF as one company to its customers, employees and the public" should result in TIAA-CREF's treatment as one publicly traded entity under the Act).

19 2003-AIR-12 (ALJ Mar. 5, 2003).

20 2002-ERA-5 (ALJ Sept. 17, 2002).

21 2003-SOX-27 (ALJ Apr. 30, 2004).

22 2006 U.S. Dist. LEXIS 75225 (D. Minn. Oct. 16, 2006).

23 *See also* Andrews v. ING North America Insurance Corp., 2005-SOX-50 (ALJ Feb. 17, 2006) (dismissing complaint because complainant could not proceed directly against the subsidiary and complainant did not attempt to add the publicly traded parent because of inability to sue the parent due to its foreign status).

24 Carnero v. Boston Sci. Corp., 433 F.3d 1 (1st Cir.) ("These registration and reporting provisions apply to U.S. and foreign companies listed on U.S. securities exchanges") *cert. denied*, 126 S. Ct. 2973 (2006); Ward v. W & H Voortman, Ltd., 685 F. Supp. 231, 232 (M.D. Ala. 1988).

25 433 F.3d 1 (1st Cir. 2006).

26 2006-SOX-3 (ALJ Aug. 1, 2006).

27 *See also* O'Mahony v. Accenture Ltd., 2005-SOX-72 (ALJ Jan. 20, 2006); Ede v. Swatch Group, 2004-SOX-68, 2004-SOX-69 (ALJ Jan. 14, 2005); Concone v. Capital One Fin. Corp., 2005-SOX-6 (ALJ Dec. 3, 2004).

28 2005-SOX-16 (ALJ Mar. 4, 2005).

29 2006-SOX-132 (ALJ Dec. 5, 2006).

30 *See* 69 Fed. Reg. 52104, 52105 (Aug. 24, 2004).

31 *See* Decisions of the Commission nationale de l'informatique et des libertés (May 26, 2005), *available at* <http://www.cnil.fr/index.php?id=1834> and <http://www.cnil.fr/index.php?id=1833>.

32 *See* Guideline document of the Commission nationale de l'informatique et des libertés (Nov. 10, 2005), *available at* http://www.cnil.fr/fileadmin/documents/uk/CNIL-recommandations-whistleblowing-VA.pdf.

33 *Available at* http://europa.eu.int/comm/justice_home/fsj/privacy/docs/wpdocs/2006/wp117_en.pdf.

34 *See* Decision of Landesarbeitsgericht, 10 TaBV 46/05 (Nov. 14, 2005).

35 Livingston v. Wyeth, Inc., 2006 U.S. Dist. LEXIS 52978 (M.D.N.C. July 28, 2006). *See also* Bishop v. PCS Admin. (USA), Inc., 2006 U.S. Dist. LEXIS 37230 (N.D. Ill. May 23, 2006) (finding that the phrase "relating to fraud against shareholder" must be read as applying to all violations enumerated under section 806); Marshall v. Northrup, 2005-SOX-0008 (ALJ June 22, 2005); Wengender v. Robert Half Int'l Inc., 2005-SOX-59 (ALJ March 30, 2006).

36 Collins v. Beazer Homes USA, Inc., 334 F. Supp. 2d 1365 (N.D. Ga. 2004) ("The threshold is intended to include all good faith and reasonable reporting of fraud"); Walton v. Nova Inf. Sys., 2005-SOX-1076; 2006-SOX-18 (ALJ March 29, 2006) (rejecting argument that report of violation or rule or regulation of the SEC must relate to fraud against shareholders).

37 2007 U.S. Dist. LEXIS 42112 (M.D. Ga. June 11, 2007).

38 *See also* Johnson v. Stein Mart, Inc., 26 I.E.R. Cas. (BNA) 470 (M.D. Fla. June 20, 2007) (allegations of accounting improprieties even in absence of fraud allegations constituted protected activity); Mahony v. KeySpan Corp., 2007 U.S. Dist. LEXIS 22042 (E.D.N.Y. Mar. 12, 2007) (finding plaintiff reasonably believed that the company was engaging in accounting practices that needed to be corrected before its financial statements misled shareholders).

39 2004-SOX-19 (ALJ May 27, 2004).

40 2003-SOX-8 (ALJ Feb. 2, 2004).

41 2007 U.S. Dist. LEXIS 52773 (W.D.N.Y. July 9, 2007).

42 2004-SOX-2 (ALJ Jan. 28, 2004).

43 *See, e.g.,* Neder v. United States, 527 U.S. 1, 4 (1999).

44 2004-SOX-2 (ALJ Jan. 28, 2004).

45 2004-SOX-20 (ALJ May 28, 2004).

46 2005-SOX-88 (ALJ Jan. 19, 2006).

47 2006-SOX-99 (ALJ July 2, 2007).

48 ARB No. 05 064, ALJ No. 2003-SOX-15 (ARB May 31, 2007).

49 2007 U.S. Dist. LEXIS 53867 (D. Nev. June 13, 2007).

50 2004-SOX-8 (ALJ June 15, 2004).

51 334 F. Supp. 2d 1365 (N.D. Ga. 2004).

52 2004 U.S. Dist. LEXIS 15753 (E.D. Pa. Aug. 6, 2004).

53 2002-AIR-21 (ALJ Oct. 18, 2002).

54 *See also* Kingoff v. Maxim Group LLC, 2004-SOX-57 (ALJ July 21, 2004) (denying complainant's attempt to add constructive discharge claims after OSHA issued its initial determination).

55 ARB No. 03-036, ALJ No. 2001-ERA-16 (ARB Aug. 26, 2004).

56 2007 U.S. Dist. LEXIS 52958 (W.D.N.Y. July 23, 2007).

57 2006 U.S. Dist. LEXIS 77885 (N.D. Ga. Sept. 12, 2006).

58 *Accord,* Hanna v. WCI Communities, Inc., 2004 U.S. Dist. LEXIS 25652 (S.D. Fla. Nov. 15, 2004) (dismissing SOX claim against individual defendant not named as respondents in plaintiff's OSHA complaint).

59 2004-SOX-74 (ALJ Oct. 19, 2004).

60 2004-SOX-39 (ALJ Aug. 9, 2004).

61 *See* 18 U.S.C. § 1513(e).

62 *See* 18 U.S.C. § 3571.

63 18 U.S.C. § 1515(a)(4).

64 *See, e.g.,* Cole v. United States Dept. of Agric., 133 F.3d 803 (11th Cir. 1998) (referring to "criminal and civil offenses").

65 2005 U.S. Dist. LEXIS 25235 (D. Utah June 13, 2005).

66 *See* Compact Disc Minimum Advertised Price Antitrust Litig., No. 05-cv-118 (D. Conn. Sept. 25, 2006); Deep v. Recording Indus. Ass'n of Am., No. 2:05-cv-00118 (D. Me. Oct. 2, 2006); Fraser v. Fiduciary Trust Co. Int'l, No. 04 Civ. 6958 (S.D.N.Y. June 23, 2005).

67 *See* 18 U.S.C. § 1961.

68 *See* Beck v. Prupis, 529 U.S. 494 (2000).

69 *See* Sedima, S.P.R.L. v. Imrex Co., Inc., 473 U.S. 479 (1985); Miranda v. Ponce Fed. Bank, 948 F.2d 41 (1st Cir. 1991).

70 No. 05-cv-118 (D. Conn. Sept. 25, 2006).

71 *See* James Hamilton, *SEC Responds to Senate Letter on Whistleblower Provisions*, 2005-32 SEC Today Online (CCH) (Feb. 17, 2005).

72 No. 3-12487 (SEC Nov. 29, 2006).

Litigation

New Opportunities for Defendants in Securities Class Actions

*By Francis J. Menton, Jr.**

The securities class action litigation industry, known for its stability the last twenty or so years, could see some real changes soon. Several developments in the past year have given defendants opportunities previously unavailable to deliver potentially fatal blows to some of the many weak cases brought every year. These include (1) the Supreme Court's decision in *Tellabs, Inc. v. Makor Issues & Rights, Ltd.*,[1] prescribing how to apply the "strong inference of scienter" requirement of the Private Securities Litigation Reform Act of 1995 ("Reform Act"); (2) decisions from the Second and Fifth Circuits, making class certification more difficult in what could be large groups of cases; and (3) consideration by the SEC of allowing issuers to provide, in their by-laws or otherwise, for arbitration of securities disputes with shareholders.

No one of these developments is itself revolutionary, and none will represent the death knell of securities class actions as we know them. Yet each provides opportunities to defendants to attempt to win, or fatally wound, at least some of the cases that would previously have been settled. Together, these developments could represent a meaningful movement away from the situation where the large majority of securities class actions, regardless of merit, reap large settlements.

Background: The Status of Securities Class Action Litigation in the United States up to 2006

The significance of these developments must be evaluated against the background from which they arose. Securities class action litigation over the last two decades has become a sophisticated financial game: very lucrative to its players, a net loss to its supposed beneficiaries (allegedly defrauded shareholders), and maddening to many business people caught up in its processes. Filings of securities class actions in the U.S. over this period have averaged about 200 per year. With about 12,000 publicly-traded companies in the U.S.(and giving some allowance for overlapping cases filed against the same company), this figure represents a probability of about 1.5% per company per year of drawing a class action lawsuit, or about 7.5% in five years for any given company.

Securities class actions have long been notoriously difficult to get dismissed on pre-trial motion, whether a motion to dismiss directed to the complaint or a summary judgment motion. Most federal trial and appeal courts have simply viewed it as the role of the jury, not the judge, to determine if the corporate disclosures made to shareholders met the standards of the federal securities laws. Congress' passage of the Reform Act in 1995 reflected intent to encourage the federal courts to exercise some gatekeeper function on the flow of cases, but after a decade of experience only a relatively small part of the potential of the Act has been realized. According to one study, the rate at which cases get dismissed at the pleading stage rose nationwide from 19.4% in 1991-1995 to 38.2% in 2000-2004.[2] However, the dismissal rate in the Second Circuit, which includes New York, remains at about the 20% pre-Reform Act level, a fact that just means that filings are increasingly concentrated in that jurisdiction. Moreover, once they have survived pre-trial motions, almost no securities class actions have gone to trial, and virtually all are settled. A study available on the American Bar Association's website reports that 99.5% of the cases in the 2002-2007 period were settled, meaning that of the 200 or so filed each year, only about one actually went to trial.[3]

The typical case begins with the company's stock having a significant and sudden drop in the market, generally several points on a single day. With only this to go on, several plaintiffs' law firms prepare complaints within a few days, alleging that whatever information about the company that came out that day was known to insiders for months. For companies of any substantial size and market capitalization, the damages claimed will quickly be huge. For example, if the stock dropped two points, trading averaged 500,000 shares per day, and the plaintiff alleges a class period of six months, the notional damages claim will be in the range of $150 million. Where trading is in the millions of shares per day, and class periods get to a year and more, notional damages quickly get into the many billions of dollars.

The notional damages take a step closer to reality when a court certifies a class of all shareholders who bought within an alleged class period. The numerous pre-requisites for class certification in Rule 23 of the Federal Rules of Civil Procedure may appear to make class certification a significant hurdle, but two Supreme Court decisions in the 1970s and 1980s swept away the hurdle and made most class certifications of securities actions a foregone conclusion. In *Eisen v. Carlisle & Jacquelin*, the Court made the frequently cited statement that "We find nothing in either the language or history of Rule 23 that gives a court any authority to conduct a preliminary inquiry into the merits of a suit in order to determine whether it may be maintained as a class action."[4] Courts around the country applied this prescription somewhat differently, with the most heavily used forum, New York's Second Circuit, leaning strongly against allowing meaningful factual inquiry at the class certification stage. Then, in *Basic, Inc. v. Levinson*, the Supreme Court accepted the notion that there should be a "rebuttable presumption" that the reliance element in securities fraud was common to all purchasers of a security because the purchasers relied on the market price as embodying all information about the security.[5]

The availability of Directors and Officers insurance against securities claims has only strengthened the already strong incentives toward settlement of securities class actions. In most circumstances, D&O insurers will pay for the defense and settlement of securities class actions, but if the case is tried and

* *Francis J. Menton, Jr. is an attorney with Willkie Farr & Gallagher, LLP.*

lost the insurers reserve their rights to refuse to pay the judgment and to seek reimbursement of legal fees already paid.

Thus, while it may seem unlikely that cases commenced on little or no investigation after a market fluctuation would have a settlement rate reported as over 99%, the dynamic is clear once one understands that classes of thousands of shareholders are almost always certified, that theoretical damages in large numbers of cases are in the hundreds of millions (and even billions) of dollars, and that insurance available for a settlement would be quite problematic to realize after a loss at trial. Indeed, while the securities class action industry has had relative stability, a problem for defendants and their insurers has been upward movement in average settlements, undoubtedly driven in part by the sense of the plaintiffs' bar that defendants cannot allow cases not to be settled. No matter how slim the chance of an adverse judgment, that result, if it occurred, would be completely ruinous to defendants, whether individual officers and directors or the corporate issuer itself.

Thus, each year we see dozens of settlements in the range from mid-seven figures ($5 million +/-) to mid eight figures, representing only a few pennies on the dollar of notional damages that could have been in the hundreds of millions or even billions. Plaintiffs' class counsel typically gets a cut of one-fifth to one-third, depending on the size of the settlement. Not uncommon are settlements where the average shareholder class member gets under $100 while plaintiffs' class counsel gets several million, or even ten million dollars, or more.

Tellabs: The Supreme Court Breathes Life into the 1995 Reform Act

Of the many provisions of the 1995 Reform Act, one that appeared at first likely to result in many successful motions to dismiss was Section 21D(b)(2), 15 U.S.C. Section 78u-4, titled "Required state of mind." That section replaced the previous Federal Rule of Civil Procedure 9(b) as to the standard for alleging intent or state of mind ("intent... and other condition of mind of a person may be averred generally") with a specific pleading standard of particularity: "the complaint shall, with respect to each act or omission alleged to violate this title, state with particularity facts giving rise to a strong inference that the defendant acted with the required state of mind."

Yet the potential for Section 21D(b)(2) to weed out frivolous cases has proved quite difficult to realize. Many courts have simply taken the view that as long as the plaintiff alleges some facts—any facts—going to scienter, then he is entitled to the usual rule of all inferences in his favor at the pleading stage. Many courts have interpreted that rule as precluding the court from weighing whether the inference of scienter is "strong" or not; if the facts alleged gave rise to *any* inference of scienter, then the plaintiff would automatically receive the presumption that the inference was "strong." As an example, the Seventh Circuit, in its decision in the *Tellabs* case, stated the standard as being whether "a reasonable person could infer [from the complaint's allegations] that the defendant acted with the required intent."[6] That standard had no element that gave any meaning at all to the statutory prescription that the inference of scienter from the facts alleged must be "strong."

With *Tellabs*,[7] the Supreme Court has now instructed the lower courts that it most definitely is their job to weigh the competing inferences as to scienter, and to allow a securities class action to proceed only if "a reasonable person would deem the inference of scienter cogent and at least as compelling as any plausible opposing inference that one could draw from the facts alleged."[8] With lower courts under explicit direction to weigh competing inferences as to scienter on a motion to dismiss, they will undoubtedly be presented with motions by defendants in most or all cases seeking to have this weighing come out in the defendant's favor.

The web sites of many defense firms have greeted the *Tellabs* decision with a degree of triumphalism, but that is premature. Until a body of case law develops, it will not be known how much the new rule moves the needle in terms of numbers of cases that will actually be dismissed. Indeed, some commentators, such as John Coffee of Columbia Law School, writing in the *New York Law Journal* of July 19, 2007, have taken the view that the shift is relatively small, and far less than the rule change advocated by the Solicitor General in the case.

Nevertheless, it is already clear that there are categories of cases that in the past routinely survived and now are in jeopardy. One such category is the case based on facts revealed to plaintiffs' counsel by alleged "confidential sources." Such unnamed sources are a common occurrence in securities class action complaints, often turning up not in the pleading filed immediately after a stock price drop, but in an amended complaint filed at a later time, such as to parry an initial motion to dismiss. Indeed, the amended complaint in the *Tellabs* case itself sought to support its inference of scienter with alleged information from some 27 confidential informants.

Following immediately upon the Supreme Court's June 2007 decision in *Tellabs*, the Seventh Circuit has ruled that information provided by alleged confidential informants will not be sufficient to provide the needed "strong inference" of scienter.[9] Affirming dismissal of the complaint, the Seventh Circuit stated:

> One upshot of the approach that Tellabs announced is that we must discount allegations that the complaint attributes to five "confidential witnesses".... Perhaps these confidential sources have axes to grind. Perhaps they are lying. Perhaps they don't even exist....
>
> Our point, rather, is that anonymity conceals information that is essential to the sort of comparative evaluation required by Tellabs.

Beyond the specific instance of confidential sources, it is fair to say that the field is now wide open. Some complaints allege specific extrinsic indicators of scienter, such as sales of stock by insiders during the class period; but many complaints lack any such extrinsic indicators of scienter, and base their alleged inference only on the allegation that the facts that ultimately came out differed from the prior disclosures. The lower courts have previously struggled to come up with some method to distinguish any one of these cases from the others. While it is not clear exactly where the new line will be drawn, it is clear that a new complaint alleging an inference of scienter from the mere fact of new disclosures differing from old ones stands a substantial chance of being in jeopardy.

New Class Certification Decisions: Has Class Certification in Securities Class Actions Now Become Contestable in Many Cases?

Class certification is an essential element of the economics of securities class actions. In most cases involving large public companies, even the largest shareholders will own less than 1% of the securities that traded during an alleged class period; and often the very largest shareholders, such as mutual funds, will not be willing to bring a case at all. Even if there is a willing plaintiff with a 1% holding of the shares at issue, defeat of class certification will turn a potential one billion dollar claim with perhaps a $20 million settlement value into a $10 million claim with a settlement value well under $1 million—not enough to justify the time and effort of the plaintiff's lawyer.

Yet, as set forth above, a pair of Supreme Court cases in the 1970s and 1980s had made class certification almost impossible to oppose in most securities class actions, to the point where many defendants conceded the issue. Now, new cases from the Second and Fifth Circuits point to substantial categories of cases where class certification may be successfully opposed.

The Second Circuit decision is *In re Initial Public Offering Securities Litigation*.[10] The holding of the IPO Laddering Cases is that the district court on a class certification motion must make determinations that each of the Rule 23 prerequisites for class certification has been met, and that such determination can be made only if the district judge resolves factual disputes relevant to the Rule 23 requirements.[11] Based on that holding, the Second Circuit reversed class certifications granted by the district court and, in an unusual step, did not remand on this issue. The Second Circuit characterized its holding as being in accord with the Supreme Court's decision in *Eisen v. Carlisle & Jacquelin*;[12] but the holding in the IPO Laddering Cases was most definitely not in accord with the prior application of *Eisen* by many courts, who had interpreted the Supreme Court's decision as precluding extensive factual inquiry into class certification issues as a prohibited "preliminary inquiry into the merits."

The IPO Laddering Cases involved no fewer than 310 consolidated class actions, brought with respect to nearly every IPO that came out during the internet and telecommunications bubbles of 2000-2001. The complaints alleged improper conduct by the issuers and their banks, particularly in setting the pricing for the issuance and trading in the immediately following after-market. After denying motions to dismiss, the district court designated six "focus cases" for consideration of class certification, and then certified broad classes in each of the cases, with the class period covering the time immediately after the IPO. The particular fact which the Second Circuit found had not been and could not be established was the efficiency of the market for the newly issued stocks in the first weeks of trading, a fact which undermined the presumption of reliance.

As with the Supreme Court's *Tellabs* decision, the consequences of the IPO Laddering Cases are likely to be significant, but the full extent will only emerge with time. Certainly the IPO Laddering decision has an immediate impact on the 304 cases that were waiting in the wings while the Second Circuit focused on only six. Next in line will be any other cases arising from IPOs, heretofore a not insignificant portion of the securities class action universe. Also at risk will be any other cases where the trading market for the securities is less than robust—for example, stocks that are thinly traded or have had their markets disrupted for some reason.

The Fifth Circuit's decision, *Oscar Private Equity Investments v. Allegiance Telecom, Inc.*, has the potential to affect an even larger category of cases.[13] The Fifth Circuit phrases its holding in *Oscar* in relatively opaque terms: "we [require] plaintiffs invoking the fraud on the market theory to demonstrate loss causation."[14] But the practical implication could be great. In *Oscar,* the stock-price-inflating and allegedly false good news at issue were buried among various other bits of good news; and the stock-price-deflating corrective disclosure was buried among other bad news. How then could one know that the stock price inflation was caused by the particular allegedly false information as opposed to something else? Lacking an answer to that question, the Fifth Circuit ruled that plaintiff had failed to show loss causation and hence that the class could not be certified.

Factual patterns similar to that in *Oscar* are present in numerous cases. Indeed, it could well be possible for issuers to orchestrate the facts in some cases to make loss causation from any one particular disclosure more difficult to establish.

It is by no means certain that all circuits will follow the holding in *Oscar*. Indeed, there was a dissent in the case. Nonetheless, there is ample room for pushback against the notion underlying a high percentage of securities class actions, that every earnings disappointment must stem from a "fraud" that was the sole cause of a stock price inflation and subsequent drop. The *Oscar* case highlights that lack of clarity as to the unique causation of stock price deflation could defeat class certification and render much securities litigation non-viable.

Will the SEC Permit Companies to Compel Arbitration of Securities Class Actions?

On April 16, 2007, the *Wall Street Journal* reported that the SEC has under consideration a proposal to permit issuers to require arbitration of securities class actions against them. The potential mechanism to effectuate this result could be a provision in the by-laws. By that mechanism, theoretically claims against officers and directors, in addition to those against the issuer itself, could be subject to the arbitration requirement.

Relative to the incrementalism of court decisions, such a move by the SEC has the potential to bring about something closer to a revolution of the current securities litigation landscape. Of course, this proposal faces the prospect of major opposition before adoption, not least from the plaintiffs' bar and shareholder activist groups.

Even assuming arbitration would still entail class action treatment to the same extent as currently allowed in litigation, there can be little doubt that many defendants would prefer arbitration over jury trials, if available. Despite the lack of real world experience resulting from the extraordinarily high settlement rate, most defendants regard the jury process as containing a high random component with a real potential

for runaway damages. In a world of arbitration, potentially defendants might view at least some percentage of cases as triable.

CONCLUSION

The changes so far are incremental; the full consequences cannot yet be known. Yet new developments offer new opportunities for defendants to make motions that can actually lead to victory in more than a handful of securities class actions. Since trial is so rarely a viable option, defendants will not fail to present these motions.

Endnotes

1 ___ U.S. ___, 127 S.Ct. 2499 (2007)("Tellabs").

2 *Available at* www.nera.com/publication.asp?p_ID:2777.

3 *Available at* www.abanet.org/buslaw/blt/2002-07-08/wagerward.html.

4 417 U.S. 156, 177 (1974).

5 485 U.S. 224, 242 (1988).

6 437 F.3d 588, 602 (7th Cir. 2006).

7 127 S.Ct. 2499.

8 *Id.* at 2510.

9 Higginbotham v. Baxter International, Inc., ____ F.3d ____, 2007 WL 2142298 (7th Cir. July 27, 2007).

10 471 F.3d 24 (2d Cir. 2006)(the "IPO Laddering Cases").

11 *Id.* at 41.

12 417 U.S. 156 (1974).

13 487 F.3d 261 (5th Cir. May 16, 2007).

14 *Id.* at 265.

West Virginia Supreme Court Strikes Down Learned Intermediary Rule

*By James M. Beck & Theodore H. Frank**

In 1999, the New Jersey Supreme Court created an unprecedented exception to the learned intermediary rule, holding that it did not apply where the manufacturer had engaged in direct-to-consumer advertising.[1] The decision was especially surprising because New Jersey was one of the few states where the legislature had explicitly written the learned intermediary doctrine into statutory law.[2] *Perez v. Wyeth Laboratories, Inc.* was not followed in other jurisdictions until 2007,[3] when the West Virginia Supreme Court used the rationalization of direct-to-consumer advertising in *State ex rel. Johnson & Johnson Corp. v. Karl* (a case where there was none at issue) to hold that it would be the first state to disregard the learned intermediary doctrine.[4]

The learned intermediary doctrine has been adopted by state courts in thirty-nine states and by the District of Columbia—including thirty-four states where the decision was made by the highest court, and the highest court of the fourteen most populous states. Federal courts have made an *Erie* prediction in eight other states, and in Puerto Rico, that the state would adopt the learned intermediary rule.[5] Thus, West Virginia's highest court has parted company with the forty-seven states that have judicially recognized this doctrine.

The West Virginia Court's Decision

But in *Karl*, a 3-2 decision with four opinions, the West Virginia Supreme Court disagrees that virtually all states have adopted the learned intermediary doctrine. The lead opinion counts "decisions of only the highest state courts," and concludes that "a mere" twenty-one states have expressly adopted the doctrine.[6] One reason for the discrepancy is the omission in *Karl* of cases involving prescription medical devices.

The majority goes on to reject all of the usual justifications for the doctrine. "At the outset," it notes, "the learned intermediary doctrine is not a modern doctrine. Rather, its origins may be traced as far back as 1925."[7] This is a curious observation in a common law system governed by stare decisis; the doctrine has been on the books, and increasingly embraced, for more than eighty years. This casual approach to precedent is of concern to more than just pharmaceutical defendants;[8] it potentially threatens other doctrines, such as product identification, remote causation, and even the burden of proof.

Criticizing the learned intermediary rule as "outdated" would hold if the rule was only rarely invoked, or if there were a general trend away from it. Neither is the case here. The rule is as routinely followed now as in prior years.[9] Indeed, *Karl* cites not a single other opinion that outright rejects the rule. That is because no such opinion exists—until now. Remarkably, the *Karl* court relied on the absence of a ruling in the highest courts in twenty-two states (some of which are erroneously included) as "precedent" for rejecting the rule.[10]

** James M. Beck is Counsel at Dechert, and represents pharmaceutical companies in his practice. Theodore H. Frank is a Resident Fellow at the American Enterprise Institute for Public Policy Research, and before joining AEI represented pharmaceutical companies in private practice.*

The majority opinion proceeds to attack direct-to-consumer (DTC) advertising. While there are certainly arguments that DTC advertising can be abused, there is also evidence that DTC advertising has been a substantial benefit to consumers and to health outcomes.[11] Nevertheless, the majority held that DTC advertising "obviates each of the premises upon which the [learned intermediary] doctrine rests."[12] In a world of DTC advertising, patients become active participants in their health care, and they ask for particular drugs by name. And the existence of DTC ads supposedly proves that it is possible to explain accurately the risks and benefits of drugs directly to patients.[13] The majority therefore saw no benefit in adopting the learned intermediary doctrine. Manufacturers should simply warn patients directly of the risks associated with prescription drugs.

This reasoning is problematic, however. First, making decisions about the optimal level of direct-to-consumer advertising might seem a usurpation of legislative, executive, or administrative prerogative—particularly in a case in which the defendant did not, in fact, engage in this practice. Moreover, it is peculiar to respond to the supposed harm of DTC advertising by enacting a rule of law that forces manufacturers to do even more of it. By abrogating the learned intermediary rule altogether—supposedly based upon its critique of DTC advertising—*Karl* virtually forces every prescription medical product manufacturer to engage (even for drugs and devices that have not previously been the subject of DTC advertising) in precisely the DTC conduct about which the decision complains. How else is a company supposed to satisfy this new duty to convey warnings about prescription medical products directly to the general public, bypassing the doctor altogether?

Second, although there is a great deal of DTC advertising these days, there are plenty of drugs that are *not* advertised this way—such as the drug in *Karl*. Generally, only a relatively few blockbuster drugs are promoted with expensive ad campaigns; the $2.38 billion/year spent on DTC advertising in 2001 is less than 2% of the annual cost of prescription drugs. Ordinary drugs, or drugs for rare conditions, do not merit the cost of DTC advertising. Why should the protections of the learned intermediary doctrine be removed from drugs that were (1) never advertised DTC, or (2) advertised DTC, but the plaintiff-patient who brought the case never saw or heard the ad? Even New Jersey applies the learned intermediary doctrine to non-DTC-advertised drugs, and appears to recognize causation as a defense.

Third, it is ordinarily impractical to warn patients directly about the risks associated with drugs. Every doctor has access to the *Physicians' Desk Reference* and can locate and understand prescription drug labeling. Patients are less able to find and read the package inserts, and less able to understand them. They will be even less likely to understand them if drug companies are forced to do additional disclosures for fear of liability. As it is, a 2002 FDA study found, only 16% of patients say they read "almost all" or "all" of even the brief summary disclosures

in DTC advertising, a problem that will surely be exacerbated when the disclosures become longer.[14] Nor is it universal for pharmacies to distribute package inserts when they dispense prescriptions.

This leads us to the fourth point, which is that the opinion is unrealistic. Who, other than physicians and product liability lawyers, actually reads drug package inserts from cover to cover? The average patient surely does not. What purpose does it serve for courts to dictate the contents of documents that go unread?

Karl states that "only" four state supreme courts have adopted the learned intermediary rule since direct-to-consumer ("DTC") drug advertising "proliferate[ed]" in 1997.[15] However no court—other than that of *Karl*—rejected the rule during that (or any other) period. Nor does *Karl* acknowledge the additional seventeen high court opinions from twelve other states that reaffirmed the learned intermediary rule during this same period.[16]

The Learned Intermediary Doctrine & Public Policy

The learned intermediary rule fills an important role in the law, ensuring harmonious operation of state product liability law with the "unique system used to distribute prescription [products]."[17]

There are several reasons why no other jurisdiction shares the West Virginia view on display in *Karl*. First and foremost, the learned intermediary rule makes sense because it reflects common practice. Ever since the FDA started regulating prescription drugs in the 1930s, those drugs and (more recently) prescription medical devices have only been available to the public after doctors have decided that they are appropriate for the treatment of particular patients. Thus, the rule reflects how prescription drugs and medical devices are actually distributed. These mandatory federal restrictions on distribution are why prescription medical products are *not* the same as a "lawnmower," the concurring opinion's example.[18] Anybody can go down to the hardware store and buy a lawnmower for any reason (or no reason) at all. That is not possible—legally—with prescription medical products.

Moreover, these FDA requirements are motivated by longstanding safety concerns. The FDA defines "[p]rescription drug" as "any drug (including any biological product...) required by Federal law... to be dispensed only by a prescription."[19] All drugs are presumed to be prescription drugs unless the FDA "finds such requirements unnecessary for the protection of the public health by reason of the drug's toxicity or other potentiality for harmful effect."[20] Similarly, a "prescription device" is "[a] device which, because of any potentiality for harmful effect... is not safe except under the supervision of a practitioner licensed by law to direct [its] use."[21]

Federal limitations upon distribution of prescription medical products are thus explicitly based upon their inherent safety risks. Nobody is telling purchasers of lawnmowers that they cannot have them unless they first go to a government-licensed professional and obtain pre-certification that the purchase is necessary.

Fundamentally, the learned intermediary rule is "based on the principle that prescribing physicians act as 'learned intermediaries' between a manufacturer and consumer and, therefore, stand in the best position to evaluate a patient's needs and assess the risks and benefits of a particular course of treatment."[22] "When the purchase of the product is recommended or prescribed by an intermediary who is a professional, the adequacy of the instructions must be judged in relationship to that professional."[23] The rule is legal recognition of something that is as true today as ever: prescription medical products are not available to the public at large precisely because the FDA has determined that such products have inherent, unavoidable risks of sufficient gravity to require a doctor's evaluation before anyone can use them.[24]

Indeed, the warnings in the package insert are designed for doctors to read, not laypeople—containing jargon like "treatment-emergent hyperglycemia-related adverse events," "mean (SD) pharmacokinetic parameters," "agranulocytosis," and "glomerular filtration rates." Such warnings are "designed for the physician and not the patient."[25]

The rule further reflects reality by recognizing that doctors make most prescribing decisions for their patients in the context of a physician-patient relationship in which patients rely upon their doctors to explain treatment decisions, and do not rely upon their own reading of product labeling.[26] It would not just be burdensome, but dangerous, for the law to demand the dumbing-down of technical medical product information. Putting aside practicality, patients should follow doctors' orders, and should not be conducting possibly ill-informed self-evaluations of their own prescribed medical treatments.[27]

Further, this sort of scientific precision in labeling is almost certainly preferable, at least as long as prescription drugs are not freely available. Better to have doctors to break things down to their patients in one-on-one conversations, rather than having patients try to read package inserts themselves. At least until *Karl*, West Virginia recognized the doctrine of informed consent, which requires doctors to do just that.[28] Thus, at best, *Karl* has just created a regime of redundant and overlapping liability with the patients stuck in the middle—not knowing whom to believe if doctors and drug companies say different things.

Thus, "[i]t is... the duty of the physician to advise the patient of any dangers or side effects associated with the use of the drug as well as how and when to take the drug."[29] "Education of the physician, on the one hand, and communication to the patient, on the other, are distinct processes, and the manufacturer's duty involves only the former."[30] If the product requires a doctor's prescription—and a doctor, in fact, prescribed it—the rule properly applies.[31]

Conversely, a tort system that requires manufacturers to bypass doctors and warn patients directly would disrupt the physician-patient relationship. "[I]mposition of a generalized duty to warn would unnecessarily interfere with the relationship between physician and patient."[32] "When the physician-patient relationship does exist... we hesitate to encourage, much less require, a drug manufacturer to intervene in it."[33]

Doctors are highly trained, with their own professional and legal obligations to their patients. Among other things, they keep those patients from overreacting to the many warnings

(often directed to limited situations or patient populations) that the FDA requires product labeling to include.[34] Thus, the learned intermediary rule keeps doctors where they should be: front and center, as the primary source of patient information about medical treatment.

The rule ensures that, once a manufacturer has warned a doctor, it "may reasonably assume that the physician will exercise his informed judgment in the patient's best interests."[35] "[F]ailure [of medical personnel] to perform their duties from that point forward do[es] not operate to create, or to extend, a manufacturer's duty to warn third-party family members, bystanders, or any persons other than the learned intermediary."[36] Thus, manufacturers are not "advisors" to physicians during the informed consent process,[37] nor are they responsible for how physicians conduct their business.[38]

An additional practicality consideration buttresses the rule. "[T]he treating physician is in a better position to warn the patient than the manufacturer."[39] It is unrealistic to depend upon pharmacies to take up the slack, as they usually do not dispense prescription medical products to patients with full warnings.[40] Thus, it is frequently impractical, or even impossible, for manufacturers to provide direct warnings to unknown patients, particularly in a medical emergency.[41]

The Concurrence

Critics have charged that "West Virginia is now routinely called a 'judicial hellhole' with the 'worst legal system in America.'"[42] The state of jurisprudence in West Virginia is exhibited most starkly by the concurring opinion in *Karl* of two justices, which states:

> Suppose Patient John Doe visits his *small-town West Virginia doctor*. Further suppose he is prescribed a drug by his doctor that causes him serious injury. Suppose that the drug is one that is heavily advertised. Patient Doe then sues his West Virginia doctor and the drug manufacturer for the injury caused by the drug. If this Court were to adopt the learned intermediary doctrine, the West Virginia doctor would remain in the lawsuit, but the drug manufacturer would not remain in the suit and would not be liable for damages if the drug manufacturer *could show that it warned the doctor of the risks* of injury associated with the drug. Thus, a *small-town West Virginia doctor* would become solely responsible for the injury to Patient Doe while an *out-of-state multi-million dollar drug manufacturer* is off the hook…. This result simply would be unfair.[43]

The characterization of the result as "unfair" is strange. The "out-of-state multi-million dollar drug manufacturer" cannot make a prescription drug that has no significant risk of injury. If it did, the FDA would not have classified the product as a prescription drug to start with, and the learned intermediary rule would never come into play. So, it cannot be the fact that the product causes injury that makes anything unfair.

Nor is the result "unfair" due to an inadequate warning. The hypothetical provides that the doctor was in fact properly warned. If the warning were inadequate, then the learned intermediary rule would not preclude liability of the manufacturer.[44] If the "out-of-state multi-million dollar drug manufacturer" adequately warned the "small-town West Virginia doctor" about the inherent risks of its product, then what wrong did it commit? "[T]he fact that a particular drug might produce unfortunate side effects makes it 'unavoidably unsafe' but not 'unreasonably dangerous', and strict liability will not obtain if 'proper warning is given, where the situation calls for it.'"[45]

Suppose the "small-town West Virginia doctor" prescribed the hypothetical drug notwithstanding adequate warnings of the harm it caused his or her patient, or did so without telling the patient what could happen. Suppose there was malpractice, or perhaps a breach of that doctor's informed consent duty. What business does the "out-of-state multi-million dollar drug manufacturer" have in the suit—except that it has deep pockets? The "unfairness" advanced as the basis for the concurrence's hypothetical does not seem to be tied to any tort concept of fault, but rather wealth ("multi-million dollar company") and residence ("out-of-state").

But it is not a court's job to impose what amounts to an extra-legislative tort tax in order to extract money (ultimately coming almost entirely from the pockets of other, out-of-state product users) for the benefit of local plaintiffs regardless of any concept of legal fault. Moreover, if West Virginia's tort system allows what amounts to absolute liability on pharmaceutical manufacturers, manufacturers will inevitably re-consider whether the state's (relatively small) market is worth serving.

Endnotes

1 Perez v. Wyeth Laboratories, Inc., 734 A.2d 1245 (N.J. 1999).

2 NJ Product Liability Act L. 1987, c. 197, § 4.

3 *In re Norplant Contraceptive Prods. Liab. Litig.*, 215 F. Supp.2d 795, 812 (E.D. Tex. 2002) ("New Jersey law is in direct conflict with the law of every other jurisdiction in the United States"); *see also*, Colacicco v. Apotex, Inc., 432 F. Supp.2d 514, 547 n.30 (E.D. Pa. 2006); *In re Meridia Prods. Liab. Litig.*, 328 F. Supp.2d 791, 812 n.19 (N.D. Ohio 2004), aff'd, 447 F.3d 861 (6th Cir. 2006); Larkin v. Pfizer, Inc., 153 S.W.3d 758, 766 (Ky. 2004); Eady v. Lansford, 92 S.W.3d 57, 60-61 (Ark. 2002); *see* Allen v. G.D. Searle & Co., 708 F. Supp. 1142, 1148 (D. Or. 1989) (rejecting similar argument before *Perez* was decided).

4 State ex rel. Johnson & Johnson Corp. v. Karl, ___ S.E.2d ___, 2007 WL 1888777 (W. Va., 2007).

5 A full list of precedents is *available at* <http://druganddevicelaw.blogspot.com/2007/07/headcount-whos-adopted-learned.html>.

6 *Karl*, 2007 WL 1888777, at *8.

7 *Id*. at *10.

8 *Cf.* State ex rel. Mobil Corp. v. Gaughan, 211 W.Va. 106, 563 S.E.2d 419 (2002) (approving trial plan of over eight thousand asbestos plaintiffs against two hundred fifty defendants).

9 *E.g.,* Ethicon Endo-Surgery, Inc. v. Meyer, ___ S.W.3d ___, 2007 WL 1095552, at *2 (Tex. App. April 12, 2007); Bodie v. Purdue Pharma LP, ___ F.3d ___, 2007 WL 1577964, at *5-6, 11 (11th Cir. June 1, 2007) (applying Alabama law); Stilwell v. Smith & Nephew, Inc., 482 F.3d 1187, 1194 (9th Cir. 2007) (applying Arizona law); In re Zyprexa Prods. Liab. Litig. v. Eli Lilly & Co., ___ F. Supp. 2d ___, 2007 WL 1678078, at *28 (E.D.N.Y. June 11, 2007) (applying Pennsylvania, North Carolina and Florida law); Soufflas v. Zimmer, Inc., 474 F. Supp.2d 737, 751 (E.D. Pa. 2007); Madsen v. American Home Prodts. Corp., 477 F. Supp.2d 1025, 1033 (E.D. Mo. 2007) (applying Iowa law); Cowley v. Abbott Laboratories, 476 F. Supp.2d 1053, 1060 (W.D. Wis. 2007) (applying North Carolina law); Hill v. Wyeth, Inc., 2007 WL 674251, at *3-4 (E.D. Mo. Feb. 28, 2007); Saraney v. Tap Pharm. Prods., 2007 WL 148845, at *6 (N.D. Ohio Jan. 16, 2007); Paparo v. Ortho McNeil Pharm., 2007 WL 121149, at *3 (S.D. Fla. Jan. 10, 2007).

10 2007 WL 1888777, at *21. For example, *Karl* erroneously lists the District of Columbia, when its highest court acknowledged the learned intermediary rule in *Mampe v. Ayerst Labs.,* 548 A.2d 798, 802 n.6 (D.C. 1988).

11 *See generally* Natonal Health Council, Direct-To-Consumer Prescription Drug Advertising: Overview & Recommendations (January 2002).

12 State ex rel. Johnson & Johnson Corp. v. Karl, ___ S.E.2d ___, 2007 WL 1888777, *17 (W. Va. 2007).

13 *Id.* at *19.

14 K.J. Aikin, "The Impact of Direct-to-Consumer Prescription Drug Advertising on the Physician-Patient Relationship," Presentation at FDA public meeting, Washington, D.C. (22–23 September 2003). www.fda.gov/cder/ddmac/aikin/index.htm (13 April 2004).

15 *Karl*, 2007 WL 1888777 at *35.

16 Hurley v. Heart Physicians, P.C., 898 A.2d 777, 783-84 (Conn. 2006); Rite Aid Corp. v. Levy-Gray, 894 A.2d 563, 577 (Md. 2006); Walls v. Alpharma USPD, 887 So.2d 881, 883 (Ala. 2004); Janssen Pharmaceutica, Inc. v. Bailey, 878 So.2d 31, 57 (Miss. 2004); Howland v. Purdue Pharma, L.P., 821 N.E.2d 141, 146 (Ohio 2004); Morguson v. 3M Corp., 857 So.2d 796. 801-02 (Ala. 2003); Schaerrer v Stewart's Plaza Pharmacy, Inc., 79 P.3d 922, 928-29 (Utah 2003); Happel v. Wal-Mart Stores, Inc., 766 N.E.2d 1118, 1127 (Ill. 2002); Hansen v. Baxter Healthcare Corp., 764 N.E.2d 35, 42 (Ill. 2002); Cottam v. CVS Pharmacy, 764 N.E.2d 814, 820 (Mass. 2002); Moore v. Memorial Hospital, 825 So.2d 658, 664 (Miss. 2002); Bennett v. Madakasira, 821 So.2d 794, 804 (Miss. 2002); Vaccariello v. Smith & Nephew Richards, Inc., 763 N.E.2d 160, 164 (Ohio 2002); Perez, 734 A.2d at 1257; Spensieri v. Lasky, 723 N.E.2d 544, 549 (N.Y. 1999); E.R. Squibb & Sons, Inc. v. Farnes, 697 So.2d 825, 827 (Fla. 1997); Edwards v. Basel Pharmaceuticals, 933 P.2d 298, 300-01 (Okla. 1997).

17 Pittman v. Upjohn Co., 890 S.W.2d 425, 429 (Tenn. 1994).

18 *Karl*, 2007 WL 1888777, *59 (Maynard, J. concurring).

19 21 C.F.R. §203.3(y).

20 21 C.F.R. §310.200(b).

21 21 C.F.R. §801.109.

22 *Vitanza*, 778 A.2d at 836-37.

23 *Mampe*, 548 A.2d at 802 n.6.

24 *See also, e.g.,* Coyle v. Richardson-Merrell, Inc., 584 A.2d 1383, 1387 (Pa. 1991); Larkin v. Pfizer, Inc., 153 S.W.3d 758, 763 (Ky. 2004); Ellis v. C.R. Bard. Inc., 311 F.3d 1272, 1287 (11th Cir. 2002) (applying Georgia law); Fane v. Zimmer, Inc., 927 F.2d 124, 129 (2d Cir. 1991) (applying New York law); Phelps v. Sherwood Medical Industries, 836 F.2d 296, 298-99 (7th Cir. 1987) (applying Indiana law).

25 Oksenholt v. Lederle Laboratories, 656 P.2d 293, 297 (Or. 1982). *See also, e.g.,* Martin v. Hacker, 628 N.E.2d 1308, 1311 (N.Y. 1993); Reyes v. Wyeth Laboratories, 498 F.2d 1264, 1276 (5th Cir. 1974) (applying Texas law); Freeman v. Hoffman-La Roche, Inc., 618 N.W.2d 827, 841-42 (Neb. 2000).

26 *See*, *e.g.*, West v. Searle & Co., 806 S.W.2d 608, 613 (Ark. 1991); Humes v. Clinton, 792 P.2d 1032, 1039 (Kan. 1990); Williams v. American Medical Systems, 548 S.E.2d 371, 375 (Ga. App. 2001); Reaves v. Ortho Pharmaceutical Corp., 765 F. Supp. 1287, 1289 (E.D. Mich. 1991).

27 *Perez.*, 734 A.2d at 1254; Dyer v. Best Pharmacal, 577 P.2d 1084, 1088 (Ariz. App. 1978); *In re Norplant Contraceptive Prods. Liab. Litig.*, 215 F. Supp.2d 795, 815 (E.D. Tex. 2002).

28 *E.g.*, Cross v. Trapp, 294 S.E.2d 446, 454-56 (W. Va. 1982).

29 584 A.2d at 1385.

30 Harwell v. American Medical Systems, Inc., 803 F. Supp. 1287, 1299 (M.D. Tenn. 1992).

31 Ellis, 311 F.3d at 1282.

32 Morgan v. Wal-Mart Stores, Inc., 30 S.W.3d 455, 467 (Tex. App. 2000).

33 Swayze v. McNeil Laboratories, Inc., 807 F.2d 464, 471 (5th Cir. 1987) (applying Mississippi law).

34 *Larkin*, 153 S.W.3d at 764; *see also* "Requirements on Content and Format of Labeling for Human Prescription Drug and Biological Products," 71 Fed. Reg. 3922 (Jan. 24, 2006) (to be codified at 21 CFR Parts 201, 314, and 601) (acknowledging danger of overwarning).

35 Tracy v. Merrell Dow Pharmaceuticals, Inc., 569 N.E.2d 875, 876 (Ohio 1991).

36 Ellis, 311 F.3d at 1283. *See also* Brooks v. Medtronic, Inc., 750 F.2d 1227, 1232 (4th Cir. 1984) (applying South Carolina law).

37 Polley v. Ciba-Geigy Corp., 658 F. Supp. 420, 421 (D. Alaska 1987).

38 *Ellis*, 311 F.3d at 1283; Prohaska v. Sofamor, S.N.C., 138 F. Supp.2d 422, 444 (W.D.N.Y. 2001).

39 McCombs v. Synthes, 587 S.E.2d 594, 595 (Ga. 2003).

40 *Tracy*, 569 N.E.2d at 880.

41 *West*, 806 S.W.2d at 613; Cunningham v. Charles Pfizer & Co., 532 P.2d 1377, 1381 (Okla. 1974); Talley v. Danek Medical, Inc., 179 F.3d 154, 163 (4th Cir. 1999) (applying Virginia law).

42 Strahin v. Sullivan, ___ S.E.2d ___, 2007 WL 1891551, at *1 (W. Va. Feb. 21, 2007) (Starcher, J, dissenting).

43 *Karl*, 2007 WL 1888777, at.*55-56 (emphasis added) (Maynard, J. concurring).

44 *E.g.,* Proctor v. Davis, 682 N.E.2d 1203, 1211, 1213 (Ill. App. 1997) ("the failure of the prescribing and treating physicians to learn of the risks of a drug from other sources does not relieve the manufacturer of liability for harm resulting from its own failure to adequately warn").

45 *Larkin*, 153 S.W.3d at 761 (quoting Restatement (Second) of Torts §402A, comment k (1965)).

Professional Responsibility & Legal Education

Attorneys' Fees in Class Actions: The Problem Remains

*By Jack Park**

In 2005, as a member of a plaintiff class in a securities lawsuit, I objected to the attorneys' fee component of a proposed settlement. Over my objection, the court approved a settlement that resulted in a class counsel's recovery of a contingency fee of 25% (plus expenses) from a settlement fund of $80 million—a figure that represented a multiplier of 4.7 on the "lodestar" figure derived by multiplying the hours worked by the typical fee. Put differently, class counsel would have had to have worked far more hours at their regular billing rate to receive that amount in fees.

Notwithstanding my lack of success, I think it worthwhile to reflect on this experience for two reasons. First, the attorneys' fee component of class action settlements has been the subject of substantial debate in recent years. One question that has been discussed is whether attorney fee awards are increasing. Secondarily, the debate continues because Congress did not address attorney fees to any substantial extent in the Class Action Fairness Act of 2005.[1]

In this article, I will use my experience to describe the practice of attorney fee litigation in the class action context. Without suggesting that it is typical, I will discuss my experience as an objector, first setting out the ethical background for determining the reasonableness of an attorney's fee. Then, I will describe the Xcel litigation and the district court's consideration of the claim for attorney fees and expenses. Finally, I will try to put my experience and this lawsuit into the broader context. In my judgment, it is only by objecting that unnamed class members can bring their interests before the courts, and they should pursue this avenue.

The Ethical Background

This discussion, like any discussion of attorneys' fees, takes place against the backdrop of the rules of legal ethics. Rule 1.5 of the ABA Model Rules of Professional Conduct requires that the fees and expenses charged by an attorney not be "unreasonable."[2] Rule 1.5 further provides:

The factors to be considered in determining the reasonableness of a fee include the following:

(1) the time and labor required, the novelty and difficulty of the questions involved, and the skill requisite to perform the legal service properly;

(2) the likelihood, if apparent to the client, that acceptance of the particular employment will preclude other employment by the lawyer;

(3) the fee customarily charged in the locality for similar legal services;

(4) the amount involved and the results obtained;

(5) the time limitations imposed by the client or the circumstances;

(6) the nature and length of the professional relationship with the client;

(7) the experience, reputation, and ability of the lawyer or lawyers performing the services; and

(8) whether the fee is fixed or contingent.[3]

In that regard, Rule 1.5 permits the use of contingent fees where they are not otherwise prohibited, and likewise requires that they be reasonable.[4]

These ethical rules prompt several observations. Ultimately, the Rule 1.5 factors should not be considered in a vacuum, but, rather, as they play out in the market.

First, an attorney's hourly billing rate presumptively reflects that lawyer's skill and experience as measured by the applicable market.[5] To the extent that skill and experience may be subsequently taken into account in leveraging an enhanced fee, either directly or by pointing to results, a measure of double-counting would appear to be going on.[6] We should expect a lawyer who bills at a higher hourly rate to achieve good results, and paying that lawyer at those rates goes a long way toward compensating that lawyer adequately. Likewise, even when engaging in discovery and preparing dispositive motions in a particular case is "particularly time-consuming or demanding... that fact would be captured in the number of hours expended on the litigation, and the plaintiffs' counsel would be compensated accordingly in any fee award."[7]

Second, Rule 1.5's focus on the client's perception that his case may preclude other employment by the lawyer is incomplete, because the lawyer is a party to that transaction. Even though some cases may consume a disproportionate amount of time or effort, the lawyer should be in a better position than the prospective client to anticipate them. In lawyer-driven litigation like securities, redistricting, and institutional reform litigation, the lawyer who does not expect to do a large amount of work is likely inexperienced. Rather than looking at the client's perception of the preclusive effect of a proposed lawsuit, the reasonable expectations of an experienced lawyer should be imputed to the lawyer.

Third, law firms that take on a portfolio of clients and matters spread the risk that any attorney will take on litigation that consumes a disproportionate of time and other resources.[8] Those law firms, and lawyers generally, know how to turn down business that does not offer a reasonable likelihood of success and remuneration.[9]

The Xcel Litigation

By notice dated January 14, 2005, I was advised of the proposed settlement of class actions involving claims made on behalf of the purchasers of the common stock of Xcel Energy, Inc., between January 31, 2001 and July 25, 2002.

** Jack Park was formerly an Assistant Attorney General for the State of Alabama. He now serves as Special Assistant to the Inspector General for the Corporation of National and Community Service.*

Xcel Energy is a public utilities holding company that, through its subsidiaries, was serving electrical and natural gas retail customers in twelve western and midwestern states. I was a member of the class because I had purchased shares of one of Xcel Energy's predecessors and had participated in the dividend reinvestment program. No doubt notice of the class action had been published somewhere (in a publication I do not read), notice of the settlement was published too (again, in publications I do not read), and the litigation was discussed in Xcel's annual reports, but the notice of January 14, 2005 was the first notice I received of the lawsuit and my plaintiff status.

Xcel Energy's difficulties arose from its NRG subsidiary, which had been created in 1989 to tap into the domestic and international market for independent power production. NRG bought, built, and operated power plants, funding its operations with debt that was paid from the proceeds of sales of power in the competitive wholesale market. Before 2002, NRG's operations were sufficiently profitable that Xcel sold some 26% of the stock to the public.[10] Market conditions changed in 2001, however, and, as the prices for independently generated power fell, NRG had trouble handling its debt load. In December 2001, Moody's placed NRG's credit rating, then investment grade, on review for potential downgrading.[11] The problem was that, if NRG were downgraded, it would have to post cash collateral that it did not have, and, if it did not post the collateral, it would be in default. The district court observed: "As NRG's liquidity problems worsened, the market began showing signs of reduced confidence in the IPP [Independent Power Producer] sector as a whole."[12]

Xcel took steps to assist NRG, including committing additional capital and purchasing the public shares of NRG to bring it back in house. NRG's problems continued, however, and Xcel was dragged down with it. At the beginning of the class action period, Xcel's stock sold for $25.47 per share, and, as of June 3, 2002, it sold for $21.20.[13] By July 1, 2002, Xcel's stock was selling for $15.93 a share, and it was still $15.00 per share at the close of business on July 15.[14] Finally, the problem became public. On July 25, 2002, Xcel Energy revealed that its credit facilities contained a cross-default provision, such that, if NRG defaulted, Xcel would be called on to cure that default and might, in turn, be in default on its covenants with its lenders. Put differently, NRG's problems threatened to drag Xcel down too.

Xcel held a conference call with analysts the next day. The price per share declined from its July 25 close of $11.94 to $7.55 on July 26, and to $5.66 the following day, before rebounding. Ultimately, the uncertainty was cleared up, but not without cost, after Xcel severed its relationship with NRG and negotiated a release of the cross-default provision in exchange for concessions that increased its cost of borrowing and limited its access to credit funds. In addition, Xcel cut its dividend in half.[15]

The first lawsuit alleging violations of the Securities Exchange Act of 1934 and SEC Rule 10b-5 was filed on July 31, 2002. That lawsuit named Xcel and its former president and CEO, its CFO, and the former Chair of its Board as defendants. The plaintiffs alleged that the defendants made false and misleading statements relating to the relationship between Xcel and NRG and the effect of NRG's problems on Xcel. In short order, thirteen more securities actions were filed, as well as an action on behalf of holders of NRG Senior Notes, a shareholder derivative action, and two ERISA lawsuits. After the lawsuits were consolidated and class representatives appointed, the defendants moved to dismiss the complaint. The district court granted that motion in part and denied it in part.[16] After reviewing the documents produced by the defendants and engaging in mediation, but before any depositions were taken, the parties reached a settlement under which the defendants would pay $80 million to the securities plaintiff class and $8 million to the ERISA plaintiff class. Class counsel for each of those classes would receive 25% of the fund plus expenses.

I joined six other individuals and entities in objecting to the proposed settlement. In my letter, dated March 17, 2005, I noted that, as of then, securities class counsel had not filed a fee application. They did so on or about March 21, 2005. In addition, I objected to the anticipated hourly rate, noting that a lawyer who billed at $1,000 an hour would have to bill 20,000 hours to produce a bill of $20 million. I noted that counsel for plaintiffs in civil rights cases in Alabama customarily identified an hourly rate of $300 or less. I also objected to the use of a multiplier that would be used to back into the requested award; I anticipated that counsel would calculate the lodestar, determine what multiplier was needed to turn the lodestar into the award, then defend the multiplier. I further recommended that the court cut the contingency fee in half, to 12 ½%, which would result in an award of $10 million plus expenses. Again, because I filed my objection before class counsel filed their fee application, my objections were necessarily general.

In their fee application, class counsel for the Xcel securities plaintiff class reported spending some 10,400 hours of time on the case. When multiplied by hourly rates ranging between $250 and $650 for lawyers and $60 and $195 for paralegal time, the resulting lodestar figure was $4,255,949.[17] A multiplier of 4.7 was required to turn that figure into $20 million. Class counsel contended that the results they achieved and the difficulty of the case justified the multiplier. They also showed that, by reference to the results in other class action settlements, neither the 25% contingency nor the 4.7 multiplier was remarkable.

The district court overruled the objections and awarded the requested fees and expenses to counsel for the securities and the ERISA plaintiff classes.[18] The court reasoned that the contingency of 25% comports with awards in similar cases, courts in other securities cases had approved multipliers in excess of four, and the 25% contingency was not unreasonable when cross-checked against the lodestar figure and the multiplier. With respect to the last, the cross-checking, the Eighth Circuit merely uses the lodestar to cross-check the contingency.[19]

Attorneys' Fees in Class Actions

This experience, typical or not, prompts consideration of several of the common criticisms of class action practice. In particular, it raises questions about the attorney-client relationship and the reasonableness of the attorney fee award. There has been a debate about whether such awards are rising, but the real question may be whether anything limits them. And, whether there are limits or not, how should class members respond to notices of settlement?

Bill Lerach, one of the more prominent class counsel, has famously said, "I have the greatest practice in the world because I have no clients. I bring the case. I hire the client. I do not have some client telling me what to do. I decide what to do."[20] As a general matter, the rules of legal ethics leave tactical decisions to the lawyer. So, in theory at least, the lawyer has plenty of latitude. Even so, those pesky clients have a say in the formation of the relationship and need to be kept informed, and some clients are peskier than others. And, whether clients are pesky or not, class counsel cannot completely dispense with them.[21] In a class action, class counsel form a relationship with a class representative, and those class representatives, deemed capable of adequately representing the unnamed class members, make the strategic decisions.[22]

In a 2003 Working Paper, Theodore Eisenberg and Geoffrey Miller suggested that, notwithstanding claims to the contrary, attorney fee awards in class actions were not increasing.[23] In a 2004 program at the American Enterprise Institute, the participants, including Eisenberg, discussed possible limitations in the data, noting an apparent absence of cases from state courts and so-called "magnet" jurisdictions.[24] Paul Rubin, from Emory University, further suggested that the maximum of the range would become the mean, something that might be bearing out in practice.[25] In Xcel, for example, class counsel justified their 25% contingency by pointing to, among other things, the Third Circuit's reliance on a declaration by John Coffee, from Columbia Law School, suggesting that the average recovery in securities actions involving settlements greater than $10 million was 31%, and percentage recoveries between 25% and 30% were "fairly standard" in cases involving settlements between $100 and $200 million.[26] Class counsel also pointed to awards in other cases that purported to show that their 25% contingency and the resulting multiplier of 4.7 were not out-of-line. With the district court's ruling upholding that multiplier, these and other class counsel can point to it to justify their own multipliers in future cases.

Arguments like this are flawed because they do not represent the operations of the market for attorney services. Instead, they represent the actions of courts justifying awards to counsel, which are akin to the creation of hot-house flowers. The Eleventh Circuit has criticized this approach, observing, "Prior awards are not direct evidence of market behavior; the court is not a legal souk."[27] It also explained that, while there was some "inferential evidentiary value" to prior awards, giving them controlling weight over evidence of a lawyer's actual billing rates and practices "equates to [improperly] giving the prior awards issue-preclusive value against a party whose interests were not even arguably represented in the prior litigation."[28]

If awards are not increasing, it is not clear what restrains them. Certainly, the process does not. Class counsel and class representatives have an incentive to settle cases when the reward in hand exceeds the likely results down the road. Defendants have little incentive to object; they want to bind as many potential plaintiffs as possible and, having negotiated the settlement, have little incentive to upset any part of it.[29] And, courts have an incentive to dispose of cases.[30] None of these actors behaves irrationally when acting in this fashion. But, where does that leave the unnamed class members? They can object, but the plaintiffs' counsel want their money; the defendants want their deal and may have conveyed their silence; and the courts want the cases gone.

Objections are not easy. In my case, the hearing was in Minneapolis. The district court graciously allowed me to participate by telephone. But my anticipated recovery was tiny, and, more importantly, the process is skewed. My objection had to be postmarked before class counsel submitted their fee application. As a result, my objections were general, and, to the extent I objected to the fact that no fee application was filed, moot by the time of the hearing. I understand that, in their fee applications, counsel want to take objections into account, but that is not entirely fair to prospective objectors. The fee application should be filed before objections are due, and class counsel can address them at the hearing. Finally, objectors find themselves lonely because few class members will join them. As class counsel would have it, that reflects satisfaction with the deal, but I believe a measure of rational ignorance is at work. If class members see that they are getting something for nothing, and the process for objecting is not friendly to them, they will have little incentive to object.

In my judgment, lawyers who are members of plaintiff classes should consider objecting to awards that appear excessive. Lawyers have the experience to judge whether a requested award appears related to the recovery. So do institutional investors, several of whom objected to the attorney fee request in Xcel. Those institutional investors need to speak up and to choose their sides. In Xcel, several submitted written objections but none participated at the hearing. Furthermore, those institutional investors are frequently active plaintiffs, a fact which compromises their objections because they hire some of the class counsel who appear frequently. Still, lawyers and institutional investors have the ability to articulate serious objections.

Objecting may look like a futile act, but it is the only game in town. And it is the only way to be heard. Likewise, it is the only way of forcing the courts to carry out their responsibility to scrutinize proposed class action settlements.[31] By becoming the squeaky wheel, objectors may help to put limits on the operations of a class action system that needs them to further interests that are not theirs.

Endnotes

1 Pub. L. 109-2, § 2, Feb. 18, 2005, 119 STAT. 4. As codified, CAFA limits attorneys fees only with respect to settlements in which the class members receive coupons. 28 U.S.C. 1712. In such cases, CAFA links the recovery of attorneys fees to the value of the coupons redeemed. 28 U.S.C. 1712(a). When coupons are not redeemed, the award "shall be based upon the amount of time class counsel reasonably expended working on the action." 28 U.S.C. 1712(b)(1). "Nothing in this subsection shall be construed to prohibit application of a lodestar with a multiplier method of determining attorney's fees." 28 U.S.C. 1712(b)(2).

2 Rule 1.5(a)

3 *Id.*

4 Rule 1.5(c)

5 Comment 3.

6 *See* Ron Rotunda & John Dzienkowski, Legal Ethics: The Lawyer's

Deskbook on Professional Responsibility (2005), § 1.5-1(h), at 143 ("The lawyer may also consider his own experience, reputation, and ability. Lawyers who are twice as good as other lawyers are justified in charging twice as much as the other lawyers charge. A client can always conclude that the lawyer is not twice as good, and that client can go elsewhere for legal services.") Thus, we should expect better results from the lawyer or firm that charges a higher fee.

7 Healey v. Leavitt, 485 F. 3d 63, 69 (2d Cir. 2007).

8 A firm's compensation policy may discourage its lawyers from taking on time-consuming litigation predicated on the recovery of fees and expenses from the defendant or a common fund under a fee-shifting statute or legal doctrine.

9 Elihu Root famously observed that "about half the practice of the decent lawyer consists in telling would-be clients that they are damned fools and should stop." That observation reinforces the notion that the lawyer's expectations should be considered when evaluating of the reasonableness of a fee.

10 In its 2000 Annual Report, Xcel stated that, in 2000, its 82% share of NRG contributed $0.46 per share to Xcel's earnings, up from $0.17 on a 100% share in 1999, and that it planned to offer more of NRG's shares to the public. The 2000 results were, in part, attributable to favorable conditions in the wholesale power market, but Xcel expected that NRG would produce almost 25% of its earnings in 2001. 2000 Annual Report of Xcel Energy, in author's possession. A second public offering was made in 2001.In the 2001 Annual Report, Xcel stated that, as of December 31, 2001, it owned 74% of NRG. 2001 Annual Report of Xcel Energy, at 18, in author's possession.

11 At that time, NRG was rated BAA with Moody's and BBB with S & P. Investments bearing those ratings are generally medium-risk, investment-grade securities, but downgrading them would put them in junk status. *See* "What is a Corporate Credit Rating," *at* www.investopedia.com/articles/03.102203.asp (last viewed August 1, 2007).

12 *In re Xcel Energy, Inc., Sec., Derivative, & ERISA Litig.*, 286 F. Supp. 2d 1047, 1051 (D. Minn. 2003).

13 Historical pricing data available on <www.bigcharts.com> (last viewed May 7, 2007).

14 *Id.*

15 Ultimately, on May 14, 2003, NRG filed a voluntary Chapter 11 bankruptcy petition. NRG emerged from bankruptcy, as one might expect for a company with power and cash generating facilities that were overburdened with debt, in December 2003. As part of NRG's reorganization, Xcel relinquished its interest in NRG and paid an additional $752 million in exchange for a complete release from NRG and its creditors. 2004 Annual Report of Xcel Energy, at Note 4 to Consolidated Statement, in author's possession.

16 The district court denied the motion to dismiss the securities actions and dismissed the claims of the noteholders. *In re Xcel Energy, Inc. Securities, Derivative, and ERISA Litig.*, 286 F. Supp. 2d 1047 (D. Minn. 2003). The district court also dismissed the shareholder derivative action. *In re Xcel Energy, Inc.*, 222 F. R. D. 603 (D. Minn. 2004). Finally, it dismissed the ERISA claims in part. *In re Xcel Energy, Inc.*, 312 F. Supp. 2d 1165 (D. Minn. 2004).

17 The lodestar figure for the ERISA class counsel, who wanted an award of $8 million, was $867,611.75 with a multiplier of 2.16 required to turn it into $8 million.

18 *In re Xcel Energy*, 364 F. Supp. 2d 980 (D. Minn. 2005). The district court also awarded a small amount to the counsel for the derivatives plaintiff class, which had reached a settlement while their appeal from the dismissal of their claims was pending in the Eighth Circuit.

19 *See* Petrovic v. Amoco Oil Co., 200 F. 3d 1140, 1157 (8th Cir. 1999).

20 *See In re Network Associates, Inc., Sec. Litig.*, 76 F. Supp. 2d 1017, 1032 (N.D. Cal. 1999) (quoting Attorney William S. Lerach).

21 In *Roe v. Alabama*, a case involving the counting of absentee ballots in the 1994 election for Chief Justice of the Supreme Court of Alabama, one of the class representatives died in December 1994. Counsel soldiered on, without discovering the death of their client until August 1995.

22 Federal Rule of Civil Procedure 23(a) provides that, among other things, the class representatives must have claims that are typical of those of the unnamed class members, present questions of fact and law that are common to those of the rest of the class, and have no conflicting interests. In 2005, one serial plaintiff for a prominent securities class action plaintiffs' firm was indicted and accused of receiving more than $2.4 million for serving as the plaintiff in more than 50 securities class action lawsuits, and, in early 2006, another serial client of the same law firm admitted that he or members of his family were paid more than $2.4 million to act as plaintiffs. Payments like those alleged, which would not be shared with the unnamed class members, would align the lead plaintiffs' interests with those of class counsel rather than those of the class. *See generally* Margaret Little, *The Milberg Weiss Indictment*, Class Action Watch 3 (March 2007).

23 Theodore Eisenberg & Geoffrey Miller, Attorneys Fees in Class Action Settlements: An Empirical Study, New York University Center for Law an Business, Working Paper #CLB 03-23.

24 Have Attorney's Fees Risen in Class Action Settlements?, unedited transcript of program at American Enterprise Institute, February 20, 2004, *available at* www.aei.org/events/filter.all,eventID.753?transcript.asp (last viewed May 10, 2007).

25 *Id.* at 6.

26 Joint Declaration of Jack L. Chestnut and Sherrie R. Savett in Support of the Motion for FinalApproval of the Settlement and the Motion for Attorney Fees, Reimbursement of Expenses , and an Award to Lead Plaintiffs, at 69-70 (citing *In re Rite-Aid Corp. Sec. Litig.*, 396 F. 3d 294, 298, 303 (3d Cir. 2005). More recently, in *Vaughn v. American Honda Motor Co.*, No. 2-04CV-142 (TJW), in the United States District Court for the Eastern District of Texas, Marshall Division, a claim arising from alleged problems with the odometer on Honda vehicles that resulted in a small but improperly fast recording of mileage , class counsel submitted a declaration by Professor Charles Silver of the University of Texas School of Law, in which Professor Silver incorporated information about the sizes of multipliers employed that he received from Professor Eisenberg. *See* Plaintiffs' Application for Award of Attorneys Fees & Reimbursement of Expenses (No. 158) at Exhibit A, at 6-9.

27 Dillard v. City of Greensboro, 213 F. 3d 1347, 1355 (11th Cir. 2000).

28 *Id.* and 1355 n. 8.

29 In *Xcel*, counsel for the defendants agreed not to contest the fee application. Such "clear sailing" agreements are not illegal. *See, e.g.,* Waters v. Intern. Precious Metals Corp., 190 F. 3d 1291, 1293 n. 4 (11th Cir. 1999). As the court explained, such agreements give defendants a "more definite idea" of their total exposure. *Id.*, at 1292 n. 2 (citing Weinberger v. Great N. Nekoosa Corp., 925 F. 2d 518, 520 n.1 (1st Cir. 1991)). Such agreements, though, shift the burden to the court and to unnamed class members.

30 *Cf.* Alexander Tabarrok & Eric Helland, Two Cheers for Contingent Fees, 39, n. 44 (AEI Press 2005) ("Recent work by Helland and Klick has suggested that judges prove to be ineffective protectors of the plaintiff class, placing their own desire to clear the court's docket of the case before the interests of the nominal plaintiffs.") (citing Helland & Klick, "The Effect of Judicial Expediency on Attorney Fees in Class Actions" (FSU College of Law Research Paper No. 138, FSU College of Law, Law and Economics Paper No. 05-07)).

31 *Cf.* Reynolds v. Beneficial National Bank, 207 F. 3d 277, 279 (7th Cir. 2002) (The district court has a "judicial duty to protect the members of a class in class action litigation from lawyers for the class who may, in derogation of their professional and fiduciary obligations, place their pecuniary self-interest ahead of that of the class. This problem... requires district judges to exercise the highest degree of vigilance in scrutinizing proposed settlements of class actions.")

Religious Liberties

"Equal Access"? *Faith Center Church Evangelistic Ministries v. Glover* And the Use of Public Facilities for Religious Worship

*By Douglas G. Smith**

The U.S. Court of Appeals for the Ninth Circuit has developed a unique jurisprudence in cases involving religious liberty. From holding recitation of the pledge of allegiance in public schools unconstitutional[1] to invalidating the display of a cross erected as part of a war memorial,[2] the court has issued a series of rulings that push the envelope in addressing the constitutionality of various government policies implicating religious liberties. *Faith Center Church Evangelistic Ministries v. Glover* represents the latest in that line of controversial decisions.[3]

In *Faith Center*, a divided Ninth Circuit panel held that the district court abused its discretion when it found that a county library must give equal access to a Christian group seeking to utilize one of its public meeting rooms. The library offered these public meeting rooms for "educational, cultural and community related meetings, programs and activities."[4] However, it specifically prohibited certain religious activities, stating that the library's meeting rooms "shall not be used for religious services."[5]

The court did not dispute that the Christian group "engaged in protected speech when its participants met in the Antioch library for prayer, praise, and worship."[6] Nonetheless, the court held that the group could not engage in religious worship because the library meeting room was a "limited public forum" and "the County's policy to exclude religious worship services from the meeting room is reasonable in light of the forum's purpose."[7] However, the county specifically defined that purpose as excluding religious worship. The court's reasoning thus allows public entities to define away religious organizations' equal access rights by merely defining the "limited" forum as one that excludes certain religious practices—a result that is inconsistent with well-settled Supreme Court precedent.

The Supreme Court has unequivocally held that "religious worship and discussion... are forms of speech and association protected by the First Amendment."[8] "The Constitution forbids a State to enforce certain exclusions from a forum generally open to the public, even if it was not required to create the forum in the first place."[9] Accordingly, "speech discussing otherwise permissible subjects cannot be excluded from a limited public forum on the ground that the subject is discussed from a religious viewpoint."[10]

The *Faith Center* court conceded that, in opening its facilities to the public, "the County's purpose was to invite the community at large to participate in use of the meeting room[s] for expressive activity."[11] Nonetheless, it asserted that the library did not open its meeting rooms "for indiscriminate use" because it required the submission of an application that "must be reviewed and approved in advance" and specifically excluded "religious services."[12] The panel majority claimed, without any citation to the record, that the library's exclusionary policy was designed to "preserve the character of the forum as a common meeting space, an alternative to the community lecture hall, the corporate board-room, or the local Starbucks."[13]

However, excluding religious "worship" has nothing to do with "preserving" such characteristics of the space. To the contrary, religious worship inherently involves utilizing the meeting rooms as a "common meeting space." Moreover, even if the library had such an objective, as the panel majority conceded, the library may not "discriminate against a speaker's viewpoint."[14] Accordingly, the library cannot discriminate against religious organizations by defining the "limitation" on the public forum to specifically exclude those with a religious viewpoint.

The Supreme Court has made clear that a public entity may not exclude "religious worship" any more than it may exclude the promotion of atheism or libertarian philosophy.[15] Nonetheless, the panel majority allowed the library to do just that, relying upon a *dissent* issued by two justices to conclude that such limitations were "reasonable" and thus constitutional, asserting that it would be "remarkable" if a "'public [building] opened for civic meetings must be opened for use as a church, synagogue, or mosque.'"[16]

Nor can the county's discriminatory behavior be justified by an "interest in screening applications and excluding meeting room activities that may interfere with the library's primary function as a sanctuary for reading, writing, and quiet contemplation." The panel merely assumed that religious worship was "controversial" and "alienating," and that the library must have reasonably wanted to exclude it.[17] Whether "offensive" or not, worship remains protected by the First Amendment, as are controversial views generally.[18] By allowing the library to define the "limited forum" to exclude religious worship, and then claiming that the Christian group seeking to use the library's facilities "exceeded the boundaries of the library's limited forum," the Ninth Circuit significantly eroded the plaintiffs' equal access rights.[19]

The Supreme Court has repeatedly held that access may not be restricted if the restriction is based on "the specific motivating ideology or the opinion or perspective of the speaker."[20] In *Widmar v. Vincent*, for example, the Court ruled unconstitutional a policy that barred the use of university buildings "for purposes of religious worship or religious teaching" on the grounds that "[t]hese are forms of speech and association protected by the First Amendment."[21] In *Rosenberger v. Rector and Visitors of the University of Virginia*, the Court likewise held that the University of Virginia's policy of excluding religious publications from eligibility for student funds violated the Constitution because the University "select[ed] for disfavored treatment those student journalistic efforts with

....

** Douglas G. Smith is a Partner at Kirkland & Ellis, LLP and filed an amicus brief in support of petitioners seeking review of the Ninth Circuit's decision in the Faith Center case.*

religious editorial viewpoints."[22] In *Lamb's Chapel v. Center Moriches Union Free School District*, the Court held that a school district acted unconstitutionally when it opened its property for "social, civic, or recreational uses," but specifically prohibited its use for "religious purposes."[23] Finally, in *Good News Club v. Milford Central School*, the Court held that a school district engaged in impermissible viewpoint discrimination when it refused to allow a Christian children's club to offer a religious perspective on moral and character development in a school forum that was open to the public. [24]

These cases involved facts legally indistinguishable from those at issue in *Faith Center*. Nonetheless, the panel majority attempted to reconcile these decisions with its holding on the ground that a footnote in *Good News Club* allegedly drew a distinction between religious speech and "'mere religious worship, divorced from any teaching of moral values.'"[25] This argument plainly misreads the Court's opinion, which made no such distinction and did not authorize the exclusion of any particular forms of religious speech. Moreover, this reading is inconsistent with the Supreme Court's decisions as a whole, which have made clear that there is no such purported distinction.

In *Widmar*, for example, the Court specifically held that such a distinction had no "intelligible content." "There is no indication when 'singing hymns, reading scripture, and teaching biblical principles,' cease to be 'singing, teaching, and reading'—all apparently forms of 'speech,' despite their religious subject matter—and become unprotected 'worship'."[26] The Court further found that "even if the distinction drew an arguably principled line, it is highly doubtful that it would lie within the judicial competence to administer." "Merely to draw the distinction would require the university—and ultimately the courts—to inquire into the significance of words and practices to different religious faiths, and in varying circumstances by the same faith. Such inquiries would tend inevitably to entangle the State with religion in a manner forbidden by our cases."[27]

More fundamentally, the majority opinion lacks internal consistency. The court conceded, for example, that plaintiffs' "Wordshop" meeting, which included "fervent... [p]rayers," "teaching" and "singing", was permissible under the county's policy and that "*Good News Club* makes clear that such speech in furtherance of communicating an idea from a religious point of view cannot be grounds for exclusion."[28] Yet, at the same time, the panel asserted that the library properly excluded plaintiffs' "religious worship."[29] It is difficult to find a principled distinction between these activities. Indeed, standard definitions of "religious service" and "religious worship" make clear that such activities are merely a form of "prayer".[30]

In any event, the Supreme Court observed in *Widmar* that, if such a distinction could be made, discerning where it applied would impermissibly entangle the government with religion. Here the county would be faced with the "impossible task" of determining "which words and activities fall within 'religious worship and religious teaching.'" "There would also be a continuing need to monitor group meetings to ensure compliance with the rule."[31] Such entanglement is not merely undesirable—it is plainly prohibited under the Court's Establishment Clause jurisprudence. By contrast, "an open-forum policy, including nondiscrimination against religious speech… would in fact *avoid* entanglement with religion."[32]

As in *Rosenberger*, the religious exclusion the Ninth Circuit sanctioned in *Faith Center* is premised on a reading of the Establishment Clause that the Supreme Court has repeatedly rejected. Thus, for example, the panel majority asserted that religious worship is inherently "controversial" and "alienating" and that the government must exclude such conduct from public property because it is "not a secular activity."[33] Likewise, in a separate concurrence Judge Karlton criticized the Supreme Court's analysis in *Good News Club* and *Lamb's Chapel* on the ground that these decisions fail to recognize that the Establishment Clause creates a "wall of separation between church and state" that expressly prohibits any government role in religious life. Rather than guaranteeing religious liberty, he asserted that the First Amendment "serves the salutary purpose of insulating civil society from the excesses of the zealous" and lamented "[t]he *Good News Club* and *Lamb's Chapel* majorities' disdain of [this] Jefferson model."[34]

As a threshold matter, engaging in such analysis, the Ninth Circuit appears to have misunderstood the issue before it. As the Supreme Court observed in *Widmar*: "The question is not whether the creation of a religious forum would violate the Establishment Clause. The [library] has opened its facilities for use by [community] groups, and the question is whether it can now exclude groups because of the content of their speech." There is "no realistic danger" that the community would think the library "was endorsing religion or any particular creed" by allowing equal access to its facilities.[35] Any "benefit to religion or to the Church" would have been incidental. [36]

More fundamentally, the Constitution neither requires nor permits the government to expunge from public property all religious speech in order to ensure a purely "secular" forum. To the contrary, the Supreme Court has repeatedly held that the expression of religious viewpoints on public property does not offend the Constitution.[37] Indeed, the Constitution expressly *protects* the right to engage in such expression.[38] Accordingly, the Court has "rejected the position that the Establishment Clause even justifies, much less requires, a refusal to extend free speech rights to religious speakers who participate in broad-reaching government programs neutral in design."[39] Thus, the *Faith Center* decision is inconsistent with not only the Court's free speech jurisprudence, but also its interpretation of the Establishment Clause.

The *Faith Center* decision has already prompted a strong reaction. In dissenting from the Ninth Circuit's denial of rehearing en banc, seven judges of that court maintained that the *Faith Center* majority "disregarded equal-access cases stretching back nearly three decades, turned a blind eye to blatant viewpoint discrimination, and endorsed disparate treatment of different religious groups."[40] Only time will tell whether the Ninth Circuit's decision represents a permanent erosion in religious organizations' equal access rights, or merely an aberration in an otherwise well-settled area of law.

Endnotes

1 *See* Newdow v. United States Congress, 328 F.3d 466 (9th Cir. 2003), *rev'd*, Elk Grove Unified Sch. Dist. v. Newdow, 542 U.S. 1 (2004).

2 *See* Buono v. Norton, 371 F.3d 543 (9th Cir. 2004).

3 462 F.3d 1194 (9th Cir. 2006), cert. denied, -- S. Ct. --, 2007 WL 1668585 (U.S. Oct. 1, 2007).

4 *Id.* at 1198.

5 *Id.* at 1198-99.

6 *Id.* at 1202.

7 *Id.* at 1204.

8 Widmar v. Vincent, 454 U.S. 263, 269 (1981).

9 *Id.* at 267-68.

10 Good News Club v. Milford Central School, 533 U.S. 98, 112 (2001); *see also* Rosenberger v. Rector and Visitors of the University of Virginia, 515 U.S. 819, 828 (1995); Lamb's Chapel v. Center Moriches Union Free School District, 508 U.S. 384, 387, 393-94 (1993).

11 *Faith Center*, 462 F.3d at 1204.

12 *Id.* at 1205-06.

13 *Id.* at 1206.

14 *Id.*

15 *See Good News Club*, 533 U.S. at 112; *Rosenberger*, 515 U.S. at 828.

16 *Faith Center*, 462 F.3d at 1206 (quoting *Good News Club*, 533 U.S. at 139 (Souter and Ginsburg, JJ., dissenting)).

17 *Id.* at 1207.

18 *See* Police Dept. of the City of Chicago v. Mosley, 408 U.S. 92, 96 (1972) ("[G]overnment may not grant the use of a forum to people whose views it finds acceptable, but deny use to those wishing to express less favored or more controversial views.").

19 *Faith Center*, 462 F.3d at 1210.

20 *Rosenberger*, 515 U.S. at 829.

21 454 U.S. at 265 & n.3, 269.

22 515 U.S. at 831.

23 508 U.S. at 387, 393-94.

24 533 U.S. at 108.

25 *Faith Center*, 462 F.3d at 1209 (quoting *Good News Club*, 533 U.S. at 112 n.4). Instead of attempting to engage in the "laborious" effort to parse the Supreme Court's prior decisions to uphold the library's exclusionary policies, Judge Karlton simply disagreed with the Court's decisions in *Good News Club* and *Lamb's Chapel*, claiming that they misread the First Amendment because they do not recognize that "religious speech is categorically different than secular speech and is subject to analysis under the Establishment and Free Exercise Clause without regard to the jurisprudence of free speech." *Id.* at 1215 (Karlton, J., concurring).

26 454 U.S. at 270 n.6 (internal citation omitted).

27 *Id.*

28 *Faith Center*, 462 F.3d at 1210.

29 *Id.* at 1204.

30 *See, e.g.,* Merriam-Webster's Online Dictionary, *available at* http://www.m-w.com (defining "prayer" as "a religious service consisting chiefly of prayers"); Cambridge Online Dictionary, *available at* http://dictionary.cambridge.org (defining "pray" as "to speak to a god either privately or in a religious ceremony" and "worship" as "when you worship God or a god, often through praying or singing"); American Heritage Dictionary (4th ed. 2000) (defining "prayer" as "a reverent petition made to God, a god, or another object of worship").

31 454 U.S. at 272 n.11.

32 Board of Education of the Westside Community Schools v. Mergens By and Through Mergens, 496 U.S. 226, 248 (1990) (citing *Widmar*, 454 U.S. at 272 n.11) (emphasis in original).

33 *Faith Center*, 462 F.3d at 1210.

34 *Id.* at 1216 (Karlton, J., concurring); *see also id.* at 1215 (lamenting the "sorry state of the law" under the Supreme Court's decisions and "pray[ing] for the court's enlightenment").

35 *Lamb's Chapel*, 508 U.S. at 395.

36 *Id.*

37 *See, e.g., Good News Club*, 533 U.S. at 112-19; *Lamb's Chapel*, 508 U.S. at 394-96; *Widmar*, 454 U.S. at 271-75.

38 U.S. Const. amend. I ("Congress shall make no law respecting an establishment of religion, or *prohibiting the free exercise thereof*….") (emphasis added).

39 *Rosenberger*, 515 U.S. at 839; *see also Lamb's Chapel,* 508 U.S. at 393-94; *Mergens,* 496 U.S. at 248, 252; *Widmar,* 454 U.S. at 274-75.

40 Faith Center Church Evangelistic Ministries v. Glover, 480 F.3d 891, 895 (9th Cir. 2007) (Bybee, J., dissenting from denial of rehearing en banc). Despite this strong reaction, the Supreme Court denied Faith Center's petition for certiorari on October 1, 2007.

Telecommunications & Electronic Media

Will the Federal Communications Commission Broadcast Flag Order Be Resurrected?

*By Stephen T. Yelverton**

In order to protect digital television broadcast programming from mass piracy through the Internet, the Federal Communications Commission (FCC) adopted rules in 2003 to impose certain technical requirements on digital television receivers. This rulemaking, formally entitled *Digital Broadcast Content Protection*, is known as the "Broadcast Flag Order."[1]

A "broadcast flag" is a digital code embedded into a digital broadcast signal. It alerts digital television receivers to limit the indiscriminate copying or redistribution of digital broadcast programming. The FCC's Broadcast Flag Order required all makers of digital television receivers to design their equipment to recognize the broadcast flag embedded in the digital signal and to provide a mechanism to limit the indiscriminate copying or redistribution of the programming.

The FCC's rationale for the rulemaking was to enhance the transition from analog to digital television. It reasoned that the broadcast flag would benefit consumers by ensuring continued access to high-value programming content on free over-the-air television. The program producers had voiced fears that digital television would provide an easy opportunity for mass piracy. Therefore, they might withhold their programming from digital broadcasting and shift it to cable and satellite, which is technically less susceptible to piracy.

Over strong protests from consumer groups and thousands of objections from members of the public, the FCC adopted the rulemaking. Their concern was that the yet-to-be-implemented broadcast flag technology would violate the privacy of the television user, prevent copying for personal use the lawful "fair use" of copyrighted materials by educational institutions and libraries, and the lawful copying of non-copyrighted material such as news, public affairs, and political discourse.

The FCC set a deadline of July 1, 2005 for the makers of digital television receivers to comply with the new technical requirements to recognize broadcast flags. In an *Order* adopted on August 12, 2004, the FCC approved thirteen different technologies by various equipment makers to implement the broadcast flag protection for programming on digital television.[2]

Before the broadcast flag was implemented, the U.S. Court of Appeals for the D.C. Circuit struck down the FCC's rulemaking: *American Library Association v. Federal Communications Commission*.[3] In an opinion by Judges Edwards, Sentelle, and Rogers, the court ruled that the FCC had exceeded its statutory authority in regulating television receivers and the equipment makers. According to the court, Title I of the Communications Act of 1934 (codified under Title 47 of the U.S. Code) confers authority on the FCC to regulate "apparatus that can receive television broadcast content, but only while those apparatus are engaged in the process of receiving a television broadcast." In the case of the broadcast flag, the FCC's regulation of the receiver is *after* the completion of the broadcast and thus beyond its jurisdiction.

The court's ruling turned on its interpretation of Congressional intent in 47 U.S.C. § 153, which defines "radio and wire communications" to include not only the "transmission of… writings, signs, signals, pictures, and sounds" by aid of wire or radio, but also "all instrumentalities, facilities, apparatus, and services (among other things, the receipt, forwarding, and delivery of communications) incidental to such transmission." The court very narrowly construed 47 U.S.C. § 153 to limit FCC jurisdiction to the regulation of apparatus used for the receipt of radio or wire communications *only* "while those apparatus are engaged in communication."

However, the plain wording of 47 U.S.C. § 153 would belie such a strict construction. Its broad language also speaks of the "forwarding" of received communications (not just the receipt) and regulating apparatus "incidental" to transmissions. Thus, the broadcast flag digital television receivers would appear to be well within these jurisdictional parameters.

Albeit questionable reasoning, the court's ruling is a rare victory for strict constructionists who favor limiting the expansion of authority of the federal regulatory state and also for those who believe that federal courts should not take sides in policy disputes, but rather require Congress to decide—as envisioned by the Founding Fathers.

Congress is now deciding whether to resurrect the Broadcast Flag Order in the pending Communications, Consumer's Choice, and Broadband Deployment Act (S. 2686, which is the Senate version of H.R. 5252). Included in this proposed legislation is a provision authorizing the FCC to reinstate its broadcast flag rules for digital television. The bill also addresses the issue of flags for digital radio. While supported by Senators Ted Stevens (R-AK) and Gordon Smith (R-OR), there are serious concerns as to how to implement broadcast flags without infringing the rights of consumers.

Likewise, pending before the House is the Digital Content Security Act (H.R. 4569). This bill covers both digital and analog television receivers in preventing piracy and enlists the assistance of the Patent and Trademark Office. Also pending before the House is the Audio Broadcast Flag Licensing Act (H.R. 4861), which authorizes the FCC to implement flags for digital radio.

These bills have lukewarm support, but the conundrum for Congress is how to allow copying for personal and "fair" use, and for non-copyrighted material, while at the same time effectively protecting intellectual property rights on the Internet. A technology that can do all of this may not yet be available. Even with technology that does it all, Congress faces the conundrum of how to effectively enforce the broadcast

* *Stephen T. Yelverton practices before the FCC in broadcast matters and formerly served as a trial attorney in the Media Bureau.*

flag. The FCC rules had no real enforcement and the Agency has no real enforcement powers over television users. Many technically savvy consumers could likely disable the anti-piracy mechanisms in the broadcast receivers. What would be the penalty? Would there be fines, payments of copyright royalties, jail sentences? Would the federal courts be overwhelmed by a flood of violations? Would Congress risk voter wrath to mandate strict enforcement?

In sum, Hollywood has an uphill battle to get the needed protection for program producers against digital piracy. But even with new legislation, will the Internet eviscerate copyrights?

Endnotes

1 18 FCC Rcd 23550 (2003).

2 FCC 04-193.

3 406 F.3d 689 (D.C. Cir. 2005).

Book Reviews

David's Hammer: The Case For An Activist Judiciary

By Clint Bolick

*First Review by Timothy Sandefur**

*Second Review by Craig S. Lerner***

In all the recent debates over "judicial activism," Clint Bolick appears to be the only writer who begins by asking what the judiciary was actually designed to do. This is profoundly refreshing, given that both liberals and conservatives have embraced the slogan of "judicial restraint"— fashioned almost a century ago as part of the Progressives's ultimately successful effort to remold the Constitution in the service of the administrative state. The one thing everyone knows about judges today is that they are "legislating from the bench," imposing elitist liberal social theories on a captive populace in violation of democratic principle. Mark Levin's *Men in Black*, and Robert Bork's *The Tempting of America* are typical examples of this genre. Yet their interpretation reflects such a warped understanding of the Judiciary's constitutional role, it is impossible to address the issue without starting, as Bolick does, at the beginning: with the Founders' vision.

In *Federalist* 78, Hamilton explains that the judiciary will act as an intermediary between the people and their legislatures, to ensure that electoral majorities do not trample on liberty. Judicial independence, he explains, is "requisite to guard the Constitution and the rights of individuals" from the "ill humors" which sometimes lead officials into "dangerous innovations in the government, and serious oppressions of the minor party in the community."

Note the logical order that Hamilton assumes: individual rights are primary to, and the justification for, democratic decision-making. *Liberty,* not *democracy,* was his primary concern, and his colleagues agreed. The Declaration of Independence, for example, holds that all men are endowed with certain rights, which government is then instituted "to secure," and that when it violates those rights, the people may alter or abolish it. The Constitution, too, unambiguously declares liberty a "blessing"—yet it imposes powerful limits on democracy. The ontological order could not be more clear. As Jefferson put it, "an *elective despotism* was not the government we fought for; but one which should not only be founded on free principles, but in which the powers of government should be so divided and balanced… that no one could transcend their legal limits, without being effectually checked and constrained by the others."

It was the Progressives who inverted this scheme. They focused their attention on mechanisms of collective decision-making, which they saw as both the source and limit of freedom. As Justice Louis Brandeis put it, "rights of property and the liberty of the individual must be remolded, from time to time, to meet the changing needs of society." You have got to love that euphemism: "remold." The Progressives forged the idea of "judicial restraint" as a way of allowing legislatures the broadest possible discretion to "remold" individual rights at will and to control realms of life that the Founders had considered off-limits. In 1934, with the adoption of "rational basis scrutiny," the Progressives won the day. Under this theory, courts simply look the other way when legislatures trample on individual rights.

But what is remarkable is the way that alleged opponents of such chicanery—including Bork—have embraced these heresies. Notwithstanding his protestations, Bork's critique of the judiciary is not rooted in the Founders' views but in the views of the Progressives. Madison held that "the sovereignty of the society as vested in & exercisable by the majority, may do anything that could be *rightfully* done by the unanimous concurrence of the members; the reserved rights of individuals (of conscience for example) in becoming parties to the original compact being beyond the legitimate reach of sovereignty." Bork, however, like his Progressive forebears, sees society as primary, the rights of individuals as secondary: as basically permissions granted by the majority, retractable when it chooses. Thus he invokes what he absurdly calls a "Madisonian" proposition: that "in wide areas of life majorities are entitled to rule, if they wish, simply because they are majorities." In fact, Madison and his contemporaries believed majorities were never *entitled* to rule: they were instead *authorized* to rule, and there were natural moral limits to that authority—limits which the judiciary, like the other branches of government, ought to enforce.

Where Bork, Levin, and other writers on "judicial activism" have ignored these matters entirely, Bolick begins with a welcome refresher on constitutional theory, supported by examples which prove his assertion that "the gravest deprivations of liberty have occurred not when the courts have exercised too much power but when they have exercised too much restraint." Indeed, a list of "great moments in judicial humility" would have to include such abominations as *Plessy v. Ferguson, Buck v. Bell, Korematsu v. United States,* and *Kelo v. New London*—all cases in which courts looked the other way, and allowed legislative majorities to act at will. (And it is typical of the murky thinking on this subject that Levin calls *Plessy* and *Korematsu* "activist" decisions. As Bolick points out, the Court in these cases *deferred* to the elected branches!)

By way of contrast, Bolick provides a series of examples from the case files of the Institute for Justice, in which courts have done their duty to enforce the Constitution, even when that meant stopping legislatures from serving the prejudices of the moment. There are three important lessons to be drawn from such cases.

First, it is legislative activism—not judicial activism—that presents the gravest threat to Americans generally. Certainly there are cases of judicial overreaching—outlandish ones, in which judges have gone beyond even the boundaries of reason—but they are insignificant compared to the mountains of unconstitutional and oppressive legislation that oozes out daily from under the domes of fifty state capitols and

* *Timothy Sandefur is a Senior Staff Attorney at the Pacific Legal Foundation. His article "The Wolves and the Sheep of Constitutional Law: A Review Essay on Kermit Roosevelt's The Myth of Judicial Activism" appeared recently in the Journal of Law And Politics.*

Washington, D.C. This is not even to mention the reams of regulation propounded by unelected lifetime bureaucrats in countless barely-known agencies, from the California Raisin Administrative Committee to the Migratory Bird Conservation Commission. It is by enforcing, not by nullifying, the orders of these thousand petty Caesars, that Americans lose their freedom.

Moreover, a judicial decision that wrongly upholds a law is far more dangerous than a decision that wrongly strikes one down. When a court declares a law unconstitutional, that decision leaves the legislature free to choose any number of alternative ways to accomplish a legitimate goal. But when the court upholds an unconstitutional law, there is practically no way to fix the damage done. The challenger is subjected to an irreparable violation of her freedom, legislatures go on to enact follow-on legislation, and dissenters are smugly told that if they want to change things, they must use the political process—a process that in reality is dominated by hostile majorities, safe incumbents, unaccountable regulatory agencies, and labyrinthine campaign finance laws that make grassroots political activism virtually impossible.

Second, deference often *is* activism. Consider an example not mentioned in Bolick's book: *Guinn v. Legislature of Nevada*. In that case, the Nevada Supreme Court ordered the state legislature simply to "disregard" the state constitution's requirement that tax increases receive a two-thirds vote of each house. The court explained that it was "concerned with the interest of preserving the democratic process," and so, because "a majority of legislators" had voted in favor of the tax increase, the court would simply set aside the two-thirds requirement just this once. Although this lawless decision was overruled only a few years later, *Guinn* stands as a stark reminder that some of the worst incidents of "judicial activism" occur when courts stand back and allow legislatures free rein.

Obviously, the judiciary has the power to manipulate the political process unfairly—and it has often abused its power. Recent cases like *Kelo, Atkins v. Virginia,* or *Roper v. Simmons,* illustrate that all too well. But the label of "judicial activism" is inadequate to address the real problem with such cases. Judges *ought* to be "active" when enforcing the Constitution. That is why they take an oath to support and defend it—not passively to allow legislators to do what they please. Courts ought only to defer when the legislature acts within its legitimate boundaries. But the choice of when to act and when to defer must be guided by an understanding of the Constitution—and that is a task for political philosophy. It is the betrayal of the Constitution's philosophy that is the real heart of our problem. To dismiss such questions as arbitrary, outdated, or matters for majorities to settle—or, worse, to say that majorities "are entitled to rule simply because they are majorities"—is to betray not only the Constitution's framers, but the purpose of the judiciary, and the rule of law itself.

It is therefore no surprise that we find the very concept of judicial review attacked by Bork and Levin, who characterize *Marbury v. Madison* as an "activist," politically motivated decision. Theirs is the natural conclusion of a political philosophy that puts majority will—and not individual freedom—at its center. A true democracy has no use for judges, or even for a constitution, which will simply get in the way of the majority's desires.

Thus the third lesson to take from Bolick's book is, once again, that liberty must come before democracy. It is because we are naturally free that we are entitled to participate in a constitutional order, and that it is wise for us to do so. It is because we have liberties that need protection from others that we employ a rule of law—and it is because those liberties are as vulnerable to the state as to our fellow citizens that we impose law on the government itself, through a Constitution. As Madison wrote, "In a society under the forms of which the stronger faction can readily unite and oppress the weaker, anarchy may as truly be said to reign as in a state of nature, where the weaker individual is not secured against the violence of the stronger." Preserving individual rights—both enumerated and unenumerated—from wrongful and impassioned majorities is the proper role for our judiciary. To countenance their violation, whether actively or passively, is to do violence to those blessings of liberty that our Founders preserved for us.

Clint Bolick is an asset to the nation. He and the Institute for Justice, the organization he helped found, have rescued Americans from the entangling web of government regulations time and again. One must admire a man who has entered the arena, battled the odds, and often emerged victorious. But this is all by way of introduction. For present purposes, he is a laudable man who has written a misguided book.

Bolick weaves the David-and-Goliath metaphor through *David's Hammer: The Case for an Activist Judiciary*. The first false note is the title. The Biblical David's preferred weapon was the slingshot. The hammer would have been far too direct and open-handed for the famously devious and underhanded David. Reading this book had one odd result: I found my sympathy swelling for the much-maligned Goliath, a point to which I will return.

Bolick's David has three senses. First, there is the humble citizen denied a fair chance by government bureaucrats, working in tandem with, or at the behest of, labor unions and corporations. Bolick is at his best capturing the plight of students seeking to escape monopolistic government schools and would-be entrepreneurs frustrated by special interest laws prescribing the number of taxicabs or the licensing of hairdressers.

Problems arise when we turn to the second sense of David, which is the "public-interest" litigator: constitutional lawyer as "hero." Bolick writes that he shares the public's "disdain" for the broader (for-profit) legal profession, which he characterizes as a "mercenary profession." But when lawyers help people manage their affairs in a private and peaceful way—*e.g.,* resolving contract disputes, forming corporations, navigating through mazes of government regulations—they are performing a public service. That they are paid for their efforts does not negate this fact. As Adam Smith long ago observed, a dollop of enlightened self-interest can go a long way toward promoting the public good.

*** Craig S. Lerner is a Professor at the George Mason University School of Law.*

Bolick is intoxicated by the vision of the "creative lawyer [who] can change the world in one fell swoop," but he glosses over the mischief sometimes committed by lawyers grandiosely pursuing the public interest. Touting the "black and white" results that can be obtained through litigation, Bolick exudes impatience with modern American democracy and its high levels of voter ignorance and low levels of participation. He complains that "[m]uch of government power at every level today is exercised by unelected agencies and authorities that have only the barest democratic accountability."

But Bolick's own experiences suggest that the corruption is not that endemic. Although his efforts to overturn government regulations have met with mixed success in the courts, the supposedly apathetic public was stirred to act, with resulting triumphs in city councils and state legislatures. The school choice programs Bolick has so effectively promoted were enacted through the democratic process, only then to be imperiled in the courts. "Public-interest lawyers" at the NAACP Legal Defense Fund, ACLU, and elsewhere have been Bolick's antagonists in these cases and others throughout his career.

It is, however, when we come to the third sense of David-and-Goliath that the book becomes most problematic. On the cover of the book is a blurred gavel crashing into a wooden block. The judge, seemingly, is David and the gavel is his hammer. According to Bolick, "[j]udges should see themselves as what they are intended to be: fearless guardians of individual liberty."

Bolick criticizes meek judges who cling to a "presumption of constitutionality" when considering challenges to government policies. Like Randy Barnett, Bolick argues that the very opposite presumption, the "presumption of liberty," is rooted in the Ninth and Tenth Amendments, as well as in the Privileges or Immunities Clause of the Fourteenth Amendment.

Claiming Clarence Thomas as his model Supreme Court Justice, Bolick adopts an interpretative method that emphasizes textualism and originalism. With respect to the Constitution's Commerce Clause, for example, Bolick convincingly shows how flawed the modern jurisprudence is in failing to pay careful attention to the text and original meaning of this congressional power. Yet Bolick charges conservatives, and even Justice Thomas to some extent, with being fair-weather textualists and orginalists. "[I]f conservatives want courts to protect economic liberties," he writes, "they must accept that the courts must protect the rights of people to engage in nonharmful, consensual activities in their own homes and of political dissenters to burn the American flag."

On the one hand, Bolick concedes that there is such a category as bad judicial activism. Most of his adduced examples, however, are instances of in-activism—that is, a judicial failure to invalidate a law or executive action. Notably missing from his roll of dishonor is *Roe v. Wade.* On the other hand, his catalog of "good judicial activism" includes *Lawrence v. Texas.* Bolick contends that there is no "principled basis" upon which conservatives could uphold the right of an organization to exclude homosexuals (as in *Boy Scouts v. Dale*) and then uphold a state law criminalizing homosexual sodomy (as in *Lawrence*). As Bolick sees it, "freedom of association" means that "a zone of privacy exists into which the government may not intrude." Not surprisingly, then, Thomas, who dissented in *Lawrence*, fades as a model justice in the book's conclusion, with a new hero of judicial activism emerging: "[P]atrolling the Court's libertarian center is Justice Anthony Kennedy, who has shown himself in many cases to be extremely mindful of the need for an activist judiciary to protect individual rights."

Bolick reserves special disdain for Robert Bork, the famous critic of judicial activism. Bork's arguments are intermittently mocked: he is an "absolutist" and an advocate of "judicial castration;" his views are "radical" and even "chilling." By contrast, Jonathan Rauch, Jeffrey Rosen, and Cass Sunstein are all described at various points as "thoughtful." They often are, of course. But so is Bork. Unfortunately, Bolick does not meaningfully engage Bork or the many thoughtful conservative scholars with whom he disagrees.

For starters, although Bolick taxes others with inconsistency in their constitutional interpretations, he might want to storm-proof his own glass house. The close attention to actual language and historical understandings that does credit to his interpretation of the Commerce Clause is jettisoned when Bolick turns to favored constitutional provisions. Bolick stretches the Ninth Amendment ("The enumeration in the Constitution, of certain rights, shall not be construed to deny or disparage others retained by the people.") to create a judicially enforceable right to privacy, extending to contraceptives and homosexual sodomy. A more natural, textually rooted interpretation of the Ninth Amendment, especially when viewed in tandem with the Tenth Amendment's reservation of powers to the states, has been suggested by Nelson Lund, drawing upon the historical work of Kurt Lash. Lund has argued that under the Ninth Amendment, "the federal government (including the federal judiciary) lacks the authority to impose its views of natural or inherent rights on the states: one of the rights 'retained by the people' is the right to govern themselves as they see fit in their own states."

With respect to the Privileges or Immunities Clause, Bolick confidently opines that "any believer in original intent who devotes even the most cursory attention to legislative history and to the problem Congress sought to correct will conclude that Congress unambiguously meant to protect economic liberty against excessive state regulation." Bolick may be unfamiliar with the work of originalists such as David Currie and John Harrison. They have argued that the Privileges or Immunities Clause is an anti-discrimination provision, not a substantive guarantee. It prevents states from discriminating against certain classes of citizens in the same way that the privileges and immunities clause of Article IV prevents states from discriminating against non-citizens.

Bolick's grand theory of constitutional design—the "presumption of liberty"—has at best a tenuous relationship with the Constitution's text. A more nuanced, textually rooted rule of presumption is suggested by Steven Calabresi and Gary Lawson. Given that the federal government is vested with enumerated powers, whereas the state governments enjoy general jurisdiction, perhaps one should say that federal courts should apply a presumption of liberty when reviewing federal laws, but a presumption of constitutionality when reviewing state laws. More generally, Bolick's advocacy of a muscular

Supreme Court strains the Constitution beyond recognition. As Calabresi has written, "There is simply no way to read the bare-bones language of Article III, in contrast to the detailed language of Article I, and conclude that the Framers meant for the Court to be a powerful institution."

The thorniest problem raised by *David's Hammer* is its treatment of the question of precedent. Supreme Court decisions in many areas of law have strayed so far from the Constitution's text that there is a question of what can or should be done to redress the matter. Bolick says: Follow the text and ignore the precedent. There are respectable arguments for this point of view, and Justice Thomas has been a forceful advocate. "Justice Thomas exhibits several qualities that make him the type of justice the Framers must have had in mind when they invested the judiciary with its central role in protecting freedom," Bolick writes. "In almost every constitutional case, he begins by examining not the Court's precedents but the language and intent of the Constitution itself."

It would strengthen Bolick's argument if he acknowledged that the issue is a complicated one. It is true that Article VI, section 2, states that "The Constitution, and the Laws of the United States [are] the supreme Law of the Land," which suggests that judges should continually repair to the text of the Constitution. But Article III places the "judicial Power" in the federal courts. And "judicial power" has long included a respect for precedent. Hamilton's *Federalist* No. 78 notes that "[t]o avoid an arbitrary discretion in the courts, it is indispensable that they should be bound down by strict rules and precedents." Given that the number of precedents will grow to a "very considerable bulk," Hamilton suggests that only those who have committed themselves to "long and laborious study" will be eligible for service on the federal judiciary. The image of a judge that emerges from *Federalist* No. 78 is a far cry from Bolick's hero. Hamilton's judge is a cautious old man, worn down by all those mind-numbing years immersed in the study of precedents. A few more such humble judges might serve America well.

Libertarian judicial activists are oddly confident that swashbuckling judges will enact their policy preferences. As Nelson Lund and John McGinnis have argued, this may be true with respect to sexual matters, but there is reason to doubt that judges in the future will be widely using the Contracts Clause to invalidate rent control laws, or the Privileges or Immunities Clause to overturn government regulations. It is more likely that judges, encouraged to view themselves as "fearless guardians of individual liberty," will see in a distended Ninth Amendment the sort of "rights" that inevitably herald an expansion of government power—a right to quality education, a right to health care, etc.

Which brings me back to the sad case of Goliath. If Bolick can write a book eulogizing judicial activism, I can manage a (tongue-in-cheek) paragraph in defense of this misunderstood giant. The Israelites won the war and got to write the definitive history, according to which he was a nine-foot ogre, but let us imagine this from the Philistine perspective. Goliath displayed himself openly and was prepared to fight honorably. David was a sneak and a weakling who could not win through honorable means and so employed deceit (which is why Machiavelli was such a fan). Those who seek to overturn laws through judicial activism can occasionally be likened to David, but not in the flattering sense that Bolick intends. Unable to win directly and honorably, *through the political process*, they have opted for an underhanded alternative. Perhaps libertarians, and all Americans, should fight in the tradition of Goliath, openly and honorably—in the state legislatures. After all, the Constitution created a framework of competitive federalism that resembles Bolick's beloved free market far more closely than the activist judiciary held out as the nation's potential messiah.

Silence and Freedom

By Louis Michael Seidman

*Reviewed by Paul Horwitz**

"Silence," A.A. Attanasio wrote, "is a text easy to misread." It is all the more brave and impressive, then, that Louis Michael Seidman has undertaken a project that places silence at its very heart. As Seidman observes, somewhat paradoxically, ever since *Miranda v. Arizona,* the words "You have the right to remain silent" have become the most famous in our popular constitutional culture. And yet, "What a strange right this is. Of all the activities that are especially worthy of human beings… why privilege silence?"

Seidman makes two basic claims in this relatively short, but wide-ranging, volume. First, he argues, silence can be a liberating "expression of freedom." It supplies meaning when our clumsy tools of language run out. And when silence is an active refusal to speak or act in the face of demands that we do so, it is something more than an absence: it is an act of defiance. Thomas More, the King's good servant, but God's first, never spoke more loudly than when he refused to acknowledge the King's marriage, on pain of his own doom. Second, Seidman claims we must protect silence "in order to give meaning to speech." While silence sometimes is a freedom worth preserving for its own sake, at other times "it is the necessary frame for freedom…. It is [ ] important to remain silent when there is nothing to say. When one confronts an ineffable mystery, breaking a silence only brings speech into disrepute."

Such oracular language certainly lets us know that we are not in for a typical doctrinal monograph. Seidman offers more of a meditation on the nature of silence and its place within the law, ranging from the mundane precincts of the police station interrogation room to the hushed mysteries of the end of life. One would do wrong to look for definitive answers to the questions he poses. As he warns, "Readers who like the hard edges of legal argument and have no taste for paradox are bound to be disappointed." For those with a taste for a more catholic and tentative journey, however, he is a faithful and careful guide, and offers what resolutions he can.

To examine a subject that is, well, *silent* is a difficult task. Seidman thus proceeds like a scientist attempting to observe a black hole: he applies his observational toolkit to what cannot be seen directly in order to perceive it by indirection—to detect

** Paul Horwitz is an Associate Professor at the University of Alabama School of Law.*

the very absence which signals its presence as a "void" that nevertheless "form[s] the core of a basic human right."

The toolkit is a simple one. Seidman examines the scope and limits of the legal right to silence through the lens of four political and philosophical concepts. The republican position emphasizes "deliberation about the common welfare as a common good," and sees active deliberation as the key to sound self-governance. With their encouragement of active public participation, republicans "are more comfortable defending the opposite of silence—that is, speech." To this he contrasts liberal thought, with its emphasis on autonomy, and its refusal to privilege the universalist aims of the state over a variety of personal commitments—to a faith, a community, or a set of individual values and beliefs. Liberals are more likely to champion silence, as "an absence that individuals can fill with plans for their own lives."

Along a rather different dimension, he plies silence with two more schools of thought, more expressions of skepticism than positive programs. First, pervasive determinists argue against the very notion of a genuine freedom to choose, emphasizing instead the psychological forces and power relations that constrain or deny human agency. For determinists, it is hard to talk about a right to silence "if all of human conduct can be reduced to unchosen or unconscious manipulation of others." Yet a determinist may ultimately see silence as the only "choice" left open to us. Finally, radical libertarians believe that none of these answers finally fills the void of meaning. In that void, we are left with nothing but the absolute freedom to choose. While a radical libertarian cannot champion a right to silence as such, he may believe that "it is better to remain silent than to attempt to fool others into believing that anything we say will require a particular choice."

These concepts are mere sketches, and somewhat thinly drawn ones at that. Those who work frequently with concepts like liberal and republican thought may find them but dimly represented here; and certainly it is true that most of us act with some combination of all of these concepts in mind. But Seidman does not pretend otherwise. These concepts are simply heuristic devices, tools by which he can examine some of the paradoxes and contradictions inherent in standard arguments for the right to silence.

Thus, Seidman argues that the growing movement to provide some safe harbor for a right of apology in the law, whether in civil or criminal settings, is beset by a host of contradictions when examined more closely. Republicans might favor the legal use of apologies, since an apology is a form of active engagement with others. But once we treat an apology as a legally significant fact, with mitigating consequences for the apologizer, we allow insincere and self-serving motives to enter the picture. Thus, Seidman observes, "there is a sense in which counting apologies as a mitigating circumstance destroys the possibility of apology." Similarly, liberals might be said to favor a right to withhold an apology, on the ground that such actions are quintessentially private and should not be subject to the coercive pressures of public life. But some apologies are quite genuine; if we do not honor them in the public sphere, we may end up discouraging praiseworthy private choices. Similar tensions are unearthed when Seidman applies the worldview of the determinists and the libertarians to the legal treatment of apology. Although he ultimately finds that the right to withhold an apology—the right to remain silent in the face of one's own wrongdoing—is the best outcome, he suggests that the path to that conclusion is strewn with doubt and contradiction.

Seidman works a similar legerdemain across a range of other topics involving silence and freedom. In two chapters, he reaches a surprising set of conclusions about the Constitution's most explicit treatment of silence: the right against self-incrimination. He questions whether the use of that right in courtrooms and other formal proceedings genuinely serves human freedom. Indeed, he argues, too much attention has been paid to the right against self-incrimination in formal settings, and not enough to its role in police interrogation. Here, too, he applies his acid bath to the received wisdom, arguing that the usual question courts ask, whether a suspect's statements are voluntary, is less important than the question of *how* police interrogate suspects. Here, his concern is that "police interrogative techniques invade a protected private sphere by abusing intimacy and illusions of intimacy." He proposes a tough remedy: permitting police to formally apply to a judge for an order requiring the suspect's cooperation, and holding the suspect in contempt if he resists.

In another chapter, Seidman skillfully dissects the Supreme Court's confused doctrine regarding compelled speech and its reverse, the right to silence in the face of compulsions to speak. Elsewhere, he suggests that we must enjoy some right to choose death, the ultimate silence. Outside the self-incrimination chapters, however, perhaps his most elegant performance is his treatment of torture. The problem with torture, he suggests, is not simply that an individual is compelled by force to reveal something; the state often coerces information from unwilling individuals. It is something deeper. Physical torture strips from us the illusion of choice and intellect, reducing us to nothing more than frail and mortal bodies. Torture thus removes our ability to maintain the sustaining illusion of human agency, and to remain silent about the gross physical nature of our existence. Seidman closes by urging us "to end our silence about torture's terrifying truth. We need to understand torture and all that it tells us about ourselves, rather than simply outlaw it."

It goes without saying that such passages do not provide an easy route to legal reform. This is just not that sort of book, although his chapter on free speech makes clear Seidman's skill at such conventional exercises. To venture one modest criticism, it seems to me that Seidman focuses too much on the individual who is subject to compulsions to speak or remain silent, and not enough on the entity that makes those demands: the state. He does address this question from time to time, but much more could be said about the question of when and whether the state may seek to compel speech or silence. What does it mean for the state, in the criminal context, to seek to elicit the thoughts of its citizens? Is the state's only concern about the use of torture purely instrumental, or does it fundamentally distort or degrade the moral legitimacy of the state when it attempts to pry the truth from a human body? Is it appropriate for the state to speak through its citizens' mouths, whether through compelled patriotic exercises or through such trivialities as the use of the motto "Live Free or Die" on a license plate? Such

questions may tell us relatively little about silence itself. But they might force us to think harder about the nature of a state that seeks to compel silence, or to break it.

This cavil notwithstanding, Seidman offers a thoughtful, and thought-provoking, exploration of the questions raised by the "right" to silence. Although he is often persuasive, he does not intend, I think, to push us to concrete conclusions. Rather, he hopes to persuade us to "reassert and think carefully about the value of silence." In this he is eminently successful.

The Constitution's Text in Foreign Affairs

By Michael D. Ramsey

*Reviewed by Jeremy Rabkin**

Michael Ramsey's new book is not likely to become a best seller. The book is probably too scrupulous in its scholarship for those whose interest in these constitutional questions is entangled in political debates of the day. Professor Ramsey, who teaches at San Diego University School of Law, seems to have no larger agenda than establishing the historical truth, as clearly as he can discern it. And truth, whatever its ultimate strength, does not always have a large market.

Two decades from now, however, when partisans have moved on to new claims, and most of today's big books are relegated to remote library annexes, serious scholars will still be consulting Ramsey. We are not likely to see *The Constitution's Text* displaced any time soon by a more penetrating or exhaustive set of historical inquiries on the subjects it covers.

As the title suggests, the book is an exercise in recovering the "original understanding" of constitutional provisions relating to the conduct of foreign affairs. It is not the kind of account a professional historian might offer with a chronological narrative of "formation" or "development." *The Constitution's Text* is very much a lawyer's history, in the sense that it is organized around claims advanced by lawyers or judges in recent decades, invoking specific constitutional provisions. Almost every chapter begins with a prominent Supreme Court ruling, and then measures the Court's assumptions against the intentions or expectations of the Founders, as Ramsey has reconstructed them.

And, on the whole, Ramsey's scholarship demonstrates that the Supreme Court did not do very well in the twentieth century in expounding the Constitution if their work is judged by the understandings of its Framers. Yet the book is not primarily a polemical debunking. Among other things, the Supreme Court itself has seemed to endorse quite different interpretations in different cases. Correcting the Court's errors may keep scholars busy, but it does not force an originalist critic into relentless reproaches against the same mistakes. Ramsey gets to look at quite a number of different issues, because the Court has made so many—and, sometimes, mutually contrary—mistakes.

Quite apart from what it contributes to specific debates regarding foreign affairs powers, the book makes a valuable addition to the literature on originalism. It does so by demonstrating that serious inquiry can help to reconstruct a quite compelling account of "original meaning," even when it comes to the clauses of the Constitution dealing with the conduct of foreign policy. That, in itself, is a notable achievement.

After all, the scattered clauses in the Constitution that make some passing reference to the conduct of foreign policy do not stand out in bold headings. For the most part they seem incidental, elliptical, perhaps even evasive. Influential commentators, such as Professor Louis Henkin, chief reporter for *the Restatement (Third) of Foreign Relations Law*, have emphasized the "strange laconic" character of constitutional provisions in this field, requiring subsequent generations (as commentators urge) to rely on imagination or experience, more than textual exegesis. Others, starting with Edward Corwin, the Princeton scholar who began what is today's *Annotated Constitution*, characterize balancing provisions in this area as "an invitation to struggle" among the different branches of government for the dominant say in foreign affairs. Historians, in fact, trace the origins of the first party system, pitting "Federalists" (or "Hamiltonians") against "Republicans" (or "Jeffersonians"), to disputes about foreign policy in President Washington's second term.

What Ramsey shows is that, if we look closely at contemporary sources, we can reconstruct reasonably definite understandings of each claim—definite enough to endorse some interpretations and exclude others. These interpretations seem to have been generally accepted at the time. Even Jefferson (as Secretary of State) agreed that the President had exclusive authority in articulating American policy to foreign governments, while Hamilton acknowledged that inherent presidential authority did not extend to ordering interference with domestic commerce without a statutory authorization. The accepted understandings add up to a framework that may seem, even today, sufficiently coherent and reasonable that we can feel some confidence in attributing them—as a shared, conscious, considered structure—to the understanding of the Framers as a collective entity.

Ramsey's method is to start with a question raised by subsequent litigation, such as the scope (or existence) of inherent executive powers, look at what was said at the Philadelphia Convention, what was endorsed or rejected in the course of the Convention, how this did or did not parallel arrangements under the Articles of Confederation or the British Constitution, and how these choices echoed (or repudiated) doctrines set out in the major "authorities" of the day—Blackstone, Montequieu, Locke, etc. Then, he compares these "literary" sources with actual practice in the first few administrations under the Constitution, and what was or was not accepted at that time as proper.

The Constitution's Text proceeds in this way through six broad topics: the general source of national authority in foreign affairs, the specific powers of the President in this area, the powers of the Senate, the powers of Congress, the powers of the states, the powers of the courts. As there are several chapters under each heading (for a total of eighteen), the exposition could easily have become wearisomely repetitive or

* *Jeremy Rabkin is a Professor of Law at George Mason University.*

suspiciously selective. But Ramsey is quite skillful in navigating the literary challenges of his chosen approach. He offers enough background to illustrate the argument, but does not stray too far from the specific interpretive query he is pursuing in any one chapter.

Commonly, the chapters begin with a proposition advanced in a famous Supreme Court ruling; then show why the evidence makes such propositions untenable. Ramsey tries to build confidence in alternative claims by showing all the evidence there is for the alternate view in the different sources. Along the way, he articulates possible objections to the counter-theses he offers, while showing why these objections have less force than at first might appear. The back-and-forth style of exposition might remind some readers of the objections and responses in Thomas Aquinas' *Summa*. But perhaps it echoes the rhythm of claims and rejoinders in legal pleadings. It certainly works through competing arguments with great care.

The historical record inspires confidence in the conclusions Ramsey reaches. Both in what they said and did (and, in some cases, what they did or did not oppose), there was considerable agreement among statesmen of the Founding era on these questions. When it comes to precise questions regarding precise clauses or powers, Hamilton and Jefferson agreed much more than they disagreed—which is perhaps why they did not regard "interpretation" as quite such a hopeless undertaking as so many scholars do today. Where the giants are in agreement, scholars of today should find it easier to accept their authority; and, seeing so many areas of agreement, we ought to be suspicious of claims that the record is too confused or contradictory to offer any guidance for disputes in our day.

Ramsey himself seems prepared to follow the trail of evidence where it leads him, without much concern about whether this aligns him with today's counterparts of Hamiltonians or Jeffersonians—or let us say, Bush defenders or Bush critics, partisans of the executive, Congress or the courts. So, for example, he rejects claims (as in Justice Black's opinion in the "Steel Seizure" case) that the Constitution offers no inherent powers to the President in foreign affairs. But he also rejects the claim, recently advanced by John Yoo, that the assignment to Congress of the power to "declare war" was understood merely as a legal formality, which still left the President with broad power to initiate military operations on his own authority. When it comes to initiating war, Ramsey argues for a restricted view of presidential authority. Similarly, he disputes the notion that the President may disregard international law. But he also rejects the claim that courts have wide authority to overrule executive interpretations of what international law requires of the United States in any particular dispute.

Ramsey sees a role for courts in upholding individual rights, even when they derive from treaties, and argues for a more restricted understanding of what it means for a treaty to be "non-self-executing" than that favored by contemporary champions of executive authority. But the role he sketches for courts remains constrained. In Ramsey's version of the Constitution, then, genuine and important presidential powers are balanced by powers retained by the Senate in certain areas, by Congress as a whole, by the states and, in some cases, by courts. It is, as he says, a structure of distributed and assigned powers—a counterpart to the constitutional architecture assumed in so many domestic areas.

The principal disappointment of this book is that it adheres so closely to its own focus—in reconstructing the original understanding of particular textual provisions—that it does not venture any reflections on why case law and practice have so often crossed lines which the Founders tried to draw. The last lines of the book affirm that the "founders' Constitution in foreign affairs… can guide us through modern controversies. That does not mean that they should. We may conclude here, only, that they can." And by "can," Ramsey seems to mean only that we can recover enough about original meaning to understand what was intended. He does not try to argue that our foreign policy would be as effective or safely conducted if we adhered more closely to the boundaries the Founder's had in mind—rather than concentrating more authority in the national government, more initiative in the President and giving more overall lee-way to arrangements Congress has found most convenient.

An historical inquiry need not, of course, culminate in a philosophic defense of a particular interpretative method, or even a particular philosophy of historical development. Still, defenders of originalism in other areas often acknowledge the need to address the wider challenge: Do changes in the U.S. economy, for example, justify (or really demand) a corresponding change in our understanding of the federal power to regulate "inter-state commerce"? It might be that changes in military technology, world power balances or the character of international conflicts now require an executive with more initiative in the deployment of force. It might be that changes in international trade patterns—such as the prevalence today of multilateral bargaining forums rather than state-to-state treaties to "establish commerce"—make it more necessary to adapt old constitutional structures, even if modern adaptations are (as he says of the legislative-executive agreements establishing the WTO and NAFTA) "sort of unconstitutional."

Ramsey's very limited focus not only leaves such questions unaddressed but leaves their implications outside his version of originalism. We get much more of the lawyerly letter and much less of the founding political spirit in this account of constitutional meaning. We learn where the Founders hoped to draw the line, not where or when or how they might have been willing to accommodate new circumstances, given their overall goals as statesmen.

To inquire into permissible adaptation is, of course, to open a door to manipulations. It risks making the Constitution so plastic and adaptable that it is no longer a constitution. But to view original meaning as a matter cut off from questions of overall purpose makes it harder to assess or distinguish different sorts of adaptation. Ramsey's posture of chaste self-denial—claiming to offer arguments that "can guide" us while disclaiming any conclusion "that they should"—implies that the Founders had nothing of value to say on how much the Constitution should or should not be binding on future generations. It is to reduce them from actual founders, with a large vision, to mere lawyers, with a lawyer's kit of line-drawing tools.

But such limitations may be a price Ramsey willingly pays to disentangle his exposition from the distorting fervor of contemporary debates. He does gain something by offering himself simply as a special consultant on historical questions. What he says is unfailingly intelligent, and often provocative, in the way a new and well-developed argument shakes complacent or thoughtless assumptions. Perhaps it is understandable that he does not feel compelled to draw out the moral for national security policy or participation in global governance. He can afford to leave much of the argument to others.

Ramsey's book will be an essential starting point for future debates. Lawyers in the executive branch will need to study and consider its arguments, even where they disagree with some of its conclusions. Critics of executive policy will also find helpful guidance in Ramsey's book. *The Constitution's Text* sets the standard for what, in the first place, should be offered to support a serious conclusion.

Until Proven Innocent: Political Correctness and the Shameful Injustices of the Duke Lacrosse Rape Case

By Stuart Taylor & K.C. Johnson

*Reviewed by Michael Madigan**

Your son, a college student and athlete—a good kid who is a year away from graduating from a well-regarded university—is about to compete for a national championship in a major college sport. You could not be more proud. Until, his and your world come crashing down. He and his teammates are falsely accused of a crime—a terrible crime.

The story gets worse. Your son is not only accused, his case is sensationalized on national television. Your son and his teammates are pursued by a crazed prosecutor desperately seeking re-election, who goes on all the TV shows and says the boys are guilty. This prosecutor is determined to do anything it takes to convict your son and put him in jail—for many years.

But there is more. Your son's University President, and many of the faculty, turn against the boys and denounce them publicly. The University then fires their coach and cancels their athletic season, ending any opportunity to compete for the National Championship.

Before the spring of last year your reaction probably would be that such a thing could not happen in the United States. The old Soviet Union maybe. Other places in the world. But not in this county. But it did. And Stuart Taylor and K.C. Johnson lay out the whole sorry, sad, and outrageous tale in *Until Proven Innocent*—a must-read for every parent with even a passing interest in the criminal justice system in our country. The book takes the reader on a long and careful journey through what really happened, from the March 16, 2006 party through the indictment of three innocent Duke Lacrosse Team players to the disgraced resignation of the prosecutor.

The book meticulously details the Duke Lacrosse case from beginning to end. Its revelations are nothing short of shocking. It reveals how an elected criminal justice official, Michael Nifong, the District Attorney for Durham County, North Carolina, became a "rogue prosecutor" (so-called by the North Carolina Attorney General after his office conducted a special investigation of the prosecutor's conduct) seeking to buttress his difficult upcoming re-election campaign by railroading innocent college students and trumpeting his actions on national television--appearances during which he not only falsely accused them but misstated and in some cases outright made up "evidence" which never existed, not only just about destroying innocent lives but intentionally inflaming race relations in a diverse community.

Nifong failed to even interview the alleged victim for months, withholding from the defense DNA evidence which confirmed the boys' innocence, using a photogenic identification process which has been outlawed for years to try to influence the accuser to identify a Duke lacrosse player. The book notes the fact that the accuser told the police at the time of the incident that the perpetrators were named Adam, Brett, and Matt and that she remembered clearly that Matt said "he was getting married tomorrow." But none of the players were getting married any time soon and the players ultimately indicted were Dave, Collin, and Reade. Nifong also refused to accept the boys' offer to take polygraph exams or look at documented exculpatory evidence such as time-stamped photos proving one of the boys was not even at the party at the time of the alleged incident.

Nifong's conduct was aided and abetted by a police officer with a grudge against the University and a series of media personalities who prejudged the case and repeatedly condemned the boys on national television, often by trumpeting facts which were demonstrably untrue. Duke's President, Richard Brodhead, also proclaimed the boys guilty before the evidence was in, and a large number of Duke faculty did the same. The condemnation reached its pinnacle when "wanted posters" (similar to those seen in post offices for criminals) appeared throughout the Duke campus with photographs and names of the entire Duke lacrosse team. As this virtual lynch mob engulfed Duke, only one faculty member stood firm and tall, supporting the need for the due process of law-- Professor Jim Coleman, who ultimately conducted the independent review that condemned the prosecutor's actions.

When the hidden evidence finally saw the light of day, Nifong was forced to step aside, and the North Carolina Attorney General conducted a thorough investigation of what really happened that night in March at the lacrosse party. He reported:

> With the weight of the state behind him, the Durham district attorney pushed forward unchecked. There were many parts in the case where caution would have served justice better than bravado. And in the rush to condemn, a community and a state lost the ability to see clearly. Regardless of the reasons this case was pushed forward, the result was wrong. Today, we need to

....

** Mr. Madigan is a Senior Partner in the law firm of Akin Gump Strauss Hauer & Feld LLP, a former federal prosecutor and specializes in white collar criminal defense.*

> learn from this and keep it from happening again to anybody... This case shows the enormous consequences of overreaching by a prosecutor. What has been learned here is that the internal checks on a criminal charge—sworn statements, reasonable grounds, proper suspect photo lineups, accurate and fair discovery—all are critically important.

How could this happen in a country which is supposed to have the finest justice system in the world? And, if it can happen to three kids whose families were able to mount the resources to fight back, how about the thousands who cannot? It should come as no shock, considering this case, that more than 200 prisoners have been exonerated since the advent of DNA evidence in the late 1980s, including fourteen innocent death row inmates. Or that the grand jury system, which has eroded greatly over time, failed in this instance. The grand jury was designed by our forefathers to serve as a strong and meaningful check on rogue prosecutors. But today, across our country, it seems, nothing like this system can be found in practice. As the authors describe, the grand jury in the Duke case did exactly what Nifong asked; it indicted the three players. It may shock the average citizen to know that

> the Durham grand jury heard no testimony from (the accuser), (the other dancer who was there at the scene and told the police that no assault happened), any lacrosse player, any doctor or nurse, or anyone else with firsthand knowledge of what had happened. The only witnesses were two cops who had already lied repeatedly to the players and the court about the case

What corrective one draws from this story, if any, is open to debate, but what Taylor and Johnson convey clearly in *Until Proven Innocent* is just how much damage a rogue prosecutor can wreak. Anyone who reads it might well come to the conclusion that, had justice prevailed, it would be Nifong who should have been indicted, convicted, and sentenced to serve substantial time in jail.

The Founders on Citizenship and Immigration: Principles and Challenges for America

By Edward J. Erler, Thomas G. West & John Marini

*Reviewed by Margaret D. Stock**

Given the prominence of immigration issues in American politics today, an up-to-date and scholarly volume on the Founders' views on immigration and citizenship issues could benefit those who seek an understanding of first principles. Unfortunately, *The Founders on Citizenship and Immigration*, a slim volume of four essays, will not answer the pressing need for an authoritative resource. The book is mostly not about the Founders, and provides little in the way of a scholarly addition to the debate. Instead, all but one of its essays simply repeat the restrictionist arguments made in the popular press in the last few years—often lacking citations, and nearly always without consideration of alternative, contradictory sources. Conservatives and libertarians who appreciate the value of understanding "original intent" when evaluating constitutional debate, and who would like to pursue this approach with immigration issues, will be disappointed.

The book's four chapters include an introduction and three essays, each written by one of the co-authors. Edward J. Erler's introduction offers a highly slanted view of the current politics of immigration—one with which many conservatives and libertarians would disagree. For example, in the first section, Erler states that "there is no special interest constituency for restricting immigration," and argues that expansive immigration legislation is repeatedly passed by a Congress oblivious to public opinion so that all immigrants—legal and illegal—can become "malleable clients for the ministrations of the welfare state." In fact, Congress in the past twenty years has found it quite difficult to pass any legislation favorable to immigrants—nearly all significant immigration legislation since 1986 has been "enforcement only" legislation. Congress has been extraordinarily sensitive to the views of angry restrictionist factions, such that a vocal group of anti-immigration "special interests" was largely responsible for the recent defeat of comprehensive immigration reform. In another example contrary to Erler's thesis, immigrants—both legal and illegal—have been mostly barred from obtaining welfare benefits since the 1996 welfare reform laws. Today, it is mostly U.S. citizens who are the clients of "the welfare state," such as it exists—as the State of Colorado found recently when it verified the status of all of its welfare recipients and found only U.S. citizens on the dole.

All three of the authors purport to rely on the Founders' views to support their conclusions, but the authors' policy prescriptions for today do not necessarily follow from the "founding principles" they discern in the Founders' words. In the introductory chapter, for example, Erler discusses Thomas Jefferson's opinions on the character of likely immigrants and what character would be necessary to make good citizens. According to Erler, Jefferson expressed concern that "most of the immigrants to America would be refugees from absolute monarchies" who would not have the habits necessary to make good American citizens because "the habits and manners of freedom are not so easily acquired." If one believes that this principle should determine which foreigners are permitted to immigrate, then presumably the United States should favor those immigrants who come from more democratic societies, and disfavor those from totalitarian regimes. But no one today would argue that Cuba is more democratic than Mexico—why then, do the book's authors apparently favor an immigration policy that gives automatic amnesty to almost all Cubans who arrive in the United States, while making it almost impossible for citizens of the far more democratic Mexico to immigrate legally?

Erler's second essay in the book focuses mostly on the birthright citizenship issue. Birthright citizenship is of intense interest to many conservatives and libertarians—but understanding whether the Fourteenth Amendment demands

** Margaret D. Stock is a Lieutenant Colonel in the Army Reserve, and an Associate Professor in the Department of Social Sciences at the U.S. Military Academy at West Point (NY). Those interested in immigration issues should read Marc Harrold's article on "sanctuary cities" in the Criminal Law and Procedure section.*

that the children of illegal migrants are "natural born" U.S. citizens is not really a matter of the Founders' intent, but rather a matter of the intent of those who ratified the Fourteenth Amendment. Erler wants to argue that these ratifiers would have denied birthright citizenship to the children of illegal immigrants. Here, however, he fails to acknowledge the key legislative sources that contradict his view. The debates in the record show that the ratifiers of the Fourteenth Amendment intended only to exclude from birthright citizenship three groups—the children of diplomats who held diplomatic immunity; Native Americans who were thought to be members of sovereign Indian nations and thus immune from U.S. civil and criminal law; and children born to an invading military force, which would be immune from U.S. civil and criminal laws under the laws of war. The ratifiers knew about the children of illegal immigrants, and certainly could have excluded them—many African Americans who were clearly granted birthright citizenship by the Fourteenth Amendment were themselves the children of African slaves who had been smuggled into the United States in violation of laws barring the importation of slaves. The ratifiers of the Fourteenth Amendment arguably did not have the intent that Erler wishes to attribute to them.

In the book's third essay, co-author Thomas G. West returns to Jefferson's somewhat contradictory views on immigrants, and goes on to conclude Jefferson would not approve of current immigration policy because recent immigrants, and especially "Hispanic immigrants," are not "behaving themselves as well as the rest of America." West does not provide much support for this assertion, and fails to mention recent and well-publicized contradictory studies suggesting that immigrants have lower crime rates than native-born Americans, are less likely to use drugs, do not drop out of military training as often, and are more likely to be employed. He also uses data that aggregates all Hispanics—both citizens and non-citizens, recent arrivals and those in the United States since the Founding—to infer that Hispanic *immigrants* are badly behaved as a race, and therefore should be barred from immigrating. Yet it is hard to be sure without further discussion that Hispanic immigrants as a group share the same characteristics as all Hispanics in the United States, and there are other explanations besides bad character for some of his "shocking" statistics. (Scholars have pointed out, for example, that a higher out-of-wedlock birth rate among Hispanic women is in part due to the fact that Hispanic women are much less likely than white women to seek an abortion for an unwanted pregnancy.) Without further details and discussion of the sources of West's assertions, it is hard to accept that race alone reflects one's statistical propensity to be a bad citizen, and thus that immigration policies should be race-based, as West apparently argues. This slip-up also undermines the book's later argument that it is not other races that restrictionists fear but newcomers of bad character. The essay is also inartfully edited, such that it is often hard for a reader to follow the argument. West jumps around from topic to topic, concluding in the end that "the immigration question is inseparable from the question of the future of the administrative state, the future of modern liberalism." With dire predictions about the death of the American Republic, West demands "an indefinite moratorium on almost all immigration" and increased enforcement and employer sanctions. He has no explanation for how this view reconciles with the Founders' grievance—found in the Declaration of Independence—against King George's failure to pass laws "to encourage [foreigners'] Migrations hither." Like the Founders who signed the Declaration, many of today's conservatives and libertarians see immigration as a source of America's strength, not a liability.

Early in the book, Erler asserts that American policymakers "no longer believe that there are regime principles or that questions of merit and character have anything to do with immigration." While this statement is easily refuted by a passing glance at the numerous grounds of character-based inadmissibility and deportability found in the U.S. immigration code, Erler argues that in contrast to today's policymakers, the Founders believed in a principle of race-blind equality in immigration matters. The last essay, by John Marini, argues that Progressive-era philosophies have embedded race and class as fundamental distinctions in modern American society, such that a return to alleged Founding principles of race-blind equality is impossible, and those who would base immigration policies on equality principles will inevitably be deemed racist by most listeners. Marini makes this argument in an attempt to justify modern restrictionist immigration policies, but his essay will not reassure those today who seek to avoid being labeled "racist" for making restrictive immigration arguments. While Marini tries to provide intellectual support for restrictionists by asserting that their arguments find support in equality principles of the Founding, his argument is so complex and cautious that only a dedicated and patient reader will appreciate its logic. (One must also be willing to ignore the numerous provisions of current immigration law that contradict his underlying thesis.) Marini's detailed explication of Progressive-era restrictionist immigration policies will likely cause most readers to hear the echoes of Progressive-era racist restrictionists in modern restrictionists—and, as Marini implies, they cannot help but do this, because they are the product of the group-based politics of today.

Like Erler, Marini and West err in repeatedly claiming that today's immigrants are highly dependent on welfare benefits, and that immigrants do not seek to become citizens. They are apparently unfamiliar with the 1996 immigration reform laws, which eliminated welfare benefits for most immigrants, required immigrants' sponsors to reimburse the government for benefits given to the immigrants, and made severe distinctions between citizens and non-citizens in many other areas of political and social life. In the past ten years, laws giving preferences to citizens over non-citizens have proliferated. At the same time, increasingly harsh immigration laws—mandatory deportation, for example, for very minor offenses—have made life in the United States increasingly uncertain even for legal immigrants. Immigrants have responded by naturalizing at record rates. Thus, the authors' concerns with the "devaluing" of U.S. citizenship seem misplaced—there may be some evidence that "natural born" citizens are not aware of the value of American citizenship, but immigrants surely are.

One of the characteristics of the current immigration debate has been that many conservatives and libertarians disagree on fundamentals. This book will do little to bridge the

gap. If the authors had simply tried to discuss the Founders' statements on immigration and citizenship, they might have met the objective promised by the title. Instead, the authors nearly always conclude that the Founders' views would support today's restrictionist views. The reality of the historical record is much more ambiguous. For a more balanced and intellectual treatment of the contradictions in the Founders' views, conservatives and libertarians would be better served by reading Rogers Smith's *Civic Ideals: Conflicting Visions of Citizenship in U.S. History*.

In a scholarly book, one would expect both sides of the argument to be presented, and a logical and factually supported argument made why one side is more persuasive than the other. Such is the essence of an informed debate. Throughout this volume, however, the authors quote selectively from only one side, and seem ignorant of current immigration law. They neglect almost entirely the reasoned scholarship on the other side. Ultimately, they fail to add to "an open and honest public debate on immigration." This book will convince only those who are already in accord with the authors' views, or those who fail to appreciate that the lack of citations to the authors' more controversial statements reflects their lack of validity, not their general acceptance.

Practice Group Publications Committees

Administrative Law & Regulation
Christian Vergonis, *Chairman*

Civil Rights
Brian T. Fitzpatrick, *Chairman*

Corporations, Securities & Antitrust
Michael Fransella, *Chairman*
Carrie Darling, James C. Creigh & Robert Witwer, *Vice Chairmen*

Criminal Law & Procedure
Stephen Higgins, *Chairman*
William A. Hall, Jr., Timothy Lynch,
Ronald J. Rychlak & Peter Thomson, *Vice Chairmen*

Environmental Law & Property Rights
Donald Kochan & Andrew Gold, *Co-Chairman*

Federalism & Separation of Powers
Ben Kelly, *Chairman*
Geoffrey William Hymans, *Vice Chairman*

Financial Services & E-Commerce
Charles M. Miller, *Chairman*

Free Speech & Election Law
Stephen M. Hoersting, *Chairman*
Francis J. Menton, Jr., *Vice Chairman*

Intellectual Property
Paul E. Dans & Mark Schultz, *Co-Chairmen*

International & National Security Law
Ata Dinlenc, *Chairman*
Jennifer A. DePalma & Eric J. Kadel, Jr., *Vice Chairmen*

Labor & Employment Law
John F. Scalia, *Chairman*
Raymond J. LaJeunesse, Jr., Daren Bakst,
& Matthew Zandi, *Vice Chairmen*

Litigation
Margaret A. Little & Christopher C. Wang,
Co-Chairmen

Professional Responsibility & Legal Education
John J. Park, Jr., *Chairman*

Religious Liberties
Becky Dummermuth, *Chairman*

Telecommunications & Electronic Media
Raymond L. Gifford, *Chairman*

The Federalist Society
for Law & Public Policy Studies
1015 18th Street, N.W., Suite 425
Washington, D.C. 20036

NONPROFIT ORG.
U.S. POSTAGE
PAID
WASHINGTON, DC
PERMIT NO. 5637