HARVARD JOURNAL
of
LAW & PUBLIC POLICY

VOLUME 29, NUMBER 1 FALL 2005

LAW AND FREEDOM

THE TWENTY-FOURTH ANNUAL NATIONAL STUDENT
FEDERALIST SOCIETY SYMPOSIUM ON
LAW AND PUBLIC POLICY—2005

FEDERALIST SOCIETY ASSISTANT EDITORS.....1

THE NATURE AND IMPORTANCE OF LIBERTY
Charles Fried.....3

FREEDOM
Michael S. Moore.....9

NOT NECESSARILY IN CONFLICT: AMERICANS CAN BE BOTH UNITED AND CULTURALLY DIVERSE
Jennifer C. Braceras.....27

THE HUMAN NATURE OF FREEDOM AND IDENTITY—WE HOLD MORE THAN RANDOM THOUGHTS
Douglas W. Kmiec.....33

FIVE THESES ON IDENTITY POLITICS
Richard D. Parker.....53

NEWS FOR THE LIBERTARIANS: THE MORAL TRADITION ALREADY CONTAINS THE LIBERTARIAN PREMISES
Hadley Arkes.....61

SAFETY *AND* FREEDOM: COMMON CONCERNS FOR CONSERVATIVES, LIBERTARIANS, AND CIVIL LIBERTARIANS
Nadine Strossen……73

THE VIRTUES OF PREEMPTIVE DETERRENCE
David B. Rivkin, Jr……85

ARTICLES

TOWARD THE FRAMERS' UNDERSTANDING OF "ADVICE AND CONSENT": A HISTORICAL AND TEXTUAL INQUIRY
Adam J. White……..103

ENDING THE WAR ON TERRORISM ONE TERRORIST AT A TIME: A NONCRIMINAL DETENTION MODEL FOR HOLDING AND RELEASING GUANTANAMO BAY DETAINEES
Tung Yin……..149

USING INFORMATION MARKETS TO IMPROVE PUBLIC DECISION MAKING
Robert W. Hahn & Paul C. Tetlock…….213

AGAINST FOREIGN LAW
Robert J. Delahunty & John Yoo……..291

RECENT DEVELOPMENTS

THE SUPREME COURT OF THE UNITED STATES, 2004 TERM

IMPOSING NECESSARY BOUNDARIES ON JUDICIAL DISCRETION IN BALLOT ACCESS CASES: *Clingman v. Beaver*, 125 S. Ct. 2029 (2005)……331

PLACING UNNECESSARY LIMITS ON VOTING AND ASSOCIATIONAL FREEDOMS: *Clingman v. Beaver*, 125 S. Ct. 2029 (2005)……..345

PAYING FOR THE SINS OF THEIR USERS: LIABILITY AND GROWING UNCERTAINTY IN A DIGITAL AGE: *Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*, 125 S. Ct. 2764 (2005)..357

CHASING FINALITY: FEDERAL COLLATERAL RELIEF IN THE WAKE OF *Pace v. DiGuglielmo*, 125 S. Ct. 1807 (2005)……..373

HARVARD JOURNAL
of
LAW & PUBLIC POLICY

Editor-in-Chief
JENNIFER CARTER

Articles Editor
PATRICK J. BUMATAY

Deputy Editor-in-Chief
JOHN C. COOK

Managing Editors
JOSHUA D. DAVEY
CASSANDRA H. WELCH

Executive Editors
DAN ALBAN
DARIN T. BEFFA
MATT COOPER
TRACY LEIGH DODDS
JENNIFER FREEBERG
JASON M. HULL
JEFFREY THEODORE
BRANDIE N. WEDDLE
EMILY D. ZIMMER

Assistant Articles Editors
ERIN SHELEY
MICHAEL KORNS
SUSAN LOUISE AMBLE
PAUL WATKINS

Business Manager
JEFFREY M. HARRIS

Subscriptions Manager
VONN R. CHRISTENSON

Notes Editor
LOWELL SCHILLER

Web Manager
MARY BEAR

Senior Editors

BRYSON L. BACHMAN
AMY BLAKEWAY*
DANIEL O. BURKHART
MARK CHAMPOUX
JAMES COLEMAN
JAY DEALY
ELAINE EWING
N. NOELLE FRANCIS
TORY JACKSON
CHRISTIAN JENNER
MATTHEW KOHLEY*
GREG KRAFKA
WILL LEDBETTER
JOHN LOBATO
MATT LONG
YOON AH E. OH
KIMBERLY OLVEY
JOSEPH Z. PERKINS*
NICHOLAS PILCHAK
A. PAUL PINEAU
SETH PRITIKIN
MIKE REDMOND
KENNETH REINKER
JUNE RHEE
JOEL SCHELLHAMMER*
M. JASON SCOGGINS
MICHAEL M. SEVI
CORAL A. SHAW
BRIAN SHULL
DANIEL J. SULLIVAN
BRETT J. TALLEY
KAREN TEPICHIN*
JERRY THOMAS
JENNIFER VOSKO
JAMES A WAWRZYNIAK, JR.
CASSANDRA H. WELCH

* *Articles Board*

Editors

LINDSAY M. ADDISON
JOSHUA D. ANDERS
SALVATORE ASTORINA
ADAM C. BAGLEY
BENJAMIN BARRON
KEN BASIN
KODIE BENNION
ANDREW M. BERNIE
MICHAEL BOLDT
BRENDAN BOWES
AARON BROOKS
ANDREW S BROWN
ROBERT M. BROWN
BRIAN R. CALLANAN
PAUL CASADABAN
RYAN CAUGHEY
KELLIN CHATFIELD
GRACE CHIEN
CANDICE CHIU
JOCELYN ANNE CHUNG
MATTHEW M. CRAIG
ELIZA CRANE
SCOTT P. DIXLER
JONATHAN DUNN
ZINA L. GELMAN
JEN GHAUSSY
DAN GOMEZ
BENJAMIN GOULD
KATHLEEN J. GRAY
MEGAN GREER
KEN GROSE
JARED M. GROSS
MICHAEL HAWRYLCHAK
KEVIN HINKLEY
SARAH M. ISGUR
DAVID KARGMAN
JOSH L. KREAMER
MATTHEW J. LATTERNER
SHIH JERN LIANG
SCOTT H. LINTON
DEREK LYONS
HAIM MACHLUF
STUART MATHESON
JASON MEHTA
MATTHEW MERKLE
PAUL ALESSIO MEZZINA
ANDREW MOIN
JONATHAN PAPIK
AMANDA PHILLIPS
JOHN PIAZZA
DANIEL PIERCE
ROBERT PORTER
ANTON PRIBYSH
ADAM C. G. RINGGUTH
JAMES ROGERS
JEFFREY C. SINDELAR, JR.
LANDE SPOTTSWOOD
ANDREW STEINMAN
AMY TAY
KATIE THOMASON
DANIEL M. UNGAR
NICOLE VALCO
YANJUN WANG
GEOFFREY C. WEIEN
DAVID MCKAY WEILER
SHEENA E. WEISNER

Founded by E. Spencer Abraham & Steven J. Eberhard

BOARD OF ADVISORS

E. Spencer Abraham, Founder

Robert H. Bork
Steven G. Calabresi
Douglas R. Cox
Charles Fried
Douglas H. Ginsburg
Orrin Hatch
Jonathan R. Macey
Michael W. McConnell
Diarmuid F. O'Scannlain
Jeremy A. Rabkin
Hal S. Scott
David B. Sentelle
Bradley Smith
Jerry E. Smith

THE *HARVARD JOURNAL OF LAW & PUBLIC POLICY* RECEIVES NO FINANCIAL SUPPORT FROM HARVARD LAW SCHOOL OR HARVARD UNIVERSITY. IT IS FUNDED EXCLUSIVELY BY SUBSCRIPTION REVENUES AND PRIVATE CHARITABLE CONTRIBUTIONS.

The *Harvard Journal of Law & Public Policy* is published three times annually by the Harvard Society for Law & Public Policy, Inc., Harvard Law School, Cambridge, Massachusetts 02138. ISSN 0193-4872. Nonprofit postage prepaid at Lincoln, Nebraska and at additional mailing offices. POSTMASTER: Send address changes to the Harvard Journal of Law & Public Policy, Harvard Law School, Cambridge, Massachusetts 02138.

Yearly subscription rates: United States, $45.00; foreign, $50.00. Special introductory rates are available for students. Subscriptions are renewed automatically unless a request for discontinuance is received.

Back issues in complete sets, volumes, or single issues may be obtained from: William S. Hein & Co., Inc., 1285 Main Street, Buffalo, New York 14209.

The *Journal* welcomes the submission of articles and book reviews. Each manuscript should be typed double-spaced, preferably in Courier 10-point typeface, and submitted in both hard copy and electronically to jlpp@law.harvard.edu, preferably prepared using the Microsoft Word word processing program or saved as "rich text format." Views expressed are those of the authors and do not necessarily reflect the views of the Society or of its officers, directors, editors, members, or staff. Unless otherwise indicated, all editors are students at the Harvard Law School.

Copyright © 2005 by the Harvard Society for Law & Public Policy, Inc.

PREFACE

Volume 29 opens in the midst of a tumultuous and potentially pivotal chapter in American legal history. As the nation anticipates a new epoch for the Supreme Court, the Senate continues to grapple with its relationship to the President and the judiciary. Judges, scholars, and lawyers struggle daily with the myriad legal questions generated in the war on terrorism, by emerging technologies, and over the content and scope of individual rights; the public has energetically taken its elected representatives to task over these legal issues. We hope that these pages offer constructive contributions in these areas, in keeping with this *Journal*'s tradition of advancing outstanding conservative and libertarian scholarship.

This Issue begins with eight papers from the Twenty-Fourth Federalist Society Student Symposium, *Law and Freedom*. We are pleased to present these Symposium highlights, which include competing views on the philosophical underpinnings of freedom, the intersection of freedom and identity, and the trade-off between freedom and security in the war on terrorism.

Next, in view of the hotly contested issue of the proper role of the Senate in confirming the President's judicial nominees, Mr. Adam White has investigated the Framers' understanding of the Advice and Consent Clause. His Article features original historical research into the practice of advice and consent in early Massachusetts, upon which the constitutional clause was based, and concludes that in the original understanding, the Senate does not have an affirmative constitutional duty to vote up or down on a nominee.

Professor Tung Yin's latest contribution to post-September 11 scholarship presents a framework for the detention and release of enemy combatants held at Guantanamo Bay. After discussing traditional military and law-enforcement mechanisms, he concludes that neither set of norms provides a workable fit for the novel challenges of the war on terrorism. His proposal for a noncriminal process draws on analogies to pretrial detention, civil commitment, and medical quarantine.

Dr. Robert Hahn and Professor Paul Tetlock present a proposal for the design and use of information markets in public decision making as an advancement over existing cost-benefit analysis tools. Their Article highlights the power of markets to quantify the value of a policy outcome through the price of a futures contract on a public market, and promises to be a major contribution to the emerging body of literature on this subject.

Professors Robert Delahunty and John Yoo confront the Supreme Court's increasing reliance on foreign law in U.S. constitutional adjudication. They contend that the outcome-

determinative use of foreign law is incompatible with fundamental American institutions such as the separation of powers and judicial review, and that the European norms on which the Justices draw are squarely at odds with America's history and values.

Finally, four student-authored Notes discuss 2004 Term decisions of the U.S. Supreme Court. We present two perspectives on *Clingman v. Beaver* (upholding Oklahoma's semi-closed primary elections) and commentary on *MGM Studios Inc. v. Grokster, Ltd.* (holding software developers liable for their users' copyright infringement) and *Pace v. DiGuglielmo* (ruling that a postconviction petition untimely filed under state law does not toll the federal statute of limitations for habeas corpus petitions).

This Issue of the *Journal* represents the efforts of these authors in collaboration with more than one hundred devoted staffers. It also could not have come about without the Federalist Society and the hard work of Symposium organizers Tessa Platt and Daniel Kelly and of the nationwide team of Symposium Editors. I am indebted to Joshua Davey, who pulled double duty as National Editor of the Symposium and Managing Editor of the *Journal*. Cassandra Welch has also been invaluable as Managing Editor; Vonn Christenson has single-handedly managed the largest circulation of any student-edited law journal in the country. Deputy Editor-in-Chief John Cook has gone above and beyond the call of duty at every opportunity.

I am particularly grateful to Amber Taylor and the staff of Volume 28, who passed along the *Journal* in such excellent condition, and to the courage and dedication of all those who have sacrificed their time and energy to the *Journal* since 1977. I hope that this Volume will live up to their high standards.

Jennifer Carter
Editor-in-Chief

THE FEDERALIST SOCIETY

Presents

The Twenty-Fourth Annual Symposium on Law and Public Policy:

Law and Freedom

February 25 and 26, 2005
Harvard Law School

The staff acknowledges the assistance of the following members of the Federalist Society in preparing this Symposium for publication:

National Editor
Joshua D. Davey
Harvard Law School

Executive Editors

Jason A. Adkins
University of Minnesota Law School

Aaron Nielson
Harvard Law School

Gregory S. McNeal
Case Western Reserve University School of Law

Raj Parekh
University of Pennsylvania Law School

General Editors

Rian J. Balfour
Duke University School of Law

Anthony Daniel Bianco
Notre Dame Law School

William H. Burgess
University of Pennsylvania Law School

Heather Ann Cameron
Notre Dame Law School

Mark Chorazak
Duke University School of Law

Seana Cranston
Catholic University of America, Columbus School of Law

Zachary Jacobs
Columbia Law School

Matthew J. Kleiman
Duke University School of Law

Irina Manta
Yale Law School

Daniel P. McGuire
Ave Maria School of Law

Brian P. Morrissey
Notre Dame Law School

Matthew Muggeridge
Ave Maria School of Law

Eugene Nardelli
Yale Law School

David C. Presser
Northwestern University School of Law

Susannah E. Prucka
University of Maryland School of Law

Kristopher Ritter
Notre Dame Law School

Amir Hassan Sadaghiani
Notre Dame Law School

Daniel J. Sullivan
Harvard Law School

Jessica Erin Tannenbaum
Notre Dame Law School

Peter M. Tempelis
University of Wisconsin Law School

Amy Y. Yeung
Duke University School of Law

THE NATURE AND IMPORTANCE OF LIBERTY

CHARLES FRIED*

What is liberty, and why is it important? Why do we care about it? The first premise that I offer here is that liberty is an expression of what is valuable about us as human beings. It is a natural law idea; that is to say, it is a moral imperative based on what is fundamental (another moral idea) about our human nature.[1]

I would say that what is important about us, what makes us moral human beings, is our individual capacities to think, reason, choose, and value. It is what Kant called our freedom and rationality.[2] Individuals, therefore, are the elementary particles of moral discourse. Our value is our taking individual responsibility for our lives, and our choices. And if a person is to count as a person—and here we have the difficult questions about the beginning and the end of life—then we are all equally valuable in this same way. It is from that base of our equal responsibility for ourselves that we choose our goods: that we choose what to make of the only life we will ever have.

My liberty, then, is my ability to choose that life. No one has the right to interfere with that choice, except as it is to further his own good. But that good of the other is worth no more than mine because he is not worth any more than I am. There is, therefore, a right of mutual noninterference: an equal right. By the same token, nobody can interfere with or draft another person to help him achieve his own good if the other person has not chosen voluntarily to enlist in that campaign.

---

* Beneficial Professor of Law, Harvard University; Associate Justice, Supreme Judicial Court of Massachusetts, 1995–1999; Solicitor General of the United States, 1985–1989. This piece comprises Professor Fried's notes for a chapter in a book on liberty that will be published in late 2006 by W.W. Norton.

1. *See generally* JOHN RAWLS, A THEORY OF JUSTICE (1971).

2. *See* IMMANUEL KANT, GROUNDWORK OF THE METAPHYSICS OF MORALS 116 (H.J. Paton ed. & trans., Harper Torchbooks 3d ed., 1956) (1948) (declaring that "to every rational being possessed of a will we must also lend the Idea of freedom as the only one under which he can act").

There are conclusions to be drawn from these postulates; for example, that liberty is a relation among persons. Liberty is violated when someone else interferes with it. Gravity, tigers, and disease do not interfere with my liberty; other people do. In addition, I can make your good, that which you choose, part of my good. But then I must work through you, through your liberty, and not upon you, not in spite of your liberty.

And what is interference? It can be hindering you for your sake, for your good. Or hindering you for my sake, for my good. When do I hinder you in that second sense; when do I hinder you for my sake? There are two ways. First, when I disregard you. Second, when I use you, as when I trick you or threaten you or force you in some way or another.

The issue of hindering others by disregard is, in fact, conceptually the most difficult. It is pretty easy to tell when I hinder you in the second sense: when I use you, when I trick you, when I force you. These are clear-cut violations of liberty. But, how can we tell when I am hindering you by, for instance, just failing to help you, by just disregarding you? The picture I have is of my driving down the middle of the highway at eighty miles per hour, disregarding that you are coming the other way. Am I hindering your liberty, or does your demand that I regard you hinder mine? It is out of this dilemma—this dilemma of when do we hinder each other by disregarding each other, just going our way, running over each other—that we derive, in a very general sense, the concept of property. Which is why property is so closely related to liberty.[3]

Property describes the whole—I will call it the *n*-dimensional space, as it is much more than three-dimensional—*n*-dimensional space in which I may operate as a free person without being hindered by you as you go about in pursuit of your goods. You may not violate my property, you may not enter my *n*-dimensional space. But that *n*-dimensional space must be defined somehow. Is it defined by liberty itself, or is it conventional, or is it a little bit of each?

If it is conventional, if our property in ourselves and our property in the outside world which we have assimilated to ourselves is wholly conventional, then of course its boundaries are established by convention. If it is the creature of conven-

3. *See* JOHN LOCKE, SECOND TREATISE OF GOVERNMENT 32 (C.B. Macpherson ed., Hackett Publ'g Co. 1980) (1690) (describing liberty as the ability to dispose and order one's whole property).

tion, it is the creature of others. And that is a serious problem for liberty and one which the friends of liberty have sought many ways to solve—the social contract tradition is such an attempt—and the enemies of liberty have sought to exploit from the very beginning.[4]

Now, there are, however, certain natural, not conventional features of my property in myself and my property in the world. Most importantly, because it derives from my liberty to reason, to choose, to decide, is my liberty of the mind.[5] That is why it is not surprising and it is not perverse that the realm of liberty of contract which, in the Nineteenth Century was celebrated, has gradually shrunk, and why the First Amendment rights have become as robust as they have.[6] It is also why, when thinkers—and there are many of them, for instance Cass Sunstein—promise us a New Deal for the First Amendment,[7] I shiver because I know what the New Deal did to our other property liberties. And I shiver to know what that New Deal holds in store for liberty of mind.

So that is one natural property right that we have that is not conventional. But I think there is another one, and it has a lot to do with why we are holding this Symposium, why we are lawyers, why this is a subject for law. When students graduate from this law school, we proclaim that they "are ready to aid in the shaping and application of those wise restraints that make

---

4. *See* Ronald Dworkin, *The Original Position*, *in* READING RAWLS 16 (Norman Daniels ed., 1975) (criticizing Rawls's conception of social contract); *see also* Anita L. Allen, *Taking Liberties: Privacy, Private Choice, and Social Contract Theory*, 56 U. CIN. L. REV. 461, 462 n.3 (1987) ("The social contract theories advanced by contemporary thinkers who continue in the contractarian tradition of Locke, Rousseau, and Kant depict 'an idealized moral conception of persons as free, rational, and equal' . . . .") (quoting DAVID A.J. RICHARDS, TOLERATION AND THE CONSTITUTION 41 (1986)).

5. *See* CHARLES FRIED, SAYING WHAT THE LAW IS: THE CONSTITUTION IN THE SUPREME COURT 170–206 (2004); *see also* Charles Fried, *The New First Amendment Jurisprudence: A Threat to Liberty*, 59 U. CHI. L. REV. 225, 233 (1992) (stating that the free speech doctrine is founded upon personal autonomy and freedom of the mind).

6. *See* Boy Scouts of Am. v. Dale, 530 U.S. 640, 644 (2000) (holding that a state's application of its public accommodations law to require the Boy Scouts to admit a homosexual member violated the organization's First Amendment right of expressive association); Cohen v. California, 403 U.S. 15, 26 (1971) (concluding that the state does not have the power to criminalize the public display of a four-letter expletive because government does not have the capacity to forbid certain words without also running a substantial risk of suppressing ideas in the process).

7. *See, e.g.*, Cass R. Sunstein, *A New Deal for Speech*, 17 HASTINGS COMM. & ENT. L.J. 137 (1994).

us free." Of course, law is that engine, that system, which defines this *n*-dimensional liberty. But if it is to do that, if it is to define our liberty, it must have certain characteristics. And its characteristics are that it must be stable and it must be formal. That is to say, it must seek to attain a certain neutrality between the contending parties who would seek to realize their various goods.

That is why a libertarian, or liberty-driven, view of law asks at least two things of a system of law. It asks that it be reasonably stable so that we can rely on it. We realize it is difficult to derive from some general principles what all of the law's details are, although the liberty of the mind can so be derived. But we insist that, whatever society gives us, it will at least stay still long enough that we can order our choices, make our plans, and live: The law must not be a moving target, so to speak.

And the second thing is why it is that the friends of liberty are also decried in the world of law as formalists.[8] I always tell my students, "To call an argument a formalist argument is not a term of abuse." And the reason for this is that, when the law is formal, when those who adjudicate do so not purposefully but formally, indeed formalistically, there is an attempt to honor the law as the definition of our property, our liberty space, and not to change that definition, not to change those boundaries for the competing goods of the persons who are contending within those spaces.

Finally, what then of love, of friendship, of generosity, of music, of art, and of science? Are these not the things that are valuable about us? Yes, they are, if they are freely chosen. If they are ours, they arise out of our deliberate personal, individual choices. And if they are not, their value disappears.

My dear friend and frequent interlocutor, Michael Sandel, is the only one on this Symposium panel who really stated the

8. *See* Thomas C. Grey, *Holmes and Legal Pragmatism*, 41 STAN. L. REV. 787, 799 (1989) (describing legal pragmatists' attack on formalism based on first principles); *see also* Duncan Kennedy, *The Stages of the Decline of the Public/Private Distinction*, 130 U. PA. L. REV. 1349, 1351–52 (1982) (describing the collapse of the formalists' distinction between the public and the private spheres); Cass R. Sunstein, *Naked Preferences and the Constitution*, 84 COLUM. L. REV. 1689, 1697, 1718 (1984) (positing that the *Lochner* decision's theoretical basis is undermined by acknowledging that the "market status quo" is truly just the result of previous government choices).

competing point of view, as best articulated by Aristotle.[9] This alternative perspective asserts that freedom sometimes requires that government and law not be neutral with respect to questions of the good life, with respect to moral and religious questions. Freedom is not a matter of autonomy or choosing whatever we happen to want. Rather, to be free is to live a certain mode of life: the good life. Professor Sandel admitted that this conception of freedom and rights could be termed judgmental or moralistic, particularly because, under this rubric, deciding what rights people possess requires one to consider the purpose and moral worth of the social practices that give rise to such questions about rights.

As an example of this understanding rightly put into practice in his estimation, Professor Sandel recounted the majority opinion in *PGA Tour, Inc. v. Martin*.[10] Writing for the majority, Justice Stevens concluded that the purpose and essence of golf has nothing to do with whether or not a golf cart is used, and therefore the Americans with Disabilities Act (ADA) requires the PGA Tour to allow carts to be used by disabled contestants.[11]

It is quite interesting, because there is another case in which Justice Stevens tried to use similar arguments, but did not succeed. In *Boy Scouts of America v. Dale*, he knew what the essence of being a Boy Scout is, and it did not include some distaste for homosexual relations.[12] That was no part of being a Boy Scout, and he knew that.

What was at stake in *Dale* was the ability of the Boy Scouts to define for themselves what it is that their particular association meant: not for Aristotle, and not for Professor Sandel, and not for Justice Stevens. The reason that *Martin* came out the way it did is that, to fall under the sway of the ADA, golf or a golf tournament has to be either commercial or a public accommodation, similar to a hotel or a movie theater.[13] If golf, however,

---

9. *See* Michael J. Sandel, Anne T. and Robert M. Bass Professor of Gov't, Harvard Univ., Remarks on Panel I: What is Liberty? Competing Philosophical and Jurisprudential Perspectives on Liberty at the 24th Annual Federalist Society Student Symposium on Law and Public Policy: Law and Freedom (Feb. 25, 2005).

10. 532 U.S. 661 (2001).

11. *See id.* at 683–86.

12. *See* 530 U.S. 640, 666–70 (2000) (Stevens, J., dissenting) (noting that the Boy Scouts' Law and Oath expresses no position on sexual matters).

13. *See Martin*, 532 U.S. at 675–77 (noting the ADA's public accommodation requirement and finding that the PGA's golf tours on golf courses meet that requirement).

had been a religious ritual then I suppose that Justice Stevens would try to the do the same number, but I devoutly hope he would, once again, have been in the dissent.

FREEDOM

MICHAEL S. MOORE*

I. INTRODUCTION

More than two centuries after our Revolution, we Americans are still arguing over the nature, and even the coherence, of three of the natural rights in whose name the Revolution was fought. We contest (because it is also truly contestable in serious ethical discussions) whether the rights to equality[1] and to property[2] are either conceptually coherent or normatively desirable. The same conceptual and ethical queries exist for liberty, conceived of either as a right or as a value: Some contest its conceptual coherence, others pooh-pooh its plausibility as a basic value or as a general right.[3] Such skepticisms about liberty raise several large questions: What is liberty? Why is it valuable? Can there be a general right to liberty?

This short Article can hardly do justice to these three large questions. Nonetheless, in short compass I intend to sketch an answer to each of them.

II. WHAT IS LIBERTY?

"What is . . . ?" questions are notoriously fuzzy in what they really seek. When attempting to answer such questions, it is wise for one first to discover what it is that is actually puzzling

---

* Charles R. Walgreen Jr. University Chair, Professor of Law, Professor of Philosophy, Professor in the Center for Advanced Study, and Co-Director of the Program in Law and Philosophy, University of Illinois.

1. *See, e.g.*, JOSEPH RAZ, THE MORALITY OF FREEDOM (1986); Peter Westen, *The Empty Idea of Equality*, 95 HARV. L. REV. 537 (1982).

2. *See, e.g.*, Thomas C. Grey, *The Disintegration of Property*, *in* PROPERTY: NOMOS XXII 69 (J. Roland Pennock & John W. Chapman eds., 1980).

3. *See, e.g.*, RONALD DWORKIN, TAKING RIGHTS SERIOUSLY 268–69 (1977); RAZ, *supra* note 1, at 13; Samuel Freeman, *Criminal Liability and the Duty to Aid the Distressed*, 142 U. PA. L. REV. 1455, 1486 (1994).

one.[4] With liberty, our puzzlement is largely normative: What could plausibly be of value and also plausibly be referred to as "liberty" or "freedom"? This suggests that the conceptual question—what do we mean by "liberty"—and the normative question—why is liberty valuable—are like a pair of trousers: Just as we can suspend temporarily putting one leg in while inserting the other, so can we suspend temporarily asking one question while we ask the other. Ultimately we must have one answer that simultaneously satisfies both. What we seek, in other words, is not some concept of liberty that is simply extracted from linguistic or other social practice. Rather, we seek a concept of liberty that at least plausibly names something of interest to us in ethics, something of value such as a good or a right. When we are asking "What is liberty?" we are thus not engaged in conceptual analyses of the old Oxford ordinary language style. We are engaged in a simultaneously normative and descriptive task.[5]

Still, as with trousers, it helps to proceed one leg at a time. What roughly are people typically picking out when they refer to liberty? We have many uses of the word that could serve as our exemplars. Our current President, for example has, in his second term, given numerous "freedom" speeches. Yet such political rhetoric is too overblown for my purposes. Consider a much more mundane example reported recently by the *Wall Street Journal*.[6] In an article entitled "Risky Riders: Touting Freedom, Bikers Take Aim at Helmet Laws," we hear that ABATE (A Brotherhood Against Totalitarian Enactments) and other motorcyclist lobbyist groups have successfully lobbied twenty-nine states to repeal or modify their helmet requirements for adult motorcycle riders. These groups call their program "five steps to freedom"; they call states where helmets are not required the "free states"; their most successful (and perhaps colorful) lobbyist, Sputnik (his legal name), has the word "FREE" tattooed across his forehead.

It is pretty plain what these motorcyclists intend to refer to by "free" and "freedom": it is the absence of criminal laws re-

4. Herbert Hart nicely illustrates this point in his treatment of the "What is law?" question. *See* H.L.A. HART, THE CONCEPT OF LAW 1–17 (2d ed. 1994). I employ a similar strategy in answering the "What is truth?" question in Michael S. Moore, *The Plain Truth About Legal Truth*, 26 HARV. J.L. & PUB. POL'Y 23 (2003).

5. *See* RAZ, *supra* note 1, at 14–16.

6. Karen Lundegaard, *Risky Riders: Touting Freedom, Bikers Take Aim at Helmet Laws*, WALL ST. J., Dec. 1, 2004, at A1.

quiring them to wear motorcycle helmets. This is the kind of usage of these words that should attract our normative interest. Notice that implicit in this usage are three restrictions: First, that the interest is in negative liberty as absence of constraint; the interest is not immediately, at least, in positive liberty, which is the ability to achieve or obtain something.[7] Second, that the interest is in political liberty, so that the restraint, the absence of which is wanted, is legal coercion. Third, that the specific form of legal coercion to be avoided is criminal sanctions; the liberty desired is part of the principled limits of the criminal law, leaving aside whether other forms of legal coercion or inducements might be acceptable.

These three restrictions define a normatively plausible concept of liberty, a concept that allows us to tell the government that some things are simply not its business. It is also a familiar concept of liberty, close to what John Stuart Mill[8] and his critics[9] and followers[10] addressed in Mill's famous essay *On Liberty*. Liberty in this conception is the absence of the coercion of the criminal law.

III. WHY IS LIBERTY VALUABLE?

As we have seen, such a conception must ultimately be justified normatively, which I now propose to do. Why should freedom from criminal sanctions be thought of as a good? Surely the critics of libertarianism have a point when they proclaim that this conception of negative liberty cannot plausibly be thought of as an intrinsic good.[11] Positive liberty—the ability to obtain or do certain things—more plausibly looks to be good in itself, that is, "intrinsically good." The negative liberty conceptualized here is surely only instrumentally good, if it is

---

7. For a famous articulation of this distinction, see ISAIAH BERLIN, TWO CONCEPTS OF LIBERTY (Clarendon Press 1958), *reprinted in* ISAIAH BERLIN, FOUR ESSAYS ON LIBERTY 118 (1969).

8. *See generally* JOHN STUART MILL, ON LIBERTY (David Bromwich & George Kateb eds., Yale Univ. Press 2003) (1859). Mill did not restrict his concept of liberty by the second and third restrictions above, but his famous harm principle is most plausible if those restrictions are honored.

9. *See generally* JAMES FITZJAMES STEPHEN, LIBERTY, EQUALITY, FRATERNITY (Stuart D. Warner ed., Liberty Fund 1993) (1873).

10. *See generally* JOHN C. REES, JOHN STUART MILL'S *ON LIBERTY* (1985).

11. Michael Sandel is a current example. For a discussion of Sandel's earlier work, see MICHAEL MOORE, PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW 647–58 (1997).

good at all. It surely, in other words, is good only because, and only insofar as, it contributes to something else that is good.

Here are six suggestions as to the basic goods to which negative liberty may be a means. First, there is positive liberty itself. Negative liberty is one of the elements necessary for positive liberty to exist. One has the ability (positive liberty) to do some action *A* only if the state does not prohibit *A* by criminal sanction. Freedom from governmental coercion is of course only one of the items needed for one to do action *A*. One also needs physical and economic capacities, for example, as critics of negative liberty are so fond of pointing out. But negative liberty is needed too, so if positive liberty is of value then so is one of its necessary means, negative liberty.

Positive liberty is itself something good. One might plausibly view it as an intrinsic good.[12] This is the view that enhancement of opportunity sets is always good, full stop. Alternatively, one might think that positive liberty is the kind of universal, instrumental good that John Rawls called a primary good: something so useful to the attainment of a wide variety of ends that, no matter what those ends might be, you would want such a good.[13] Either way, positive liberty is valuable and thus so is one of its indispensable means, negative liberty.

Second, there is what might be called Millian autonomy, so called to distinguish it from Kantian autonomy, which is discussed shortly. Millian autonomy is the hard-to-pin-down idea that human choice gains value when freely made. Even bad choices have value in this sense, so long as they are free choices. The usual metaphor is that of authorship: The autonomous person is a being who is author of his or her actions, an agent in some full sense.[14] Mill contrasts this with a mere machine such as a steam engine: Although it can do many things, it is not the author of any of them, so its actions possess no value as autonomously chosen.[15]

If Millian autonomy is valuable, it is easy to see how legal coercion prevents its attainment. A coerced choice is not an

---

12. An intrinsic good is something that is good not because it contributes to something else that is good, but because it is good in itself. Every theory of value, by its commitment to instrumental values, is also committed to there being intrinsic values, on pain of infinite regress. *See id.* at 157–58, 160–62.

13. *See* JOHN RAWLS, A THEORY OF JUSTICE 62 (1971).

14. *See, e.g.*, Charles Fried, *The Nature and Importance of Liberty*, 29 HARV. J.L. & PUB. POL'Y 3 (2005).

15. *See* MILL, *supra* note 8, at 125.

autonomous choice expressive of the agent's authorship. The threat of the law can rob our actions of any of those decisional processes one rightly values as autonomous or free choice.

The third value, Kantian autonomy, is to be distinguished from Millian autonomy. For Kant, an autonomous action is a right act done for a right reason.[16] Kant's thought is that there is value not just in doing the right act, but also, indeed mostly, in doing it for the right reason. The emphasis is on the motives for action, not the processes of choice. Although choice may well be involved whenever an act is motivated by reasons, valuing correct reasons is quite different from valuing free-choice processes.

Note that in the Kantian sense of autonomy, there is no value in an "autonomously" chosen wrong action, nor is there value in an "autonomously" chosen right action when that action is done for the wrong reason. No matter how free one's decisional process may have been, Kantian autonomy is not achieved in such cases because the right reason did not motivate the right action.[17]

Again, if there is value in acting for the right reasons, it is easy to see how legal coercion can prevent the attainment of that value, for the avoidance of legal sanctions can easily supplant the more virtuous motivations for an action that might otherwise have moved the agent to act. Consider, for example, gifts to the poor. If Kant is right about the locus of value residing in reasons and not just in actions, charitable giving is valuable not just because of the act of transferring wealth to those in need, but also because of the benevolence motivating that act. Legally coerce such acts, and the fear of sanctions will often supplant that benevolent motivation. In such cases the virtue of benevolent giving has become the necessity of paying one's taxes.

The fourth, fifth, and sixth considerations showing negative liberty to be instrumentally valuable are all broadly utilitarian in character. To begin with, a preference utilitarian might well regard negative liberty as a kind of presupposition of utilitarianism. Preference-utilitarianism is a monistic theory of the good in which there is only one intrinsically good thing, preference-satisfaction, and the right action or institution is one

16. *See* BARBARA HERMAN, THE PRACTICE OF MORAL JUDGMENT 1–22 (1993).

17. Joseph Raz, who uses autonomy in its Kantian sense, draws both of these inferences. *See* RAZ, *supra* note 1, at 380–81, 411–12.

that maximizes the production of this one good. One might well think that part of what makes human preferences so valuable, and thus worth summing in the utilitarian calculus, is that they are arrived at free of governmental coercion. This is not just avoidance of the circularity problem, which is to wonder how governmentally coerced preferences can guide government policy when they are one of the products of that policy. More importantly here, preference-utilitarianism is committed to a monistic view of what is intrinsically good, which is preference-satisfaction. It bolsters the plausibility of this extraordinary ethical postulate if one limits the preferences so valued to those formed free of coercion of any kind, governmental included. As Mill himself exclaimed, what is the value of a person or his preferences if they are not freely formed?[18]

Fifth, there is the simple psychological truth that people generally prefer to make their own decisions free of coercion. Apart from some perhaps mythical happy slaves,[19] people almost universally prefer their decisionmaking to be free rather than constrained. Therefore, from this fact alone, a committed utilitarian must give some positive evaluation to negative liberty.

Finally, there are well known costs attendant upon the use of the criminal sanction. I refer not only to direct enforcement costs such as police, courts, and so on, although these are omnipresent costs of criminalizing behavior and punishing it. There are also less obvious costs of criminalizing certain behaviors that are (1) typically carried on in private and witnessed by no one other than the participants; (2) strongly motivated so that criminalization only minimally deters the conduct in question; and (3) typically harmful to no one other than those who willingly participate. Costs of criminalizing such behaviors are well known.[20] They include privacy costs in enforcement; costs in terms of disrespect for laws that predictably will be regularly ignored; opportunities for selective enforcement of laws that predictably will be underenforced; and the costs of funding organized crime by the "crime tariff," namely, the artificial restriction of supply of the goods or services in question in the

18. *See* MILL, *supra* note 8, at 124.

19. *See generally* DON HERZOG, HAPPY SLAVES: A CRITIQUE OF CONSENT THEORY (1989).

20. These costs and the literature on them are summarized in MOORE, *supra* note 11, at 663–65.

face of relatively inelastic demand. These costs can be significant for behaviors with the three characteristics in question, and add weight to the value of not prohibiting such activities.

These six considerations strongly support the idea that negative liberty is of value, even if only instrumentally. Except for part of the sixth, these considerations argue in favor of negative liberty no matter what the character of the acts in question. In general, that is, they suggest reasons that the state should not criminalize behavior. Because of their generality, such values cannot justify the existence of a general right to negative liberty, for that would justify anarchy. The very generality of the value of negative liberty precludes it from having the kind of strength one characterizes as a right. For this reason, the general, instrumental value of negative liberty is best characterized in the most modest terms of a presumption in favor of liberty.[21] Such presumption requires that the burden be placed on the state to justify criminal legislation with reasons of sufficient weight to outweigh the reasons just outlined favoring liberty.

IV. THE GENERAL RIGHT TO LIBERTY

A. Introduction

Even those who are quite sympathetic to the values supporting liberty sketched above are often skeptical about any general right to liberty. The basis for such skepticism is easy to see: From Burke[22] to Bork,[23] certain conservatives, often called "Burkeans," have equated a general right to liberty with anarchy. As they point out, a right to be free of state coercion so general that it is applied to all behaviors would forbid the criminalization of murder just as much as it would forbid the criminalization of riding a motorcycle without a helmet. Criminal law is not compatible with such a general right to liberty.

---

21. The presumption of liberty is usually associated with Joel Feinberg. *See* JOEL FEINBERG, SOCIAL PHILOSOPHY 20–22 (1973). I discuss this presumption in MOORE, *supra* note 11, at 746–50.

22. *See generally* EDMUND BURKE, REFLECTIONS ON THE REVOLUTION IN FRANCE (1790), *reprinted in* REFLECTIONS ON THE FRENCH REVOLUTION 1, 6 (J.M. Dent & Sons 1910). I discuss this and other aspects of Burke's conservatism in Michael S. Moore, *The Dead Hand of Constitutional Tradition*, 19 HARV. J.L. & PUB. POL'Y 263 (1996).

23. *See generally* ROBERT H. BORK, THE TEMPTING OF AMERICA: THE POLITICAL SEDUCTION OF THE LAW (1990).

Such skepticism is compatible both with an insistence that individuals possess many discrete natural rights, such as the right to free speech, free exercise of religion, and the like,[24] and with a positive evaluation of liberty in some sense. Recognizing that there are discrete rights does not collapse into anarchism because such rights do not protect all behaviors. And, as Locke saw, natural-rights philosophies generally are already doctrines of freedom in a sense; a right (at least on the choice conception of it)[25] essentially gives its holder the power to waive it or to exercise it. This is unlike moralities, such as Thomistic natural-law views, where duty, instead of rights, is basic.[26] In some sense, therefore, even discrete rights are part of our notion of liberty.

But liberals of the libertarian cast have always hoped for more. They have hoped to make sense of some general right to liberty in a way that avoids both the Burkean reductio to anarchism and the skeptical flight to liberties (discrete rights). In my own writings, I have suggested that there are two such general rights, what I call the derived and the basic rights to liberty.[27] I will briefly sketch each below.

B. The Derived Right to Liberty

The derived right to liberty derives its name from the fact that it is not basic. The citizen's derived right to liberty is merely the correlative of the more basic duty of a legislature. To understand the content of such a right will thus require us to understand the content of the more basic legislative duty.

The central idea is that it is the duty of a legislature to promote certain ends, but not others, in its use of the criminal sanction. Thus, a theory of proper legislative motivation is required: a theory of what sort of reasons are the right reasons with which a legislature can justify coercing its citizens with the threat of criminal sanctions. For such a derived right to exist, one must thus put aside J.F. Stephen's well known, majori-

---

24. DWORKIN, *supra* note 3, and RAWLS, *supra* note 13, rely on just such a distinction: They urge that we each have many liberties (discrete rights), even though we have no general right to liberty.

25. On a choice-oriented theory of rights, it is essential to rights that their holders can choose to exercise them or not.

26. This is a contrast that was crucial (at least pre-Anita Hill) in the Clarence Thomas confirmation hearings. *See* Joseph R. Biden Jr., *Law and Natural Law: Questions for Judge Thomas*, WASH. POST, Sept. 8, 1991, at C1.

27. *See* MOORE, *supra* note 11, at 750–77.

tarian-based skepticism about such a theory.[28] Stephen argued that, in a democracy, the only limits are procedural, so that if the majority thinks that a practice should be criminalized, that is justification enough.[29]

The plausibility of Stephen's view depends in large part on the implausibility of any proposed limits on the proper ends of criminal legislation. The most famous proposal is John Stuart Mill's: Legislation aimed at deterring behavior harmful to others might be justified, but legislation aimed at deterring behavior harmful only to the actor, offensive to others, or thought to be immoral, cannot in principle be justified.[30]

Stephen objected to Mill's "harm principle" on the ground that there is no behavior harmful to self that is not also in some way harmful to others.[31] Against Mill's interpretation of a derived right of liberty, this is a toothless objection. The derived right focuses on legislative aim; even if all behavior harmful to self is also harmful to others, that does not mean that a legislator, motivated by the prevention of one sort of harm, must also be motivated by the prevention of the other sort of harm. Even if helmetless motorcycle riding always harms others when it harms the motorcycle rider (for example, bloody street scenes, higher medical insurance premiums for all of us, and so on) that does not mean we cannot distinguish the prevention of these harms as separate aims of criminal legislation.

Even so, Mill's harm principle is not a good statement of the content of the derived right to liberty. The fault lies deep in the four-part taxonomy of possible legislative reasons that Mill uses: harm to others, harm to self, offense to others, and harmless wrongdoings.[32] In a nutshell, Mill should have formulated his principle and his taxonomy in terms of wrongs rather than harms.[33] The derived right to liberty then becomes the right to have criminal legislation aim only at the punishment or prevention of moral wrongdoing, not other goods. The harm principle has some plausibility only because most serious moral

---

28. *See* STEPHEN, *supra* note 9.

29. *Id.* at 96–97.

30. *See* MILL, *supra* note 8, at 80–81.

31. *See* STEPHEN, *supra* note 9, at 17.

32. This taxonomy was developed in great detail in Joel Feinberg's magisterial quartet of books: JOEL FEINBERG, HARM TO OTHERS (1984); JOEL FEINBERG, OFFENSE TO OTHERS (1985); JOEL FEINBERG, HARM TO SELF (1986); JOEL FEINBERG, HARMLESS WRONGDOING (1988).

33. *See* MOORE, *supra* note 11, at 642–52, 753–62.

wrongs happen to involve causing harm to others. But there are harmless wrongs, such as cruelty to animals or desecration of the dead, and wrongless harms, such as economic injury by fair competition. The proper focus of criminal legislation is wrongdoing, harm-causing being reduced to a rough proxy, but no more than that.

Notice that, on this view of the derived right to liberty, legislators must make substantive moral judgments about the behavior that they would criminalize,[34] and so must their critics, including courts exercising the power of judicial review. The derived right to liberty thus provides no amoral, neutral basis for objecting to the criminalization of homosexuality, abortion, assisted suicide, and so on. Quite the contrary: To argue that such criminalizations violate the derived right to liberty is to argue that they are not immoral behaviors. The disagreements between conservatives and liberals on such issues thus cannot be cast, as liberals are wont to do, as some second-order disagreement about what are and are not proper aims of legislation. Rather, the disagreement is a first-order, moral one: liberals denying that, for example, homosexuality is immoral and conservatives insisting to the contrary. Honestly expressed disagreements about the propriety of various criminalizations must really be about the morality of the behaviors being criminalized, not about the propriety of considering such morality.

C. The Basic Right to Liberty

I call this second right the basic right to liberty because it is not merely the correlative of some more basic legislative duty. Rather, the basic right is basic in the sense that the legislature's duty is the correlative of such a right, not the other way around. From this difference between the two rights flows a second difference: The basic right protects acts themselves from criminal coercion; it does not merely protect citizens from badly motivated legislation about such acts. In the motorcycle helmet case, for example, the basic right protects the act of riding without a helmet from governmental coercion, no matter how that coercion may be motivated, whereas the derivative right protects riders against legal coercion of helmetless riding only if that coercion is motivated by paternalistic or other illegitimate concerns.

---

34. On this Michael Sandel and I have long agreed. *See id.* at 652.

This more ambitious version of a general right to liberty runs head-on into the Burkean objection that such a general right is equivalent to anarchy. Unlike the derived right, the basic right to liberty secures an immunity against legal coercion for actions themselves. To avoid the Burkean objection, one has to limit the sphere of actions protected by such a general right, on pain of the entire criminal code being shown to be an illegitimate intrusion upon liberty. The trick is to specify a content to such a basic right that is limited to the protection of only certain kinds of actions.

It might seem that we already have such content if we combine the presumption of liberty and the derived right to liberty. That is, one might, from the foregoing discussion, rightly conclude that (1) all acts are presumptively protected from legal coercion by the six values behind the presumption of liberty; (2) not just any good state of affairs can justify the overcoming of this presumption—only the punishment or prevention of moral wrongdoing is eligible to justify the use of criminal sanctions; and therefore, (3) the basic right to liberty protects those acts insufficiently wrong that the goodness of their punishment or prevention is outweighed by the strength of the six values behind the presumption in favor of liberty. Less abstractly, (1) riding a motorcycle without a helmet is presumptively protected against legal coercion; (2) the only relevant motivation for coercing this behavior with criminal sanctions is the moral wrongness of it; and (3) this behavior is insufficiently wrong that the goodness of its punishment or prevention does not outweigh the six good things obtained if we do not punish it.

There is nothing wrong with this balancing calculus yielding in some sense a basic right to liberty. Such a calculus protects acts themselves, not just illicitly motivated coercion of acts. And the outcome of such a calculus does not protect all acts from criminalization, so it does not justify anarchy. Still, there are reasons to seek a content to a differently defined, more stringent, basic right to liberty. One reason is the rather contingent nature of the basic right to liberty if defined by the balancing calculus above. Whether such right protects a specific type of action like motorcycle riding depends on a complex calculation that easily can come out otherwise if either side of the balance changes. This contingency corresponds poorly to the crisp, absolute, almost strident character of the basic right to liberty as it was propounded by Mill. In Mill's famous lan-

guage, with such a right we seek "one very simple principle, as entitled to govern absolutely the dealings of society with the individual in the way of compulsion and control."[35] Instead of such an absolute, simple principle, the balancing calculus above is no more than an application of what consequentialist rationality generally demands: Balance up the competing goods to be obtained by coercing or not coercing a behavior, and coerce only when the balance is favorable in that direction.

Second, any basic right to liberty defined by the balancing calculus above shares with the presumption of liberty that is part of the calculus an indifference to the character of the acts protected by liberty. The liberty to murder is as valuable as the liberty to ride a motorcycle without a helmet, the difference between them lying elsewhere: namely, in the justifications for overriding that value. This goes against many people's sense that such liberties are not of equal value.[36] Rather, the intuition is that some kinds of behaviors are strongly protected by the value of liberty, so strongly that even strong and proper legislative motivations to coerce cannot overcome them. Abortion, for example, is thought by many to be a serious moral wrong, approaching the wrongness of killing a person. Yet for some, the libertarian intuition is compelling: Although the state acts with the right motive if it prohibits abortion because of its moral wrongness, and although the punishment or prevention of that wrongful behavior is a very important good, still the woman's right to liberty trumps these justifications for prohibition. There is something special, in other words, about this kind of behavior (abortion) that protects it from well motivated, seemingly justified, legal coercion. It is whatever this "something special" might be that we seek when we look to give some content to the basic right to liberty different from the content yielded by the balancing calculus above. Such alternative content should be both narrower in the class of acts protected and more stringent in the degree to which such acts are protected against legal coercion.

This was Mill's quest: to define a sphere of action absolutely immune to state coercion.[37] Mill's own proposal, here as well as

---

35. MILL, *supra* note 8, at 80.

36. *See, e.g.*, RAZ, *supra* note 1, at 13 ("We feel intuitively that some liberties are more important than others.").

37. This was Mill's second quest in *On Liberty*, as I interpret him. The first quest was to develop a theory of proper legislative motivation for coercive legislation,

with the derived right to liberty, was in terms of his harm principle: only behavior harmful to others could be prohibited. Behavior that harms only the actor or his consenting co-participants, or which harms no one, is to be protected in Mill's version of the basic right to liberty.

This is not the place to rehearse in depth the well worn objections of myself and others showing that the harm principle cannot give the content of the basic right to liberty.[38] My main objection is the one mentioned earlier with regard to the analogous use of the harm principle to give content to the derived right to liberty. Concern with harm is undermotivated here because what matters for justified criminalization is the moral wrongness of the behavior, not its harmful effects on others. It is only the large overlap between serious moral wrongs and the causing of serious harms to others that lends superficial plausibility to Mill's harm principle here. But it is a plausibility that dissolves with the realization that there are some serious moral wrongs that harm no one, and that thus may be prohibited, and there are some serious harms to others which are not morally wrong, and thus should not be prohibited. One could, of course, amend Mill's harm principle, making it into a "moral wrong principle," but this would return us to a basic right to liberty of no greater stringency than that defined by the balancing calculus discussed above.

Mill's quest to define the basic right to liberty in terms of a sphere of action immune to state coercion is thus still very much an uncompleted quest. Not only did Mill himself fail to complete such a quest, but alternative attempts at doing so have not been very reassuring. I will briefly mention three of them.

The first is based on a return to the balancing calculus earlier described. Suppose we were to impose on it an indirect consequentalist structure. This is the familiar strategy of restricting consequentialist balancing—here, a balancing of the six values behind the presumption of liberty against the goodness of punishing or preventing certain acts—to general classes of actions only. And then, when the question to be resolved is whether to prohibit more specific kinds of action, one restricts oneself to the formal question of whether the behavior to be prohibited

---

which gives the content of what I call the derived right of liberty. *See* MILL, *supra* note 8, at 82–83.

38. *See* MOORE, *supra* note 11, at 751–62.

does or does not instantiate the general type of behavior generally justified as immune to state prohibition. So, for example, one might say, à la an amended Mill, that all and only seriously immoral actions may be criminally prohibited, because that is where, in general, the balancing calculus above comes out. Then, when considering a prohibition on abortion or helmetless motorcycle riding, one only asks whether these are seriously immoral behaviors.

My own view (widely shared, I believe) is that all such two-level (or "indirect" or "rule") consequentialisms are literally incoherent.[39] This stems from the tendency of such two-level consequentialisms to collapse either to some kind of long-range act-consequentialism, with rules as mere heuristics guiding the long range calculations required by this kind of act-consequentialism, or to a non-consequentialist absolutism. Either of these moves is quite sensible; one can be a dedicated act-consequentialist or an equally dedicated deontologist without evidencing in either case obvious irrationality. What one cannot be is a rule-consequentialist. For a consequentialist to adhere to a rule in the face of a consequentialist calculation showing that on this occasion more good would come from not following the rule, amounts to inexplicable, reason-less, and arbitrary "rule-worship."[40] One cannot thus simply superimpose a categorical flavor to the balancing calculus earlier described. That there is a familiar strategy for doing so ("rule-consequentialism") does not mean that the strategy makes any sense.

Likewise, there is no hope for a second, non-Millian definition of the content of the basic right to liberty. This is the definition of the content of the right in terms of the related distinctions between acts and omissions, doings and allowings, doings and enablings (or aidings), doings and preventings, and doings and double-preventions.[41] Examples here are clearer than generalities, given the philosophical thickets in which these distinctions are now embedded. Consider, therefore, the

---

39. The locus classicus of this view is DAVID LYONS, FORMS AND LIMITS OF UTILITARIANISM (1965).

40. This is the conclusion (and terminology) of a dedicated act-utilitarian, Jack Smart. *See* J.J.C. SMART & BERNARD WILLIAMS, UTILITARIANISM: FOR AND AGAINST (1973).

41. Most of these distinctions are explored in MOORE, *supra* note 11, at 689–703, and more recently in Michael S. Moore, *Causal Relata*, 14 ANN. REV. L. & ETHICS (forthcoming 2006).

famous example of Judy Thomson:[42] You go to the symphony and the lead violinist collapses. You uniquely possess the life-saving potential to save him if you will only allow yourself to remain hooked to him for nine months. You are hooked up to him over your protest or while you are unconscious. Thomson's question: May you now unhook? Thomson's libertarian intuition: Liberty permits it even though it means the violinist will certainly die.

A lot of work is done in sustaining Thomson's intuition by the doing-allowing distinction. While unhooking is an act, not an omission, it is still only an allowing, not a doing or a causing. By unhooking you return the violinist to the appropriate moral baseline, which is, he is not yet hooked up, and you then allow nature to take its course. Such allowing here takes the form of a double prevention: You prevent by unhooking what would have prevented his death.

From such examples one might conclude that the sphere of liberty immune to state prohibition consists of all the things that are not doings or causings: preventions, omissions (which are failures to prevent), double preventions (which are preventions of what would have prevented the death), allowings, enablings, aidings, and the like. On this view, the state rightly prohibits killings; it cannot prohibit failures to rescue (omissions) or allowings of death.

Everyone recognizes to some degree the intuitive pull of this brand of libertarianism. That is because everyone recognizes the intuitive difference in moral stringency between positive and negative duties, as well as the intuitive difference in allied distinctions. Yet the differences here are not strong enough to yield the categorical nature needed by a basic right to liberty. Contrary to the assertions of some rabid libertarians, it is wrong not to rescue a stranger; it is just not nearly as wrong as it would be to positively kill him. Seeing this, one then could agree with Mill that the goodness achievable by the prohibition of omissions is usually outweighed by Millian autonomy and the other values lying behind the presumption of liberty. But

42. *See* Judith Jarvis Thomson, *A Defense of Abortion*, 1 PHIL. & PUB. AFF. 47, 48–49 (1971).

not always: If the omission is a serious moral wrong, the balance may come out the other way.[43]

All of this is fine, but it fails to give categorical force to the basic right to liberty. Rather, it returns us to the balancing calculus mentioned earlier. Only those more radical libertarians—John Locke, Robert Nozick, Richard Epstein, Eric Mack—who believe that there is no moral wrongness involved in omissions, allowings, and so on, have much of a chance of deriving a categorical right here. Yet, to reverse an old saying, one person's valid inference is another person's reductio: Isn't it morally absurd to think we have no duties to strangers to prevent their death or that there is nothing wrong in allowing them to die? I myself have argued that we have no categorical (deontological, agent-relative) duties here, but we do have duties of other, less stringent kinds.[44]

The third non-Millian attempt to define a content to the basic right to liberty is loosely Aristotelian in character. On this view, the sphere of liberty is to be confined to those fundamental choices that define who we are as a person. I call this broadly Aristotelian because it was Aristotle who famously proclaimed that we each in some sense could choose our character. Choices such as those to beget a child or with whom to be sexually and emotionally intimate are protected; choices to ride a motorcycle without a helmet are not. This is the view animating the last forty years of U.S. Supreme Court jurisprudence on the misnamed "right of privacy," from *Griswold v. Connecticut*[45] to *Lawrence v. Texas*.[46] Its most forthright expression is found in the dissents of Justices Stevens and Blackmun in *Bowers v. Hardwick*.[47]

This third view too is beset by a number of difficulties, not least of which is clarifying exactly what is meant by "self-defining choices."[48] But I shall content myself here with one worry, in line with like worries expressed earlier about the other views of the basic right to liberty. The worry stems from the substantive thinness of the "Sixties values" of authenticity,

---

43. Notice that, for omissions, Mill uses the moral wrongness of omitting to separate the prohibitable from that which is protected by liberty. *See* MILL, *supra* note 8, at 81–82.

44. *See* MOORE, *supra* note 11, at 688–705.

45. 381 U.S. 479 (1965).

46. 539 U.S. 558 (2003).

47. 478 U.S. 186, 199–214, 214–20 (1986).

48. I explore some of the worries in MOORE, *supra* note 11, at 771–77.

self-actualization, self-realization, and self-definition. However these notions are clarified, they seem insufficiently bounded in the possible choices they might protect to ward off the Burkean objection of anarchy. After all, couldn't one raised in a long family tradition of contract killing find a uniquely meaningful life in choosing to become a competent, flourishing practitioner of this particular line of work? Yet if such choices can be "fundamental" or "self-defining," what is to rule out protecting any form of anarchy under the banner of liberty, so long as such anarchism is "authentic"?

One might of course seek to limit the protected self-defining choices to those that are not immoral or perhaps not sufficiently immoral to warrant prohibition. But this would be to return the basic right to liberty to the balancing calculus conception earlier discussed. The usual move here is to seek to show how "authentic" or "true" self-definition is impossible for seriously immoral choices because of the constraints of human nature.

V. CONCLUSION

James Wilson, often called the legal architect of the American Revolution, told his students at the University of Pennsylvania in the winter of 1791 that one of the three great rights for which the Revolution was fought was the right to liberty.[49] Undergraduate philosophy students in America in the early Twenty-First Century share Wilson's enthusiasm for liberty, at least if my own such classes at several universities, including the University of Pennsylvania, are representative. Liberty is important and we have long appreciated this fact.

Still, it is not unknown in history for people to have gone to the barricades in the name of ideals that, upon careful examination, make no sense. Some have thought liberty to be one of those. As we have seen with the presumption of liberty, liberty as a value is unproblematic. Negative liberty is of value, if only instrumentally so. With respect to a right to liberty, things are more complicated. It is easy enough to make sense of a general right to liberty when viewed in what I have called its derivative sense. More troublesome is the idea that there is a general right to liberty when viewed in what I have called its basic sense.

---

49. JAMES WILSON, AN INTRODUCTORY LECTURE TO A COURSE OF LAW LECTURES 8 (Press of T. Dobson 1791).

This last is no small matter because this is the sense most people have in mind when they think about a right to liberty.

My own current take on this last issue is agnostic. On the one hand, I share the strong inclination of many to think that there must be such a right. Abortion, assisted suicide, a variety of parenting decisions, occupational choices, harmful speech, and so on, all suggest the protection of a right more stringent than could be generated by only the presumption of liberty in combination with the derived right to liberty. On the other hand, I have yet to see a morally defensible perimeter to be drawn for Mill's sphere of liberty, despite no dearth of attempts by many talented people.

NOT NECESSARILY IN CONFLICT: AMERICANS CAN BE BOTH UNITED AND CULTURALLY DIVERSE

JENNIFER C. BRACERAS*

In considering whether identity is freedom-limiting or freedom-enhancing, it is important to focus on the interplay between one's identity as an American and one's identity as a member of a particular racial or ethnic subgroup of Americans. Our collective identity as Americans is based on shared values of individual freedom, political democracy, and a commitment to the rule of law. These ideals transcend racial and ethnic differences.[1] They do not, however, eliminate the cultural differences between Americans, nor should they.

Unfortunately, however, some commentators on both the political right and the political left have suggested that one's identity as an American and one's ethnic or racial identity are contradictory or incompatible, and that in order to celebrate one, it is necessary to suppress the other.[2]

There are some commentators on the political right, for example, who wish to ignore and even eliminate most racial and ethnic differences. These commentators incorrectly believe that the suppression of racial and ethnic pride is necessary for effective assimilation.[3] These critics misinterpret the desire to pre-

---

* Commissioner, United States Commission on Civil Rights; J.D., Harvard Law School; B.A., University of Massachusetts-Amherst.

1. *See* Jennifer C. Braceras, *E Pluribus Unum: We Must All Be Americans First*, LEGAL TIMES, Dec. 24, 2001, at 50.

2. *Compare* SAMUEL P. HUNTINGTON, WHO ARE WE?: THE CHALLENGES TO AMERICA'S NATIONAL IDENTITY 339–40 (2004) (arguing for assimilation and acculturation of immigrants in order to save the American national identity) *with* PETER D. SALINS, ASSIMILATION, AMERICAN STYLE 171–72 (1997) (describing the theory of "ethnic federalism" advocated by Louis Farrakhan and Al Sharpton, which places primacy on one's racial subgroup affiliation).

3. *See, e.g.*, HUNTINGTON, *supra* note 2, at 340 (arguing that, in order to reverse the process of decline in the United States, the American people need to renew their Anglo-Protestant identity); PATRICK J. BUCHANAN, THE DEATH OF THE WEST: HOW DYING POPULATIONS AND IMMIGRANT INVASIONS IMPERIL OUR COUNTRY AND CIVILIZATION (2002) (arguing that the United States needs to reclaim its

serve cultural traditions, celebrate the achievements of other group members, or honor a group's distinctive past as a rejection of our shared American culture. They thus find it difficult to understand why, for example, some immigrant parents want their children to learn their native language as well as English; why some Latinos feel special pride in Alberto Gonzales's service as Attorney General of the United States; or why some African-Americans have lobbied for the placement of a National Museum of Black History on the National Mall.

Many of these conservative commentators fear that an emphasis on differences in the *cultural* arena is a slippery slope that will inevitably lead to different—and ultimately unequal—treatment in the *legal* arena. To preserve the idea of equal treatment under the law and the anti-discrimination principle embodied in the text of the Fourteenth Amendment, these conservatives argue that we should ignore race and ethnicity altogether and reject multiculturalism in any form.[4]

If realized, this desire to suppress identity would limit freedom in at least four ways.

First, it would limit the freedom of individuals to celebrate their culture and history in the ways they see fit.

Second, it would limit the freedom of minority groups to honor members of their community and to highlight successful models of assimilation into the American mainstream.

Third, it would limit academic and intellectual freedom to study various racial or ethnic cultures.

Fourth, it would limit the freedom to engage in democratic politics, which is inherently and necessarily a collection of factions or identity group interests.[5]

---

common culture and its Judeo-Christian values in order to prevent the fall of American civilization).

4. This perspective goes well beyond opposition to racial preferences. *See, e.g.*, Brief of Amicus Curiae The Claremont Institute Center for Constitutional Jurisprudence at 23–25, Grutter v. Bollinger, 539 U.S. 306 (2003) (Nos. 02-241, 02-516) (arguing that racial distinctions and classifications are fundamentally at odds with the equality principle of the United States Constitution). Indeed, from this perspective, the mere *collection* of sociological racial and ethnic data is considered perverse. *See, e.g.*, Evelyn Nieves, *California Battles over Racial Identification*, WASH. POST, Sept. 13, 2003, at A4 (describing California's Proposition 54, the "Racial Privacy Initiative," which, if passed, would have amended the state constitution to bar "most governmental agencies in California from gathering information on a person's race, ethnicity or national origin").

5. *See* Richard Parker, *Five Theses on Identity Politics*, 29 HARV. J.L. & PUB. POL'Y 53, 53–54 (2005).

But, while the suppression of ethnic identity is freedom-limiting in these and other respects, the exaltation of ethnic identity over all else[6] poses an even greater danger to the American polity and to freedom generally.

First, an overemphasis on racial and ethnic identity does, as some conservatives fear, undermine our national identity as Americans. When schools fail to inculcate American values, giving short shrift to the history of the American Revolution, the American Civil War, and the American Civil Rights Movement, while emphasizing the history of Africa, Latin America, or Asia, they are severing the ties that bind Americans together in the name of diversity. This is simply wrong.

Second, an overemphasis on racial or ethnic identity encourages segregation in schools, the workforce, the political sphere, and elsewhere. It encourages school districts to herd Latino students into failed bilingual education programs that reinforce the retention of foreign customs and language at the expense of learning English.[7] It encourages employers to relegate Latinos and African-Americans to ethnic enclaves, requiring them to work primarily on minority issues.[8] And it encourages legislatures to draw up racially gerrymandered political districts that serve only to marginalize minority voters.[9]

Third, an overemphasis on subgroup affiliation can, at times, promote a culture of victimhood. Those who define themselves first and foremost by race or ethnicity often naturally come to view themselves not as active participants in civic society, but as part of a disadvantaged and powerless—even disenfranchised—minority. This perspective not only leads to political apathy, it threatens to prevent many individuals from becom-

---

6. *See, e.g.*, HUNTINGTON, *supra* note 2, at 316–24 (noting that some Hispanic activists advocate the rejection of assimilation into America's Anglo-Protestant culture and a separate "'cultural citizenship' involving 'a distinct space for Latinos'") (citations omitted); SALINS, *supra* note 2, at 167–82 (observing the existence of the antiassimilation or "ethnic federalism" movement present in the African-American community).

7. *See, e.g.*, ARTHUR M. SCHLESINGER, JR., THE DISUNITING OF AMERICA 108 (1992); Susan Headden, *Tongue-Tied in the Schools*, U.S. NEWS & WORLD REP., Sept. 25, 1997, at 44; SALINS, *supra* note 2, at 175–76, 188–89.

8. For example, in the name of diversity, a Hispanic reporter might be assigned to cover stories of interest to the Latino community instead of stories related to business and finance; likewise, an African-American law professor might be expected to teach classes on civil rights rather than corporate law.

9. *See generally* ABIGAIL THERNSTROM, WHOSE VOTES COUNT?: AFFIRMATIVE ACTION AND MINORITY VOTING RIGHTS (1987).

ing part of the American mainstream and reinforces regretable segregationist impulses.

Fourth, an overemphasis on racial or ethnic identity promotes stereotypes and limits the freedom of individual members of those ethnic and racial groups. When society encourages people to define themselves primarily by racial and ethnic characteristics, a presumption arises that members of those groups share common values. Although this is often true *as a descriptive matter*, particularly in the cultural arena, the danger lies in the assumption that ethnic authenticity can and should be measured *as a normative matter* by a person's adherence to a particular political orthodoxy.

The dangers of this approach can be seen most publicly in the context of nominations to the judiciary and other top governmental posts. Take, for example, the treatment of conservative minorities, such as Supreme Court Justice Clarence Thomas or attorney Miguel Estrada, President Bush's failed nominee to the United States Court of Appeals for the District of Columbia Circuit. Many on the political left have attacked Justice Thomas for being a black conservative, with some critics going so far as to call him an "Uncle Tom" and an "embarrassment to the Court."[10] Likewise, some on the political left have been particularly vitriolic in their criticism of Mr. Estrada, who has been described as Hispanic "in name only."[11]

Make no mistake about it: The public criticism of blacks and Latinos who dissent from the liberal orthodoxy is different in kind than the criticism directed at white conservatives.[12] As

---

10. *See* KEN FOSKETT, JUDGING THOMAS: THE LIFE AND TIMES OF CLARENCE THOMAS 270 (2004) (describing vitriolic criticism of Justice Thomas); Michael A. Fletcher, *Reid Says He Could Back Scalia for Chief Justice*, WASH. POST, Dec. 7, 2004, at A4 (reporting Senate Minority Leader Harry Reid's criticisms of Justice Thomas).

11. *A Setback for Schumer*, N.Y. POST, Feb. 1, 2003, at 16 (reporting Senator Charles Schumer's description of Estrada as Latino "in name only"). *See also* Jennifer Braceras, *Misguided Judicial Politics*, WASH. TIMES, Feb. 14, 2003, at A21 (describing the mistreatment of Estrada by Democratic senators); Kathryn Jean Lopez, *Power Struggle: Hispanic Republicans in Congress Have Banded Together to Challenge the Powerful Congressional Hispanic Caucus*, HISP. MAG., Aug. 2003, at 21–22 (discussing the response of Republican Latinos in Congress to the Democratic attack on Estrada).

12. For example, people often criticize Justice Thomas while, virtually in the same breath, praising Thomas's colleague and fellow conservative Justice Antonin Scalia. *See, e.g.*, Fletcher, *supra* note 10 (reporting comments of Senate Minority Leader Harry Reid praising the intelligence of Justice Scalia and criticizing Justice Thomas as "an embarrassment to the Court").

Professor Stephen Carter of the Yale Law School has written of the difference in the treatment of white and black conservatives, "To be white and conservative is to be a part of the American mainstream. To be black and conservative is to be a part of the lunatic fringe."[13]

For many on the political left, then, identity has become a means of control, a tool for enforcing conformity. Only those Latinos or African-Americans who genuflect to the liberal establishment are considered "authentic." Those whom the left cannot control are considered "self-haters" or "race-traitors" who should be blocked from public service. In this sense, racial and ethnic identity, as defined by the liberal political elite, has become not only a freedom-limiting concept, but a discriminatory one as well.

Racial and ethnic identity is not irrelevant. As Americans, we cannot ignore our roots. But, although racial and ethnic identity can be something to celebrate, we must be careful to prevent cultural appreciation from evolving into cultural stereotypes that marginalize individuals or become weapons of political control.

What is the best way to strike the balance between multiculturalism, on the one hand, and national unity, on the other?

First, our laws must remain color-blind. To avoid creating resentment between racial and ethnic groups, the legal system must treat all similarly situated Americans the same. This is necessary not only to preserve freedom, but to preserve the rule of law.

Second, it is essential that as a society, we take steps to reinforce the ties that bind citizens together. This can be done by teaching and honoring America's history, government, and the principles of participatory democracy.[14]

Third, we must recognize that although America's laws and social policies must be color-blind, our culture need not be. Individuals do not threaten national unity by celebrating and sharing their distinct cultural traditions or establishing organi-

---

13. STEPHEN L. CARTER, REFLECTIONS OF AN AFFIRMATIVE ACTION BABY 154 (1991).

14. *See* SCHLESINGER, *supra* note 7, at 134 ("The genius of America lies in its capacity to forge a single nation from peoples of remarkably diverse racial, religious, and ethnic origins. It has done so because democratic principles provide both the philosophical bond of union and practical experience in civic participation.").

zations aimed at advancing the interests of one ethnic community or another. To the contrary, as the great liberal historian Arthur Schlesinger has written, "[w]ithin the overarching political commitment, people are free to live as they choose, ethnically and otherwise."[15] This is the very essence of freedom.

Striking the balance between multiculturalism and national unity is not easy.[16] Despite the difficulty of the task, the best way to preserve freedom for all Americans is to maintain a neutral legal system and honor our common culture while simultaneously encouraging all citizens to treasure their distinctive traditions.

15. *Id.* at 135.

16. *See id.* at 138 ("The question America confronts as a pluralistic society is how to vindicate cherished cultures and traditions without breaking the bonds of cohesion . . . .").

THE HUMAN NATURE OF FREEDOM AND IDENTITY— WE HOLD MORE THAN RANDOM THOUGHTS

DOUGLAS W. KMIEC*

In contemplating the relation of freedom and identity, the Latin maxim *libertas non datur sine veritate* aptly reminds us that there can be no freedom without truth. While certain aspects of who we are, such as nationality or ethnic ancestry, may be culturally or serendipitously determined, there is a truth to human nature which, if not observed, corrupts or destroys life and any exercise of freedom dependent upon it. Human nature and the natural law it reflects are inescapable, and, insofar as the Constitution of the United States was consciously fashioned with an outline of human nature in mind, natural law is an indispensable aid to proper constitutional interpretation.

This essay explores the founding conception of liberty and its interrelationship with human nature. It then addresses how the Constitution reflects these aspects of human nature. Finally, it contains some concluding perspectives on aspects of human nature understated in the constitutional design and what ought to be done when there are disputing conceptions of human nature.

I. LIBERTY

The founding view of liberty was taken up directly by Hamilton. In *Federalist 15*, Hamilton asks "why," if man[1] is naturally

* Caruso Family Chair and Professor of Constitutional Law, Pepperdine University; Dean and St. Thomas More Professor, The Catholic University of America Columbus School of Law, 2001–2003; Professor and Director of the Center on Law & Government at the University of Notre Dame, 1980–1999; Assistant Attorney General and Head of the Office of Legal Counsel, Department of Justice, 1985–1989.

1. The use of the masculine in this essay is intended to include the feminine; the masculine usage is continued in the essay so as not to raise in the mind of the reader any inference that the thoughts expressed are somehow at odds with the quoted material from the founding period, which reflected a different custom in

free,[2] "has government been instituted at all?"[3] Hamilton's answer is blunt and rests squarely on a claim about human nature. Government is instituted, Hamilton asserts, "[b]ecause the passions of men will not conform to the dictates of reason and justice without constraint."[4] Liberty without restraint will not lead to private or public good.

How does Hamilton know this? Well, he says, just look around; and further, if the evidence of our own eyes does not convince us and we seek something beyond this empirical claim, he urges us to draw yet another inference about human nature: It is to be expected that men in a collective or group will act badly because the "[r]egard to reputation has a less active influence."[5] Think about it, Hamilton admonishes: Liberty will be badly used if joining together obscures accountability. Moreover, "a spirit of faction" will aggravate these intrinsic human aspects, thereby magnifying the resulting harms.[6] In a group, we will ally with others of like mind in a shameless way to disadvantage or harm others. We will be inclined to use our liberty to pursue "improprieties and excesses, for which [we] would blush in a private capacity."[7]

The desire for liberty to be well used, once "we the people" were united in political society, greatly motivated the Founders. It will be argued below that this founding conception of liberty informed by human nature accounts for much of the constitutional structure and the express limitations upon government power within and appended to it. The justification for the new Constitution is forthrightly anchored in the maintenance of human nature as the "great principle of self-preservation."[8] As such, the precondition for liberty to be used well is honoring the core principle of preserving the truth of oneself—a proposition traceable, as *Federalist 43* expressly af-

---

using the masculine pronoun alone, but which this author believes is applicable to all persons without gender distinction.

2. Note that any other presupposition is counterfactual, except to extreme behaviorists. *See, e.g.*, Thomas Szasz, *Against Behaviorism: A Review of B.F. Skinner's* About Behaviorism, 5 PSYCHOL. NOTES (1991), *available at* http://www.libertarian.co.uk/lapubs/psycn/psycn005.pdf.

3. THE FEDERALIST NO. 15, at 110 (Alexander Hamilton) (Clinton Rossiter ed., 1961).

4. *Id.*

5. *Id.* at 111.

6. *Id.*

7. *Id.*

8. THE FEDERALIST NO. 43 (James Madison), *supra* note 3, at 279.

firms, to "the transcendent law of Nature and of Nature's God" upon which the nation is founded and all human action depends.[9] The question whether identity is a limitation or starting point for freedom may be a puzzler for twenty-first century man, but it is an easier question when tossed the way of Publius. The authors of *The Federalist Papers* knew human nature or identity to be *the* starting point for human freedom or liberty.

II. HUMAN NATURE

What is the law of nature? An early twentieth century lecturer put matters nicely:

> Every living creature is the embodiment of some form of natural law. Its duration of life depends upon its obedience to the law of its nature, as embodied in its organism. It lives by being itself, by persisting in being itself, and when it violates the law of the kind of being it is, it renounces life and perishes . . . . All animated beings are subject to the laws of cause and effect, as Nature has prescribed them for each species . . . . [I]n any complex organization, like human society, something must be freely granted to the individual. This is what we mean politically by "liberty." On the other hand, something must be insisted upon for the benefit of the group. This is what we mean by "law," in its social sense . . . . Without liberty, there is no initiative, and hence no progress. Without law, there is no survival of the group.[10]

It is within the will of man to have positive law either advance human nature or undermine it. It would be a mistake, however, to suppose that the constructed, positive law of society can disregard the law of nature without consequence. We can construct governments and other social structures beyond our individual natures, but these perform well only if nature's truths are observed. "What we must never forget is that Nature never ceases to govern; and that, if men wish to govern, they must govern under Nature's Laws, or they will be doomed to failure."[11]

---

9. *Id.*
10. DAVID JAYNE HILL, HUMAN NATURE IN THE CONSTITUTION 24–25 (1926).
11. *Id.* at 29.

A. Human Liberty or "Right" Derived from Assumed Duty

The founding generation was studied in the dismal history of societies that sought the false freedom of governing *against* human nature. If one begins the history of human government with the patriarchal clan, one sees force, but little acknowledgment of human liberty or freedom.[12] As the clan gave way to various forms of warrior chiefs and kingships, there was a natural mindfulness of the well-being of one's group. Several thousand years before Christ, Hammurabi's famous legal code would describe the clan leader as a shepherd chosen "to care for the people [and cause them] to dwell in peace and security, that the great should not oppress the weak."[13] The Greeks would give a name to these assumed natural duties of care, and these in turn would later become encapsulated into the notion of rights or liberties. Rights, therefore, arose as correlatives from the reasoned objection of man's intellect when leaders defaulted on their expected duty of care and irrationally deprived man of the necessary goods or sustenance to survive. Stoic philosophers like Cicero would bring this conception of human right or liberty derived from duty to Rome, but, with Rome's fall, barbarian kings once again obscured the concept of natural rights. It would not re-emerge until the American Founders decided to build a government upon human nature and its associated rights.

B. Affirmation of Creation as Source of Natural Right or Liberty

"When . . . the Laws of Nature and of Nature's God entitle them, [and] a decent respect to the opinions of mankind requires . . . .":[14] With these words, the Founders gave explanation not only for the formation of a new sovereignty, but also the human liberty the newly established United States sought to advance. It was an explanation premised upon the proclaimed truth that man is not self-creation, but created. That the handiwork of the Creator came with a conscious endowment of unalienable rights—life, liberty, and the pursuit of happiness—led to the conclusion that any governmental form

---

12. This is not to say that force within the clan was always contrary to human nature.

13. HILL, *supra* note 10, at 37 (quoting Hammurabi's legal code from approximately 2250 B.C.).

14. THE DECLARATION OF INDEPENDENCE para. 1 (U.S. 1776).

that followed would need to have one overriding end: to secure human nature and the rights associated with it.

There is much speculation as to why America in the late Eighteenth Century was the locus of natural law rediscovery. Possibly, it was the wide sweep of land, and nature, itself, that the colonists daily inhabited and sought to harness. Perhaps it was the extraordinary discoveries of the era in natural science. Or it may simply have been that no people so distant from their country of origin could rationally continue to think of themselves as "subjects." American colonists were *persons* enjoying natural liberty. However it was, "[t]he American colonists came upon this idea in their own way . . . . It was the result of their own experience in self-government, coupled with their faith that their human nature had a Divine origin and involved a moral responsibility of which freedom was a necessary correlate."[15]

If freedom, and the new government that aspired to it, were to be guided by human nature, then that nature would need to be understood. At a very basic level, giving proper attribution to a Creator put human nature off-limits to human redefinition[16] and secured unalienable rights against the government,[17] but a workable government would require some greater identi-

---

15. HILL, *supra* note 10, at 51–52.

16. Such attempted human redefinition of truth, unfortunately, is all too common. For instance, disputes abound over what legal protection to extend to an unborn child ever since *Roe v. Wade*, 410 U.S. 113 (1973), but the legal dispute has absolutely no effect on the truth of the child's humanity. So too, any association of persons can be legally called a marriage, but such domination has no effect on the truth of what marriage is in terms of conjugal unity and procreative potential. Moreover, because legal assertions have no discernible power to redefine the natural essence of these matters, man ought not to seek to have positive law and nature work toward different ends. Justice James Wilson, who signed both the Declaration of Independence and the Constitution, sagely counseled that "law can never attain either the extent or the elevation of a science, unless it be raised upon the science of man." JAMES WILSON, *Man as an Individual—Abstractly Treated*, *in* 1 THE WORKS OF JAMES WILSON 206, 207 (James DeWitt Andrews ed., 1896). It was obvious in the 1850s that black men and women were human. Nonetheless, the law pointedly chose to treat them inhumanly. The bloody consequences of the law's impertinence in ignoring human nature are etched in history.

17. *See* Thomas L. Pangle, *The Philosophic Understandings of Human Nature Informing the Constitution*, *in* CONFRONTING THE CONSTITUTION 9, 74 (Allan Bloom ed., 1990) ("This means . . . while the majority retains supreme political power, it does not retain and never had unlimited power. The supreme (irresistible) power governing every rational person's behavior is the desire for self-preservation and every individual retains the inalienable right to resist perceived threats to his property and existence, no matter what the source of those threats.").

fication of the particulars of human nature. A few of these particulars are explored below.

III. MAN IS FREE, BUT NOT APART FROM OR ABOVE, SOCIETY

In the Seventeenth and Eighteenth Centuries, it was understood that man was not intended to live alone, but *in* society. Of course, part of this sociability was a product of pure necessity. "From his beginning [man] was born into [society], and without it he could never have been. Helpless in his isolation, he could be exterminated even by swarms of insects."[18] But the yearning for community was more than a utilitarian means of defense against predatory animals or other threats to physical existence. The Founders read Aristotle and accepted his proposition that "man is by nature a political animal . . . . There is . . . a natural impulse in all men towards an association [with others]."[19] This natural desire, according to Aristotle, arose from two sources: reasoned reflection on right and wrong (which is only a comprehensible exercise in relation to others) and our love of others.[20]

A. Jefferson: Man Has a Moral Sense Developed Out of Service to Others

Thomas Jefferson most notably made reference to man's social side, observing in correspondence to John Adams that man is "an animal destined to live in society."[21] For this reason, Jefferson would deliberately criticize the anti-social, atomistic conceptions of Hobbes as a "humiliation to human nature."[22] Thomas Pangle records that Jefferson had derived from the Enlightenment philosopher Helvetius that we experience pleasure "when we aid or even when we *seem* to sacrifice for others."[23] Jefferson was not fully satisfied that Helvetius had explained the origin of the pleasure derived from the service to

---

18. HILL, *supra* note 10, at 17.

19. ARISTOTLE, THE POLITICS 10–11 (Ernest Barker trans., 1995).

20. *Id.* at 106.

21. Letter from Thomas Jefferson to John Adams (Oct. 14, 1816), *in* 2 THE ADAMS-JEFFERSON LETTERS 492 (Lester J. Cappon ed., 1959).

22. Letter from Thomas Jefferson to Frances Gilmer (June 7, 1816), *in* THE WRITINGS OF THOMAS JEFFERSON 24 (Albert Ellery Bergh ed., 1905).

23. THOMAS L. PANGLE, THE SPIRIT OF MODERN REPUBLICANISM: THE MORAL VISION OF THE AMERICAN FOUNDERS AND THE PHILOSOPHY OF LOCKE 120 (1988).

others and was unprepared to ascribe the origin of man's moral sense solely to God since that would leave unaccounted for the moral sense or like sensation in a disbeliever. Therefore, on a philosophical level, Jefferson would conclude that, like other aspects of the moral sense in man, nature simply reveals the pleasure of service.[24] As he grew older, Jefferson would come to value tranquility over continued public service,[25] but he would continue to lean upon the theorem that the pursuit of happiness was dependent upon the virtue of knowing oneself and being useful to others. The "moral instinct" that inclines us to do good out of a love of others is, Jefferson would conclude, "the brightest gem with which the human character is studded, and the want of it as more degrading than the most hideous of bodily deformities."[26]

B. Wilson: Man Has Moral Sense Because He Has an Innate Conscience

James Wilson would question Jefferson's reliance upon the pleasure or utility of serving others as a sufficiently reliable basis for the development of a moral sense. Unlike Jefferson, Wilson would insist that human nature intrinsically includes not only a desire to be social and socially useful, but also a conscience.[27] Relying upon Thomas Aquinas by way of Richard Hooker, Wilson would insist that it is conscience that guides reason.[28] The first principles of virtue are self-evident to man, and, were it otherwise, most men would find the pursuit of virtue to be impossible and beyond their capacity.

Wilson's attachment to innate conscience contrasts with John Locke, who, in his *Essay Concerning Human Understanding*, opines that not even the Golden Rule is innately known. Locke is often quoted in a way that makes his writing seem highly relativistic, and certainly the statement, "[*c*]*onscience* . . . is nothing else, but our own Opinion," seems to be just that.[29] Locke was obviously a stronger influence on Jefferson than on Wil-

---

24. *See id.*

25. Pangle quotes Jefferson as advising a young James Monroe that "public service and private misery [are] inseparably linked together." *Id.* at 121.

26. *Id.* at 120.

27. *Id.* at 121–22.

28. JAMES WILSON, *Of the General Principles of Law and Obligation*, *in* SELECTED POLITICAL ESSAYS OF JAMES WILSON 215, 222–24 (Randolph G. Adams ed., 1930).

29. JOHN LOCKE, AN ESSAY CONCERNING HUMAN UNDERSTANDING 70 (Peter H. Nidditch ed., 1975) (1689).

son, as Locke's denial of conscience as innate fits nicely with Jefferson's proposition that men always inquire further to seek an underlying reason for a moral rule.[30] For Jefferson again, it was the utility of service that brought happiness, not following an inner voice guided by an objective, knowable virtue. Locke would similarly write that "[p]ower and riches, nay Vertue [sic] it self, are valued only as Conducing to our Happiness."[31]

C. Man's Created Nature Bridges Jefferson and Wilson

But Jefferson (and Locke) may not be as far from Wilson as it first would seem. What unifies them is reference to the transcendent. All three concede that acknowledgment of a Creator influences man's moral sense. Locke makes repeated reference to man's creation, and Jefferson's "endowed by their Creator" reference in the Declaration is well known. Nevertheless, Jefferson is sometimes described as a "materialist," a term he borrowed from Locke, or often as a "deist." These terms obscure more than they clarify because it was Jefferson's concession of a Creator God that had real consequence for filling out his conception of human nature. As Father John Hardon, S.J., wrote in apprising the so-called Jefferson Bible, the *Life and Morals of Jesus of Nazareth*:

> That Jefferson believed in God is evident first from his ready acceptance of the teachings of Christ on the subject, the Lord's Prayer, the Eight Beatitudes, the Parables of the Unjust Steward and the Ten Talents, the Sermon on the Mount—all of which presuppose a belief in the existence of God, the Creator of heaven and earth. Correlative with this goes the belief in prayer and some kind of Providence, and to that extent, at least, an acceptance of some kind of grace, requested for example in the petition, "Deliver us from evil," in the *Pater Noster*.
>
> Also the *Morals of Jesus* allows us to conclude that Jefferson believed in some sort of future life, where the good are rewarded and the wicked punished. Besides the Parables of Lazarus and Dives, of the Pharisee and Publican, and the Wedding Feast, Jefferson accepted and extracted the whole discourse of Christ about the Day of Judgment, in the twenty-fifth chapter of Matthew, not excluding the classic

30. *See id.* at 65.

31. John Locke, *Some Thoughts Concerning Education*, *in* THE EDUCATIONAL WRITINGS OF JOHN LOCKE 109, 249 (James Axtell ed., 1968).

> verse 46, in which Christ foretells: "These will go into everlasting punishment, but the just into everlasting life."[32]

What, then, of Jefferson's self-description as a materialist in the Lockean sense? Hardon writes that it was not a denial of the spirituality of the human soul, but merely the humble confession that there is no human proof anchored solely in reason of the soul's spiritual nature.[33] Hardon's explanation is persuasive. Even though Locke (and by extension Jefferson) was unable to prove the imprint of a moral sense in man and questioned whether reason is naturally inclined toward seeking the good, as Aquinas taught, Locke nevertheless insisted on the existence of natural law, knowable only by means of the Divine creator and legislator.[34] By this, Locke meant that human beings are the creation or "workmanship" of God; therefore, they belong to God and are His property.

From this declared status as created beings, a set of prescriptions under the natural law can be deduced. For example, the presupposition of creation allows man to deduce moral precepts in support of "unalienable rights" derived out of his relationship with a Creator-Owner and other created human beings. These moral precepts themselves then encourage habits of virtue, especially including Jefferson's insight of service to others. Habits of virtue yield happiness. Disregard the presupposition of man as a created being, however, and think of man as his own self-creation living outside or above society, and the process would work in reverse: unhappiness resulting from practices of vice and self-interest unchecked by any moral sense derived from human nature. Without the public acknowledgment of man's created nature, the derivation of moral sense would be impossible, because there would then be no

---

32. Fr. John A. Hardon, S.J., *The Jefferson Bible*, AM. ECCLESIASTICAL REV., June 1954, *available at* http://www.catholicculture.org/docs/doc_view.cfm?recnum'6040.

33. *See id.* John Locke writes, "we do not owe our origin to ourselves . . . ." JOHN LOCKE, QUESTIONS CONCERNING THE LAW OF NATURE 161 (Robert Horowitz et al. trans., Cornell Univ. Press 1990) (1664). Locke was sure that this is not a religious doctrine, even as such doctrines may confirm "the truth of our argument that man can, by making use of sense and reason together, arrive at knowledge of some supreme power . . . ." *Id*. at 165. Locke admitted that reason may prompt some to doubt God's existence, but he said "there exists nowhere a race so barbarous, so far removed from all humanity" that is not suited to "infer from sensible things that there exists some powerful and wise being who has jurisdiction and power over men themselves." *Id*. at 165, 167.

34. *See* Michael P. Zuckert, *Do Natural Rights Derive from Natural Law?*, 20 HARV. J.L. & PUB. POL'Y 695, 721 (1997).

stable conception of human nature. Human nature would, of course, be factually constant, but insofar as it would be subject to legal redefinition by those in the possession of force, it would not yield moral clarity for public or private decision.

Of course, man is not assured of happiness merely by public acknowledgment of his created nature. When man enacts laws or undertakes personal action in defiance of that created nature, he is acting in a way that is contrary to a state of happiness. For this reason, if a government of law is to be successful, it must be formed to meet the reality of man's nature: a reality which recognizes both man's created nobility and rebelling imperfection. Hence, Wilson insightfully comments:

> [G]overnment is the scaffolding of society: and if society could be built and kept entire without government, the scaffolding might be thrown down, without the least inconvenience or cause of regret.
>
> Government is, indeed, highly necessary . . . to a fallen state. Had man continued innocent, society, without the aids of government, would have shed its benign influence even over the bowers of Paradise.[35]

The Founders believed man had not "continued innocent" and so shaped American government to meet his shortcomings.

IV. MAN'S IMPERFECT NOBILITY

The seventeenth-to-eighteenth-century period out of which the Constitution emerged was, as Arthur O. Lovejoy records, a period of transition between the denigration of man and the celebration of his potential.[36] Theologians and religious writers reminded the Founders of man's creation in the image and likeness of God and man's supernatural destiny, but one satire writer after another demonstrated that man, in action, failed regularly to live up to this nobility. These satires of the Seventeenth Century were but the flowering of earlier writing. Father James Gillis writes:

> Shakespeare—the myriad-minded Shakespeare—probably knew man better than any other poet or dramatist or philosopher. Certainly he made a life study of man; he tracked

35. JAMES WILSON, *Of the Study of Law in the United States*, *in* SELECTED POLITICAL ESSAYS OF JAMES WILSON, *supra* note 28, at 210.

36. *See* ARTHUR O. LOVEJOY, REFLECTIONS ON HUMAN NATURE 20 (1961).

> every emotion and mood and thought and passion of man to its secret lair in the human heart, dragged it out, incarnated it in man or woman, king, peasant, soldier, student, lover, clown, clothed it in ermine or fustian or in mourner's weeds, and made it "strut and fret its hour upon the stage." If ever a man revealed ourselves to ourselves, it was that all-but-omniscient Shakespeare. But even he was compelled in the end to confess that he couldn't solve the riddle of man. Witness the famous monologue, "[w]hat a piece of work is man!" continuing "how like an angel!" but concluding, "this quintessence of dust!"[37]

A. Man Rationalizes Himself as an Exception: The Self-Interest Problem

Thirty years before the Declaration, the French writer Marquis de Vauvenargues noted how much we enjoy pointing out human defect, thinking we can somehow exempt ourselves from the same criticism. Vauvenargues lamented, "We are so presumptuous that we imagine we can separate our personal interest from that of humanity in general, and malign the human race without implicating ourselves."[38]

In response to this criticism, man would assert as a defense his commitment to reason. However, one would have to cautiously wonder if "reason" was itself rationalization and self-deception. Man reaches a conclusion favoring passion over reason, then finds reasons to justify the passion and deceives himself into thinking the reasons discovered were the cause for the initial decision. Again, satirists of the Seventeenth Century regularly pointed out that "[t]he passions always seek to justify themselves and persuade us insensibly that we have reason for following them. The gratification and pleasure to which they give rise in the mind which should be judging them, corrupt its judgment in their favor."[39]

These insights were best represented in the founding generation by John Adams. He observed that men tend to act first and think after. Men have a tendency to flatter themselves, and Adams thought this self-deception was responsible for many ca-

---

37. JAMES M. GILLIS, THIS MYSTERIOUS HUMAN NATURE 5 (1956) (quoting WILLIAM SHAKESPEARE, HAMLET act 2, sc. 2).

38. LOVEJOY, *supra* note 36, at 20 (citation omitted).

39. *Id*. at 26.

lamities.[40] Caught up in the various political controversies of his day, Adams wondered why those against him sought to "blacken and discredit" his motives, rather than address underlying issues.[41] This trait of human nature has not changed.[42]

B. A Government to Bring Perfection from Imperfect Human Nature

Richard Hooker had faithfully recorded the noble, but imperfect, aspects of human nature:

> Laws politic, ordained for external order . . . are never framed as they should be, unless presuming the will of man to be inwardly obstinate, rebellious, and averse from all obedience unto the sacred laws of nature; in a word, unless presuming man to be in regard of his depraved mind little better than a wild beast, they do accordingly provide notwithstanding so to frame his outward actions, that they be no hindrance unto the common good for which societies are instituted: unless they do this, they are not perfect.[43]

Yet all was not lost. God had created a universe by counterbalancing the forces of physical science, as Newton explained, and man could likewise construct a successful polity by following His model. So the Constitution came to be, following this instruction of counterpoise or balance, reflected in the planets as well as literature. The Founders, already having declared their fidelity to "the Laws of Nature and Nature's God," pro-

---

40. *See generally* JOHN ADAMS, ON SELF-DELUSION, *in* 3 THE WORKS OF JOHN ADAMS, SECOND PRESIDENT OF THE UNITED STATES 432, 433–36 (Charles C. Little & James Brown eds. 1851).

41. *Id*. at 436.

42. Why, for instance, did environmental groups seek to demonize then-Judge John Roberts as anti-environment because of his dissent in *Rancho Viejo v. Norton*, 334 F.3d 1158 (D.C. Cir. 2003), rather than take up the jurisprudential difficulty that animated the Supreme Court decisions upon which precedent Judge Roberts based his dissent—the proper division of state and federal power under the Commerce Clause? *See*, *e.g.*, Katharine Mieszkowski, *The John Roberts Dossier: Everything You Need to Know about Bush's Nominee, Before the Battle Begins*, Salon.com, July 20, 2005, http://www.salon.com/news/feature/2005/07/20/roberts_dossier/index_np.html ("Roberts doesn't seem to be a big fan of the Endangered Species Act, at least considering his attitude toward the arroyo toad . . . . In 2003, Roberts wanted the court to reconsider a panel's decision that upheld a Fish and Wildlife Service regulation protecting the toads under the act. The court declined to hear the case, but in his dissent Roberts maintained there could be no interstate commerce rationale for protecting the toad.").

43. 1 RICHARD HOOKER, OF THE LAWS OF ECCLESIASTICAL POLITY 188 (Everyman's Library ed., 1954) (1593).

ceeded to the practical task of constructing a government, not on the basis of what man ought to do, but what man would, in fact, likely do.

Enter, for example, James Madison and the Federalist solution for faction. Factions—special interests today—pursue their own agendas, while giving little care to the aggregate or common interests of the whole. This could be corrected either by removing the ability of a special interest to act in this way or by mitigating the effect of this natural action. Had Madison and the Framers chosen the former, it is fair to say that in view of an accurate understanding of human nature as imperfect and rebellious, it would have meant greatly constrained freedom. This, of course, had largely been the history of the world to that time. The wills and passions of men running counter to the choices of those in control of armies and great wealth were simply crushed. Even a benign despot nourishing the idea that men could be reformed to see the wisdom of opposing interests would fail. Such indoctrination would not be costless to freedom, and, in any event, it would assume that human nature was malleable in ways that it is not. Since ignoring the truth of human identity would sacrifice freedom, Madison—taking his cue from the Creator of the universe—decided, instead, to mitigate the effects of man freely choosing to follow his own passions. Ambition would check ambition.

A primary means for Madison to reach counterpoise was his belief in an extended republic. National power over truly national subjects would help submerge factional interests in ways that more localized governance could not. Consider a topical Exhibit A: the Supreme Court's decision in *Kelo v. City of New London*.[44] It was far easier for Pfizer and the proponents of commercial redevelopment to capture locally exercised eminent domain power than it was to convince the Supreme Court of the United States to disregard the original meaning of the Public Use Clause to affirm the exercise of that power. Of course, in this instance, a narrow 5-4 majority *did* ignore the words of the Fifth Amendment and approve a regrettable result in favor of faction. Still, the unfortunate *Kelo* result prompted several Justices to point out explicitly how the decision would lead to advantages primarily for those who had legislative friends, to the great disadvantage of racial minorities

44. 125 S. Ct. 2655 (2005).

and those of modest means.[45] Justice Stevens, for the majority, put his faith in comprehensive planning and town meetings.[46] This is a faith in the intrinsic rationality of man, something the Founders might say indulges a delusional conception of human nature. The town meetings merely facilitated a faction's ability to take what was not theirs and reallocate it.

It is a fair objection to contend that if the city of New London could not undertake planning in 2005 without succumbing to faction, was not the same true of the participants in the 1787 Constitutional Convention? Large states versus small ones, farmers pitted against merchants, the convention was rife with faction. Why should not the resulting constitutional product be seen to be as suspect as New London's condemnation plan? Because, unlike Justice Stevens, Madison and his fellow delegates consciously understood that they could not draft a document that exempted man from his natural impulses. They could only fashion one that would place his passions in counterpoise. The Founders were not like the satirists who made fun of man's limitations, thinking themselves above them. Rather, the Founders fully expected man to obey his natural inclinations, which is why they deliberately framed the document to set one interest against another to yield a more perfect or satisfactory result.

V. SOME OVERLOOKED OR UNDERSTATED ASPECTS OF HUMAN NATURE

Two aspects of human nature are less well reflected in the Constitution's design. The first is the desire to be held in the esteem of our fellows. Locke noted in *The Rambler* that:

> Every man, however hopeless his pretensions may appear, to all but himself, has some project by which he hopes to rise to reputation; some art by which he imagines that the notice of the world will be attracted; some quality, good or bad, which discriminates him from the common herd of mortals,

45. *See, e.g.*, *id.* at 2686–87 (Thomas, J., dissenting) (noting that "[s]o-called urban renewal programs" will disproportionately harm poor and politically weak communities).

46. *See id*. at 2668–69 (majority opinion) (noting that the necessity and wisdom of using eminent domain are matters for public debate).

> and by which others may be persuaded to love, or compelled to fear him.[47]

Jesuit thinking, it is supposed, traced this desire for approval to the Creator, who, it was thought, ingeniously implanted it within our natures as a substitute for reason or virtue.[48] Adam Smith in his *Theory of Moral Sentiments* likewise observed that "the love of virtue [is] the noblest and the best passion in human nature" and that the love of true glory is "a passion inferior no doubt to the former, but which in dignity appears to come immediately after it."[49]

This desire for esteem did not totally elude the Framers, for it lends explanation to how Jefferson could so forthrightly identify the pursuit of happiness with the service of others. Nonetheless, it did not receive sufficient attention. A desire for praise stirs human action, causing men to take up arms in common defense when they otherwise might not and to forego leisure for study (and even write essays on human nature) in hopes of some immortality. But where, it might be asked, is this powerful motivation honored as a source of human action in the Constitution? It might be supposed that the various offices of profit and trust provided or enabled there, including the Executive itself, satisfy this craving. However, John Adams thought them wholly inadequate to this task. Adams commented that

> [T]he desire of the esteem of others is as real a want of nature as hunger; and the neglect and contempt of the world as severe a pain as the gout or stone . . . . Every personal quality, and every blessing of fortune, is cherished in proportion to its capacity of gratifying this universal affection for the esteem, the sympathy, admiration and congratulations of the public.[50]

The founding focus was to create checks and balances that counteracted the passions of man, including the excessive de-

---

47. John Locke, *No. 164*, RAMBLER, Oct. 12, 1751, at 282.

48. *See, e.g.*, JUAN DE MARIANA, DE REGE ET REGIS INSTITUTIONE LIBRI III 235–45 (Scientia Verlag Aalen 1969) (1599) ("Who does not see that it is the desire for praise [implanted by Nature] that men are powerfully stirred to undertake great deeds.").

49. ADAM SMITH, THE THEORY OF MORAL SENTIMENTS 309 (Liberty Classics 1981) (1759).

50. 6 JOHN ADAMS, *On Government*, *in* THE WORKS OF JOHN ADAMS, SECOND PRESIDENT OF THE UNITED STATES, *supra* note 40, at 181, 234.

sire for power or fame, but in regards to this feature of human nature there seem to be insufficiently few fonts of public honor to motivate or to supply emulation, which Adams described as "imitation and something more—a desire not only to equal or resemble, but to excel."[51] It may well be that a government "of the people" risks too much inequality if it promotes such honorific distinctions. But if that is the case, we might spend some time contemplating whether the space for unregulated activity in the Constitution sufficiently allows for human achievement to be noted in the spheres of economics and the observance of property rights.[52]

A second overlooked aspect of human nature relates to the failure to work out more completely in the Constitution itself the significance of the corporate acknowledgment in the Declaration of Independence of man's being made, and arguably therefore "owned," by a creator. Where, for example, in constitutional jurisprudence is Locke's sentiment that "[m]an not having the power of his own life, man does not own himself, cannot by compact, or his own consent, enslave himself to any one . . . . Nobody can give more power than he has himself."[53] That is an intriguing philosophical insight, and one that is in a good deal of tension with modern expansive, and expanding, claims of autonomy.

It was observed earlier that the Founders deduced a set of moral precepts from our created human nature. These prescriptions mainly took the form of limitations on what human beings may do: They may not use force against or to directly harm each other, for they belong to God and not to each other; they may not harm themselves, such as through suicide, for the same reason; and they may not indirectly harm each other through taking more than their fair share of the goods from the external world, which is of Divine making.

Does the modern mind resist deducing additional guidance from the public acknowledgment of man's created nature? The Declaration, the founding positive law incorporating the United States as a sovereign nation, puts Creator-skepticism

---

51. *Id.* at 267.

52. *See* THORSTEIN B. VEBLEN, THE THEORY OF THE LEISURE CLASS: AN ECONOMIC STUDY OF INSTITUTIONS 25 (1905) (observing that "[t]he motive that lies at the root of ownership is emulation").

53. JOHN LOCKE, TREATISE OF CIVIL GOVERNMENT AND A LETTER CONCERNING TOLERATION 16 (Charles L. Sherman ed., 1937) (1690).

off-limits as a matter of sovereign, corporate statement. The Constitution, in relation to individual inquiry, does just the opposite, as the Religion Clauses properly ensure that such skepticism is not impeded. But perhaps the Establishment and Free Exercise Clauses can still be fully honored with a fuller understanding of other structural or substantive aspects of the Constitution in light of the corporate acceptance of our created natures. As Walter Berns has observed, the Founders "were fully convinced that the Constitution of the United States derived from a 'self-evident' truth respecting man's nature and the government appropriate to it."[54] As I have argued in an earlier volume of this Journal:

> For similar reasons, the Declaration of Independence cannot be relegated to a constitutionally insignificant aspiration. Jefferson himself understood the Declaration as the embodiment of "the American Mind" and the "harmonizing sentiments of the day." Few originalist traditions will come so well-described in the search for original meaning. As one author described it, "the clarity of [the natural-rights] doctrine in the minds of the Americans is indicated by the first phrase. The statement that 'we hold these truths to be self-evident' . . . indicates that to the colonists the principled argument is well-known and widely accepted." When Jefferson undertook to prescribe a curriculum for the University of Virginia, he instructed the Board of Visitors that it was their "duty . . . to provide that no [principle of government] be inculcated which [is] incompatible with those on which the Constitution of this State, and of the United States were genuinely based." The first text on Jefferson's prescribed list was the Declaration of Independence. The enduring significance of the Declaration can also be seen in the recourse made to it by Abraham Lincoln in the midst of the gravest threat to the Union. . . . [O]ne finds in the Book of Proverbs the meditation that "[a] word fitly spoken is like an apple of gold in a frame of silver." Lincoln likened the Constitution to the frame, and the Declaration to the apple, noting, "the frame is made for the apple, not the apple for the frame."[55]

The apple of gold—human identity—is not clarified but obscured when the Supreme Court understates our created origin

---

54. WALTER BERNS, THE FIRST AMENDMENT AND THE FUTURE OF AMERICAN DEMOCRACY 146 (1985).

55. Douglas W. Kmiec, *Natural Law Originalism—Or Why Justice Scalia (Almost) Gets It Right*, 20 HARV. J.L. & PUB. POL'Y 627, 642–43 (1996–97) (internal citations omitted).

in constitutional jurisprudence. No one should be coerced under law to believe any religious teaching, but would not taking account of man as a created being also mean a lesser readiness to exclude religious reference in the public square?[56] We are free to be atheists, but, given the created identity identified in our founding document of incorporation, the Supreme Court arguably is not.

Our human identity as created beings may be given insufficient attention in other contexts as well. For instance, the implications of created human identity might also enhance an appreciation for federalism. Recently, the federal government sought to enforce federal drug laws in a way that obliterated the distinction between what is national and what is local.[57] In so doing, did it also short-change the understanding of law as designed to advance the well being of created human persons? To many, recreational drug use is self-destructive. It is not irrational to understand drug crime as facilitating everything from gang violence to terrorism. Nevertheless, a woman with a brain tumor using marijuana under a doctor's supervision is not engaged in interstate commercial activity.[58] The failure of the United States in its briefing and arguments to appreciate the difference between locally provided and authorized medical care and an illegal drug market was distressing and deeply threatening to Mrs. Raich's created identity, if not her very existence. Human identity is also obscured when race is made the basis of public decision. The overly casual applications of *Grutter v. Bollinger*[59] may be the proof of this. The Court in *Grutter* proclaimed that it was not relying upon race as a person's sole defining characteristic.[60] Yet a lower court decision has relied upon race as such a characteristic, in approving another public-school diversity plan.[61] An acknowledgment of being created equal would seem to stand in refutation.

---

56. *Cf.* McCreary County v. ACLU, 125 S. Ct. 2722 (2005) (affirming the Sixth Circuit's holding that the county's purpose in erecting the Ten Commandments was not integrated with other material so as to carry a secular message and hence was unconstitutional).

57. *See* Gonzales v. Raich, 125 S. Ct. 2195 (2005).

58. *Contra id.* at 2207.

59. 539 U.S. 306 (2003).

60. *See id.* at 337–38; *see also* Gratz v. Bollinger, 539 U.S. 244, 271–72 (2003).

61. *See* Parents Involved in Cmty. Sch. v. Seattle Sch. Dist., 426 F.3d 1162, 1191 (9th Cir. 2005) (en banc) (using race as a tie-breaker to racially balance public schools).

VI. CONCLUSION

Although other examples of case law seemingly disregarding our created natures might be given, this essay will conclude by returning to the assigned question: Is human identity the starting point or the ending point for liberty? If human identity is separated from human nature, it is an ending point. Chief Justice Taney's infamous identification of Dred Scott as a part of a group that he considered a subordinate and inferior class of beings is probably the most glaring example.[62] But at least for some, Justice Blackmun's identification of the unborn as constitutional nonpersons is not far behind.[63]

In 2005, it is self-evident that slavery is not consistent with the created human identity specified in the Declaration of Independence, Chief Justice Taney notwithstanding. There is no similar consensus with regard to abortion. So then, how are such competing claims of identity and freedom to be resolved? One answer is to say that when a claim of freedom is in disregard of created human nature or identity, it is specious and should be judicially disavowed. That is obviously a difficult answer for those properly mindful of a restrained judiciary. The answer cannot, however, be dismissed out of hand without sacrificing all fidelity to the founding aspect of the self-evident truth of our creation. After all, how many of us can comfortably read in the *Dred Scott* decision that it is not the province of a court to decide upon the justice or injustice of slavery, without sensing this to be a profound judicial default?[64] But of course, a very large part of the human identity is, as mentioned before, imperfection—not fully grasping the truth of even our own natures—and, in part, that is why such sensitive issues as abortion and marriage are contested today. There is a truth of identity in each case and the Court is frightfully wrong when it says that human beings define their own place and their own existence in the universe.[65] The Founders did not declare, "We

62. *See* Scott v. Sanford, 60 U.S. (19 How.) 393, 404–05 (1857).

63. *See* Roe v. Wade, 410 U.S. 113 (1973).

64. *Cf.* Town of Castle Rock v. Gonzales, 125 S. Ct. 2796 (2005) (refusing to find a positive substantive or procedural obligation to implement a court order against an abusive husband, who kidnapped and then murdered his three children in open defiance of the order, even as the mother of the children begged multiple times for the police to help or intervene).

65. *See* Planned Parenthood of Se. Pa. v. Casey, 505 U.S. 833, 851 (1992).

hold these random thoughts that, however we got here, we have the right to be left alone."

If the truth of human nature is contested, however, democracy suggests that any provisional answer must be given by the people in legislative assembly. "We the people" may very well get it wrong, but at least then the conversation will not be stopped dead in its constitutional tracks by a judicial imposition of a mistaken view of human identity that puts us at war with ourselves.

FIVE THESES ON IDENTITY POLITICS

RICHARD D. PARKER*

The kind of freedom I want to address is the most vital kind: political freedom. By that I mean the summoning and exertion of energy to engage one another on matters of collective government. What I have in mind specifically is democratic political freedom. By that I mean political freedom in a context shaped by three simple norms: political equality, popular sovereignty and, therefore, majority rule.

What should we make of identity politics as an exercise of democratic political freedom? Let me respond with five connected theses.

Number One. All politics is identity politics. Political activity is—and, at its best, is—animated by efforts to define and defend who I am, or we are, or you are, or hope to be, or hope to be seen to be.[1] By extension, it is motivated by our imagination of what is or ought to be mine or ours or yours. It is not *only* about self-government. Nor does it always involve much in the way of public debate. What structures it, often beneath the surface, is the always unfinished enterprise of self-construction and self-presentation.

The reason, first of all, is that politics[2] involves making comparisons and choices among—and commitments to—values and interests and groups and individuals (including choices not to choose among available choices). The choices and the commitments we make in politics are ones with which we mean to—or by which we cannot help but—identify ourselves.[3]

---

* Williams Professor of Law, Harvard Law School. Some of the assertions made here will be developed toward the end of an essay to be entitled *Constitutional Lawmaking at the Ballot Box: A Populist Critique of a Progressive Critique*.

1. This is a premise of RICHARD D. PARKER, HERE, THE PEOPLE RULE: A CONSTITUTIONAL POPULIST MANIFESTO (1994) (arguing that political activity is an effort to express and defend identity).

2. "Politics," as I understand it, may go on in contexts we call "private" as well as in ones we call "public."

3. All that may be true of shopping: We define ourselves to others and ourselves by what we buy, after all. Hence, there is more to it.

What is more, politics involves comparison, choice, and commitment under conditions of conflict. There are winners and losers. Crucially, over time, it is an open-ended conflict: The first ones now may later be (and often are) last. And, over time, political conflict is open in another respect. It is without permanent bounds or rules. The most unexpected issues may one day become salient political issues; allegiances and alliances shift; and, at some point, any mode of struggle, even war, may turn out to be politically decisive. This contingency of politics tends, in turn, to open up the enterprise of self-identification that animates it—keeping it on edge and, so, alive.

In democratic politics, moreover, the conflict is among putative equals. The norm of political equality not only destabilizes temporary victories. It also unsettles taken-for-granted hierarchies and, so, identities—and thus renews the spring of political energy.

In this way, identity politics and democratic political freedom are, in principle and often in practice, mutually supportive, each of them enabling the vitality of the other.

Number Two. However, it can also work the other way around. Identity politics can dampen or smother democratic political freedom. And democratic politics itself sometimes seems to sponsor this tendency, undermining itself by fostering a perversion of identify politics. The question is: What accounts for that? What sorts of identity politics, what aspects of identity politics, are pathological to democratic political freedom and where do they come from?

Number Three. Certain familiar answers[4] to the question are deeply misguided. They are as follows: That the pathology of identity politics has to do with its promotion of a self-regarding (rather than a public-regarding) political culture. Or of "stereotypes." Or that identity politics tends to portray and purvey differences and grievances (rather than similarities and bonds) among groups and individuals. Such diagnoses are wrong not simply because they flush out the baby with the bathwater, but because they seem not to recognize the baby—to understand the value of identity politics—in the first place.

---

4. These are answers that I assume are familiar to the multitude assembled at the Federalist Society conference where I made my remarks.

Of course, identity politics is self-regarding. It is, after all, about the construction and presentation of oneself. That matters to everyone. That is what accounts for the energy, the motivation, which identity politics can infuse into democratic politics.[5] It shouldn't take Adam Smith to remind us that self-concern is not necessarily antithetical to—that it can accompany and foster, even be indispensable to—promotion of the well-being of others.

Of course, identity politics involves construction and presentation of personae—caricatures, even "stereotypes" if you like—imagining and portraying particular individuals in terms of certain general traits. Needless to say, they distort "reality." And they are often prefabricated. But so what? That, as they say, is life. It is part and parcel of the life that identity politics contributes to democratic politics. In and of itself, nothing should be said against it.

That identity politics promotes difference and grievance is the most misguided complaint—and, in the end, the least troubling—of them all. Difference, of course, is vital to self-definition and self-assertion. And grievance is a fuel that motivates the exercise of political freedom. Neither is necessarily inconsistent with respect for one's antagonists.[6] Indeed, in a nation like ours, the respectful democratic antagonism they can foster is woven into the patriotism that binds us together.[7]

If there is a problem here, it is not in the ills highlighted by these diagnoses. It is, instead, in the diagnoses themselves. For it is not the supposed ills but the cure they imply—a neo-Victorian regimen of highfalutin "deliberation" or "civic virtue" or "public reason"—that would tend to stifle democratic political energy at its source.

Number Four. So, in what ways can identity politics become pathological? Put simply: It becomes pathological insofar as it is perverted to stifle, rather than to animate, the openness of democratic political conflict. The irony—a working-out of the old dialectic?—is that the very exigency of open conflict is what

---

5. *See* Richard D. Parker, *Taking Politics Personally*, 12 CARDOZO STUD. L. & LITERATURE 103, 107–08 (2000) (defending the role of the "personal" in politics).

6. *See* Richard D. Parker, *Democratic Honor: Liberal and Populist*, 39 HARV. C.R.-C.L. L. REV. 291, 292 (2004) (describing how personal political honor provides a source of agency in populist democratic politics).

7. *See* Richard D. Parker, *Homeland: An Essay on Patriotism*, 25 HARV. J.L. & PUB. POL'Y 407 (2002).

motivates tendencies to suppress it. Hence, any such tendencies of identity politics are likely to be, at once, resistant and vulnerable to revision. Let me point to three.

The first is a well known tendency to "essentialism." The academic fashion is to treat this as just an "idea"—the idea that individuals sharing a particular trait (such as race or gender) are essentially the same, share the same identity, in myriad other respects. As such, right-thinking (which is to say left-thinking) professors are against it. But it is not just an idea. It is an embedded practice. It is embedded in the practice of elites who seek to establish and maintain a position as "spokesmen" or "advocates for" one or another "affinity group." And it is embedded in the practice of some rank and file "members" of such groups who, by means of group opinion, seek to keep other "members" in line. No doubt, such practices can be rationalized as a necessary girding for battle in the wider political arena. But the effect is to truncate that conflict, erecting walls that keep some in, some out—calcifying the arteries of democratic politics.

The second pathological tendency of identity politics arises as healthy grievance against others turns to consuming blame and then to taken-for-granted prejudice. This is what academics describe and deplore as "demonization." Again, however, it is not simply a bad idea that can be calmly corrected or an emotion susceptible to control by adjusting the volume knob. It begins as what may seem a useful weapon in political struggle. Yet it ends as another, even more deadening, barrier to open political engagement. If "conservatives" or "religious fundamentalists" or "whites"—or, more likely "white men"—are dismissed as wrongdoers having no "identity" worth respecting,[8] the possibilities of political conflict (including compromise, its concomitant) tend to shrink. And politics turns into a sort of theater of self-righteous narcissism, segmented to gratify little clots of the like-minded and angry.

It is the third tendency that can be most devastating. An extension of the second, it arises when blame and prejudice against wrongdoers induce a conviction of impotence—a conviction which then can become central to the group's very iden-

8. A moment of illumination for me was the afternoon I walked by our open-doored faculty lounge and overheard a group of my colleagues earnestly attesting that, needless to say, it would be terrible if identity politics were ever to excite the consciousness of white males.

tity. Ultimately, its identity may be sucked into such black-hole characterizations as "the powerless" or "the disenfranchised" or "the subordinate." This is the "victimhood" syndrome. Like the other pathologies, it is much deplored. But like the others, once entrenched, it is difficult to dislodge.[9] The reason, again, is that it is so useful a weapon in democratic conflict. It is a weapon with which to denigrate the terms of the conflict and then to short-circuit the political fray. This, however, is not the worst of it. The worst is its potential for self-fulfilling prophesy, the tendency of "victim talk" to undermine self-responsibility, even to sever the nerve of political action, among voters whose supposed impotence has been made to define who they are.

Number Five. It follows that one of the more problematic identities in today's identity politics may be one of the most common. What I have in mind is the identification of a group not in terms of a trait such as race or ethnicity or sexual orientation, but as a "minority" group—and, worse, the identification of an individual "member" of such a group as "a minority." The banality of the label blinds us to its significance. Consider the ways it can exacerbate the three pathologies.

You might think there is no essentialist tendency to worry about in this case. How could so abstract and ambiguous a term excite *that* concern? If, however, it is as a practice, rather than as an idea, that essentialism is worrisome, then things look different. For, compared with more concrete identity categories, the greater abstraction and ambiguity of "minority"—even of "disadvantaged minority" or "discrete and insular minority"—creates even greater leeway for political manipulation and discipline. Authorities hand down and enforce definitions of "minorities" and their "members," definitions used, in effect, to keep some groups inside and others outside.[10] However useful this enterprise may be in the wider political struggle, it tends to accelerate the sclerosis of democratic engagement.

To that tendency, add the pathologies of demonization and victimhood. Shuffling the labels in order to construct and present certain individuals and groups, not as "Latinos" or "Span-

9. Indeed, it is difficult to dislodge even from the rhetoric of those who purport to deplore it.

10. *See, e.g.*, JOHN D. SKRENTNY, THE MINORITY RIGHTS REVOLUTION 142 (2002) ("[The Equal Employment Opportunity Commission's and Office of Federal Contract Compliance's] designations of official minorityhood shaped the trajectory of American politics and the structure of opportunity for the next several decades.").

ish-speaking" (or whatever), but as "minorities" tends to do a subtle—but deep—kind of damage. The term does not point to any actual trait of a person (ethnicity, gender, and so forth). Instead, it places a person in a group that is defined by its place in a larger structure. Not its place in a palpable social structure (say, class). But in a political structure—specifically, a democratic political structure, a system of majority rule. The implication of identification as "a minority" is that the group to which a person "belongs" is not, has not been, and will not be part of any majority coalition (which is, of course, all that shifting majorities can ever be) and, therefore, that she is bound to be ignored or abused (victimized) by that (demonized) majority. As to nearly all persons nowadays, this representation is absurd. How then to account for its currency? It is useful. It offers a leg up in democratic political conflict. But the cost—quite explicitly—is the trashing of majority rule, political equality, and popular sovereignty. When educational and economic elites, who have much to gain from such trashing, take up the cause and massage "minoritarian" identity politics into high-minded conventional wisdom, democratic political freedom—forever a threat to the privileged—is undermined.

What should a Federalist make of this? The Society's logo is a silhouette of James Madison. Didn't he famously (maybe too famously) fret about "majority factions"? Isn't he often (if unjustly) associated with Tocqueville's snooty hysteria about the "tyranny of the majority"? Haven't members of the Federalist Society often constructed and presented an identity as a suppressed minority in American law schools?

All true. But consider the practice of the Society. Though spurred by a healthy sense of grievance, it has not insulated itself. It has not sought to short-circuit engagement with, or set up enclaves protected from, dominant points of view. It has, instead, thrown itself into the fray. Indeed, it has invited the fray into its own events. Most Federalist get-togethers, as far as I know, involve debate with non-Federalists and anti-Federalists. What's more, they tend to include debate among Federalists, debate about the very identity of the group. Most everything is up for grabs. These are hallmarks of an identity politics that is open—not obsessed with victimhood or with demonization of its adversaries, not defensive. It invites the heterodox into its meetings because it has taken to heart an old political secret, the secret of the power of optimism: believing

that one day against present odds, it may prevail, it vastly enhances the odds that it may be right.

Hence, I offer a (not very serious) suggestion. It's time to give up the Madison silhouette. Madison's association with a depressive-defensive democratic politics really doesn't fit the Federalist Society. With what logo, then, might that one be replaced?

Here's a (still less serious) suggestion. What about a silhouette of Lenin? It would convey two messages. On one hand, a warning of the dangers of identity politics: When the construction and presentation of identities becomes a weapon of manipulation with which to stifle conflict—especially if taken to the extreme of claiming to create a "new man"—it becomes the deadly enemy of political freedom. On the other hand, Lenin's silhouette might offer a lesson in political optimism: His faction was in the minority. Hence, the name he chose for it was nothing less than "Bolshevik" . . . which, of course, means "the majority."

NEWS FOR THE LIBERTARIANS: THE MORAL TRADITION ALREADY CONTAINS THE LIBERTARIAN PREMISES

HADLEY ARKES*

This is a discussion long overdue in our circles, for the libertarian temptation threatens to do for us what the lure of moral relativism and postmodernism has done for liberalism in our own day, which is to render liberalism incoherent. When liberals enacted the Civil Rights Act of 1964, there was no tinge of doubt, at least, on the wrongness of racial discrimination, and liberals had no doubt about the claims of that law to render justice. But liberals in our own day invoke the claims of privacy, while they relentlessly legislate to deny claims of privacy in the management of private corporations, small private businesses, and private clubs. Liberals invoke equality at every turn, but they do not take seriously the notion that "all men are created equal" expresses a genuine moral truth. In other words, they invoke equality, but they can no longer explain what makes equality genuinely good, right, and just. The temptation has been to deny that there are grounds of moral judgment—mainly, I think, to fend off the temptation to cast moral judgments on others through legislation. The main motive, I suspect, among liberals is to insulate themselves from judgments cast on them in their sexual lives, while remaining free to legislate to restrict our private lives in all other respects.

That stance of liberals—moralism without morality and the drive to keep legislating—may account for the recoil of libertarians from people with the passion to legislate for others. But many libertarians have been drawn to the same fount of moral skepticism, and that moral relativism threatens our own side with the same incoherence that now afflicts the liberals. My claims here may be condensed in this way before I explain them: The libertarians do not seem to understand that their

* Edward N. Ney Professor of Jurisprudence and American Institutions, Amherst College, and the author most recently of NATURAL RIGHTS AND THE RIGHT TO CHOOSE (2002).

concerns are already incorporated in the classic understanding of the connection between the logic of morals and the logic of law. To that classic understanding, they run the risk of offering nothing more than a corrosive moral skepticism. For that traditional understanding already contains in its moral logic built-in burdens on people who would legislate casually, with moral pretensions but with no moral substance. The libertarians seem to have missed these anchoring points: that the claims of liberty flow only to those beings we call "moral agents"; that the libertarian argument is irreducibly a *moral* argument; and that when we understand the genuine moral requirements that attach to the making of laws, we realize that we have lifted the bars for legislation. To understand the moral ground that justifies the law is to make it *harder, not easier*, for people to legislate.

As Aristotle explained, at the beginning, the defining mark of the polis, or the polity, was the capacity to make decisions binding on everyone through the force of law. Human beings were fitted by nature to live in a polis, because law arose distinctly out of the nature of human beings: They were the only beings capable of giving and understanding reasons over matters of right and wrong.[1] As the American Founders understood, we may not rule beings of that kind in the way that we rule dogs and horses. Even today, in the age of animal liberation, we don't sign labor contracts with horses and dogs or seek the informed consent of our household pets before we authorize surgery on them. But we continue to think that beings who can give and understand reasons deserve to be ruled with a rendering of reasons, and politicians have an incentive to offer those reasons when they are seeking the consent of the governed in elections.

Abraham Lincoln said of those slaves who did not throw in with John Brown that, as ignorant and as uneducated as they were, they still had the wit to see that the schemes of this crazy white man would not conduce to their well-being.[2] Those blacks might have been *unlettered*, but they had the capacity to reason about the grounds of their own well-being, and for that reason they were not rightly annexed to the purposes of other men without their consent. Claims of liberty flow distinctly to

1. *See* HADLEY ARKES, FIRST THINGS 11–30 (1986).

2. *See* Abraham Lincoln, Address at Cooper Institute, New York City (Feb. 27, 1860), *in* 3 THE COLLECTED WORKS OF ABRAHAM LINCOLN 522, 541 (Roy P. Basler ed., 1953).

these beings who can reason about right and wrong. There is a corollary: Those beings who understand the difference between right and wrong may also grasp—as Aquinas and Lincoln recognized—that one cannot coherently claim a "right to do a wrong." And so they may understand the things that *they may never claim a right to do, even in the name of their own liberty*. William Blackstone famously said that we give up the liberty to do mischief, the liberty we enjoyed in the state of nature, and we exchange it for a more limited set of liberties and rights under civil society.[3] Our James Wilson responded, "Is it a part of natural liberty to do mischief to any one?"[4] Even in the state of nature we had no right to kill without justification, to murder, or to rape. And therefore the laws that restrained us from murdering and raping did not deprive us of anything that we ever had a rightful liberty to do. The upshot of this understanding is that, contrary to Blackstone, as Wilson and Alexander Hamilton explained, we did not surrender our natural rights when we entered civil society.[5] As Hamilton said in *The Federalist* No. 84, "Here, in strictness [that is, under the new Constitution], the people surrender nothing."[6] They did not enter the Constitution under the condition of surrendering or reducing their natural rights (or, we might say, their natural liberties), and the purpose of the Constitution was not to create new rights. It was, as Wilson said, to secure and enlarge rights that we already had, by nature.[7]

John Stuart Mill said that we stop using the language of "like" and "dislike" and start using the language of right and wrong when we think that people should be punished for what they are doing.[8] In that rough way, Mill caught the difference between a statement of merely personal taste and something that stands in the class of a moral proposition. In moral judgments, we move beyond personal taste and private choice, and we begin to speak about the things that are right or wrong, just

3. 1 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 121–22 (Univ. of Chi. Press 1979) (1765).

4. 2 JAMES WILSON, *Of the Natural Rights of Individuals*, *in* THE WORKS OF JAMES WILSON 585, 587 (Robert Green McCloskey ed., 1967).

5. *See* THE FEDERALIST NO. 84, at 534 (Alexander Hamilton) (Benjamin Fletcher Wright ed., Harvard Univ. Press 1961) (1788); WILSON, *supra* note 4, at 586–87.

6. THE FEDERALIST NO. 84, *supra* note 5, at 534.

7. The fuller argument here, part of the original argument on the Bill of Rights, can be found in HADLEY ARKES, BEYOND THE CONSTITUTION 58–80 (1990).

8. *See* JOHN STUART MILL, UTILITARIANISM 60–61 (Oskar Piest ed., Liberal Arts Press 1957) (1861).

or unjust, for others as well as for ourselves. To paraphrase Aquinas, by "wrong" we mean that which no one ought to do, that anyone may rightly be restrained from doing, and if we are incapable of restraining any person in committing a wrong, then we are justified in punishing him.[9] With that logic, Lincoln smoked out Stephen Douglas when Douglas said that he was neutral on slavery, that he had reached no moral judgment on the matter, and so people ought to be free to vote slavery up or down in the territories as it suited their interests. Lincoln pointed out that Douglas had indeed reached a moral judgment: To say that slavery was a legitimate choice was to say that it fell into a class of things "not wrong."[10] We would not say, "it is wrong for parents to torture their infants," and then go on to say, "therefore, let them make this judgment on their own, to torture or not, as it suits their interests or their conscience." Nor would we say, "it is wrong to torture infants, therefore, let us give parents tax incentives, or a DVD player, in order to induce them to stop." Hence, the traditional understanding of the connection between the logic of morals and the logic of law: When we come to the recognition that any act—say, the torturing of infants—stands in the class of a wrong, we lay the groundwork for forbidding that act to people generally or universally; which is to say, we may forbid it with the force of law.

This does not strictly mean that we are obliged to make any and every act illegal once we can judge it as "wrong." Prudence may well enter with a sense of the wrongs that it may be impolitic to reach without generating more evils, and with the sense that we are administering to a population of human beings, not angels. With the caution of Aquinas, we know that we

---

9. *See* 2 THOMAS AQUINAS, SUMMA THEOLOGICA 1431–34 (Fathers of the English Dominican Province trans., Benzinger Bros. 1947) (1485).

10. See Lincoln's classic rejoinder to Douglas during the debate at Quincy, Illinois on October 13, 1858:

> [W]hen Judge Douglas says he "don't care whether slavery is voted up or voted down," . . . he can thus argue logically if he don't see anything wrong in it . . . . He cannot say that he would as soon see a wrong voted up as voted down. When Judge Douglas says that whoever, or whatever community, wants slaves, they have a right to have them, he is perfectly logical if there is nothing wrong in the institution; but if you admit that it is wrong, he cannot logically say that anybody has a right to do wrong.

Abraham Lincoln, Sixth Debate with Stephen A. Douglas, at Quincy, Illinois (Oct. 13, 1858), *in* 3 THE COLLECTED WORKS OF ABRAHAM LINCOLN, *supra* note 2, at 245, 256–57.

cannot hope to eradicate all vice; we would seek, rather, to keep evils constrained to certain tolerable limits. But when we say that there is a moral groundwork for the law, we mean strictly that, before we remove private choice and restrict personal freedom, the law carries the burden of showing that there is something truly wrong in principle—with the sense of a "wrong" that meets the more stringent tests that attach to a moral proposition. That standard cannot be satisfied merely by the claim that this policy reflects a strong belief shared by most people who are voting—for example, a policy to bar young lawyers from working overly long hours in offices hazardous to their health. That is not good enough to explain why it is wrong for any particular young lawyer to work additional hours if he thinks it necessary for his case or his career. As Rufus Peckham warned in *Lochner v. New York*, we could as aptly reach that kind of conclusion on the same loose reasoning that justifies restrictions on the hours of bakers who wish to work overtime.[11]

But quite apart from the beliefs or passions of the majority, it does not meet a moral standard to produce a policy based merely on *generalizations*—even those generalizations that are quite unmistakably true and persuasive. I would submit to you that we rarely encounter findings, based on aggregate data, as compelling as the findings assembled years ago by my friend George Gilder when he collected the data describing the plight of that hapless character known as the *single male*. Bachelors were twenty-two times more likely than married men to be committed to hospitals for mental disease (ten times more likely to suffer chronic diseases of all kinds); single men have nearly double the mortality rate of married men and three times the mortality rate of single women; divorced men are approximately three times more likely than divorced women to commit suicide or die by murder, and are six times more likely to die of heart disease.[12]

These data are as clear in their import as anything we have seen emerging from such aggregate data: These single men are massively at risk! Their deliverance from this condition could have the most profound effect, not only in improving public health, but actually rescuing lives. Yet, we take it that nothing

11. *See* Lochner v. New York, 198 U.S. 45, 59–60 (1905).

12. *See* George Gilder, *In Defense of Monogamy*, COMMENTARY, Nov. 1974, at 31, 32.

in this set of compelling data would furnish a ground for legislating on marriage by contravening the freedom of men *and women* by assigning these men to brides. In the same way, the old Coleman report told us that if we kept the population of black youngsters in any school below twenty-five percent, we could have the effect of raising slightly the median reading scores for black students.[13] That is to say, there was no assurance that any particular black student would benefit from this scheme of assignments based on race. But we seem to understand that, if we cannot say in principle why this arrangement is good or just of necessity for any student, we do not have the grounds for contravening the freedom of any student (or family) by assigning these students to schools explicitly on the basis of race to meet a racial ratio considered optimal for the schools. In other words, when we take seriously the moral ground of the law and the moral requirements that must be met before we are justified in making laws, then *the result is not that we legislate more, but legislate less*. For the bar has been raised notably higher: The standard is more demanding, and that standard rules out the more cavalier grounds for legislating that people have been all too ready to accept in the past.

That standard of a moral justification would not be satisfied by empirical predictions merely contingent in nature. "If Saks Fifth Avenue moves into the neighborhood, the neighborhood will prosper." That may be true, but even if it were true most of the time, it would not be true of necessity. Whether it is true or not, or good in its effects, will be contingent on many other things, probabilistic in nature, and therefore problematic. There is a striking difference between that kind of a proposition and something closer to the proposition Thomas Reid took as one of the first principles in legal and moral reasoning, which I render in this way: We cannot hold people blameworthy and responsible for acts they were powerless to affect.[14] There are no con-

---

13. I draw these findings from the project at Harvard to reexamine the data in the Coleman Report. *See* David Armor, *School and Family Effects on Black and White Achievement: A Reexamination of the USOE Data*, *in* ON EQUALITY OF EDUCATIONAL OPPORTUNITY 168, 197 (Frederick Mosteller & Daniel P. Moynihan eds., 1972). For the original Coleman Report, see JAMES S. COLEMAN ET AL., U.S. DEP'T OF HEALTH, EDUC., AND WELFARE, EQUALITY OF EDUCATIONAL OPPORTUNITY 330–33 (1966).

14. Reid rendered the matter in this way: "What is done from unavoidable necessity . . . cannot be the object either of blame or of moral approbation." THOMAS REID, ESSAYS ON THE ACTIVE POWERS OF MAN 370 (Garland Publ'g 1977) (1788).

tingencies under which that proposition will fail to be true, and its truth will be entirely unaffected by the question of whether its recognition in any case will bring good or bad results. And yet that proposition may explain quite a bit of the ground of our law in things like the insanity defense or even, in part, the wrong of racial discrimination.[15]

In striking contrast, so many of our judgments in the area of civil rights have been put forth as contingent propositions of that kind. For example, if blacks could not attend the same law school with whites in Missouri—if they had to take a voucher to attend school in another state—then, the argument ran, they could not make the connections useful to their careers in the law.[16] But it does not follow, as a matter of necessity, that any black student in the school will have conversations, form friendships, and make connections that lead to a rising career. Cecil Partee, the late black ward committeeman in Chicago, was barred from attending the law school at the University of Arkansas in 1938. With a voucher, he had to choose instead between the law schools at Northwestern and the University of Chicago. As he later said, "I laughed all the way to Chicago."[17] Cecil Partee had not suffered a material harm as a result of being turned away on the basis of race. But he had been wronged; he had been treated according to the maxims of an unjust principle. If we considered again the cases of racial segregation in the schools, had it been wrong *in principle* to separate children on the basis of race? Or had the wrong been contingent on its effects in impairing the will of black children to learn, in depressing their performance in the schools? If the students had been separated on the basis of race, and the performance of the black students had risen (as it did at the Dunbar High School in Washington),[18] would the segregation have ceased to be wrong?

---

15. *See generally* ARKES, *supra* note 1, at 85–99.

16. *See* Missouri *ex rel.* Gaines v. Canada, 305 U.S. 337, 349 (1938) (describing petitioner's insistence that "for one intending to practice in Missouri there are special advantages in attending a law school there, both in relation to the opportunities for the particular study of Missouri law and for the observation of the local courts, and also in view of the prestige of the Missouri law school among the citizens of the State, his prospective clients") (internal citations omitted).

17. MILTON L. RAKOVE, WE DON'T WANT NOBODY NOBODY SENT: AN ORAL HISTORY OF THE DALEY YEARS 155 (1979).

18. *See* Thomas Sowell, *Black Excellence: The Case of Dunbar High School*, PUB. INT., Spring 1974, at 3.

But if there is a principle engaged in these cases of racial discrimination, that principle covers all of the instances in which it may be manifested. If we know the principle that determines the speed by which the ball travels down the inclined plane, as it alters with the angle of inclination, we no longer have to ask: Was it a wooden ball or a metal one, a blue one or a yellow one? And yet, as Philip Kurland once asked, how did we get from *Brown v. Board of Education*[19] to *Palmer v. Thompson*?[20] If the wrong lay in impairing the performance of black children in the schools, did we claim that the barring of children from swimming pools affected their capacity to read and learn? But if there was a wrong in principle in segregation, then that principle applies to every instance in which it could be manifested: The racial discrimination could take place on tennis courts or in access to Xerox machines, and there would not be an obligation to articulate a constitutional "right to play tennis" or a "right to use the Xerox machine." And that is what marked the obtuseness of Chief Justice Warren's opening line in *Loving v. Virginia* that the Court had never encountered a case of this kind before: a restriction of marriage along lines of race.[21] Never encountered before? Never encountered the problem of racial discrimination? Or never articulated the principle engaged in the case, the principle that would deliver the Justices from the surprise of new cases, and rescue them from the position of that hapless character who discovers the system of positive integers—one at a time.

Immanuel Kant warned, in one of his more luminous moments, that even a unanimity of feeling cannot provide the surrogate, or replacement, for a moral proposition.[22] For if that conviction rested solely in feeling, it was contingent upon the possibility of that feeling passing away. It may not be the same tomorrow as today. And so, even if we were unanimous in our passion for Coca-Cola or frozen yogurt, we would still not have the ground for making the cola or the yogurt compulsory. My friend Jim Wilson says that we can show people sonograms on babies in the womb at eight to ten weeks, and most of them

19. 347 U.S. 483 (1954).

20. 403 U.S. 217 (1971).

21. 388 U.S. 1, 2 (1967).

22. *See* IMMANUEL KANT, FOUNDATIONS OF THE METAPHYSIC OF MORALS (1785), *reprinted in* FOUNDATIONS OF THE METAPHYSICS OF MORALS AND WHAT IS ENLIGHTENMENT? 3, 60 (Lewis White Beck trans., Macmillan Publ'g Co., 2d ed., rev. 1990).

will say, "That looks like a baby to me."[23] But if the decision turned solely on the question of "what that looks like to me," then the law would have nothing to say to that small fraction of the population who look at the sonogram and say, "Well, it doesn't look like a baby *to me*." For the law, for the question in principle, it does not matter what the child in the womb "looks like" or how people feel about it. The question is "what it is," or the ground of principle upon which we would claim that the offspring of homo sapiens in the womb is anything less than human at any of its stages.[24]

The discipline for the law, and for jurists, is to get clear on the grounds of principle, or the properties of a moral argument, that can alone justify a law: a binding rule that forecloses personal freedom and private choice. A law that is tested in that demanding way—with a stringent test of the moral ground of the law—is a law that will legislate less, and be far more protective of human liberty. Part of my claim here is that a law that cannot give that moral account—that account of what strictly "justifies" the law—is condemned to incoherence. A jurisprudence that seeks to avoid the moral question will appear to engage in charades, where the rationale brought forward for the law has no connection with the ground upon which people of ordinary sensibility understand the wrong of the case. And so, for example, employing the rationale of the Commerce Clause, if black people expect discrimination in public accommodations, they will travel less, and there will be fewer orders for meat and linens and tableware. What was the wrong then in racial discrimination in this construal? It depressed the interstate flow of meat and other goods. As a friend of mine remarked, that problem would be relieved by having the racists in the country eat more meat. Consider the same argument, reworked to fit the problem of abortion: that private abortion may be yours alone, as the wheat on Roscoe Filburn's farm was his alone,[25] but when abortion is summed up in a volume of 1.3 million every year and accumulated over the years, it has the most depressing effect on the interstate sales of diapers, baby

23. *See* James Q. Wilson, *On Abortion*, COMMENTARY, Jan. 1994, at 21, 26.

24. *See* Hadley Arkes, *Abortion Facts and Feelings*, FIRST THINGS, Apr. 1994, at 34, 36–37; James Q. Wilson & Hadley Arkes, *Abortion Facts and Feelings II: An Exchange*, FIRST THINGS, May 1994, at 39, 42.

25. *See* Wickard v. Filburn, 317 U.S. 111 (1942).

food, bassinettes, to say nothing of college educations and cohorts of workers to fund the system of Social Security.

One of the enduring sources of confusion among people who banter over moral judgments is that we cannot cast absolute moral judgments in terms of any empirical thing or activity. As Kant recognized, anything we can name, even the most prosaic things, can be part of a means-end chain leading to a wrongful end.[26] We can use the skill of driving to drive for the Mafia or to drive an ambulance. We cannot even say that it is wrong under all circumstances to kill your mother, for she may be trying to kill you, and we do have examples, after all, of mothers willing to kill or order the deaths of their offspring. We can make our way to the ground of moral judgment in the law only by making our way back to those logical maxims, with a necessary force, that underlie our acts.[27] At times those maxims have been woven into the aphorisms of the law, taught by our great jurists: The presence of a wrong justifies a remedy.[28] We should visit punishment only on people guilty of wrongs, and we should respect the difference between the guilty and the innocent by insisting on making those discriminations in the most demanding way—by insisting on evidence, tested by the canons of reason, rather than having defendants outrun mobs in the street. Justice Story reminded us that what we may not do directly, we may not do indirectly.[29] If it is wrong to kill Jones, it is wrong to hire someone to kill Jones. If the federal government cannot reach a certain subject, it may not use the tax system as a surrogate or a disguise for legislation.[30]

We do not know the moral significance that goes along with driving cars or using computers unless we know the end to which the activity is directed. For the same reason, we cannot pronounce a judgment on the categorical innocence of anything that takes place in a domain of privacy. A murder in a bedroom is as much a murder as a murder in the public streets, and the law may reach the most private places—even people storing

---

26. *See* KANT, *supra* note 22, at 9.

27. For an elaboration of this argument, see ARKES, *supra* note 1, at 87–88.

28. *See* 3 BLACKSTONE, *supra* note 3, at *23.

29. *See* Charles River Bridge v. Warren Bridge, 36 U.S. (11 Pet.) 420, 617 (1837) (Story, J., dissenting) ("Can the legislature have power to do that indirectly, which it cannot do directly?"); La Nereyda, 21 U.S. (8 Wheat.) 108, 124 (1823) (Story, J.) ("[T]he Court will not permit that to be done indirectly, which could not be done directly.").

30. *See* United States v. Butler, 297 U.S. 1, 68–70 (1936) (Roberts, J.).

heroin in condoms placed in body cavities. That understanding, long established, explains what is so mindless in those glib formulas that suggest that anything taking place in a private setting is insulated from moral judgment, or that the law cannot cast moral judgments on "sexual orientations." That term, grown so familiar, is so abstract that it could cover bestiality or sadomasochism. There is the strangest overlooking of the fact that even gay activists find certain "orientations" illegitimate. It has been a point of contention for years as to whether the Gay Pride Parade in New York will include the North American Man/Boy Love Association (NAMBLA). Even the activists suggest that there is something not legitimate about this "sexual orientation." And even the Board of Education in New York City was willing to fire a teacher at the Bronx High School of Science because he was an officer of NAMBLA, and quite clearly committed to its principles (including, quite notably, a willingness to deny the legitimacy of establishing an "age of consent" for sexual engagements).[31] Ordinarily we allow parents to give their consent for certain involvements of their minor children, but even the libertarians do not seem inclined to permit parents to give their assent to a sexual relation between a teenage son and a male teacher.

If we are even faintly rigorous, we would have to recognize that there is no ground for legislation that bars, in a sweeping way, all discriminations based on "sexual orientation." If the partisans of gay rights had to stop using that term, they could be pressed into saying more precisely what they have in mind. And yet it would be hard for them to do that, for it would become quite as implausible to say that we do not cast moral judgments on people in the way they achieve orgasms. There is the matter, obviously, of rape, but also things that are not so obvious, for our legal judgments are not all cast in the criminal law. Statutes on adultery are almost never enforced, and yet women have lost custody of their children because of a pattern of promiscuity, suggesting a want of fitness as a parent. In *Lawrence v. Texas*, Justice Kennedy suggested with sweeping language that the law must recede from casting moral judgments on certain intimate, sexual practices taking place in private.[32] Just about every conservative I know would have held back the law, or repealed that criminal law on sodomy in Texas. And yet

31. *See* Melzer v. Bd. of Educ., 336 F.3d 185, 189, 191–92 (2d Cir. 2003).
32. *See* 539 U.S. 558 *passim* (2003).

it marks a flight from sobriety—or any faintly rigorous juridical sense—to say that the law may never cast an adverse moral judgment on sexual practices taking place in private settings. If there are men who seek to be adoptive fathers, it could make the most notable difference, in the weighing of their claims, if they happen to be, like Peter Melzer in New York, committed members of NAMBLA, or people who might be committed quite openly to sadomasochistic sex. Even judges who shy away from moral language may find reservoirs of judgment as they discover plausible grounds to believe that a man committed to sadomasochism, containing in his house whips and chains and fearsome devices, would not provide the most wholesome household for the nurturing of a child.

To take a line from Henry James, Justice Kennedy is the victim of perplexities from which a single spark of direct perception might have saved him.[33] For a rather seasoned man of the law, he has shown a strange willingness to be seduced by airy clichés and to pronounce moral judgments utterly detached from any grounding in moral substance. The corrective is not to ban, from our reasoning and our judgments, a distinctly moral perspective, for in that path lies incoherence. The main remedy is to get clear again about the moral ground of the law, and to remind ourselves of how rigorous in fact the test must be before we pronounce upon the things that are categorically, in principle, right and wrong—and that deserve to furnish the ground of the law. To rephrase a line of Woody Allen's, there is no question there is a moral world out there; the only question is, how far is it from midtown and how late does it stay open?[34]

---

33. *See* HENRY JAMES, *The Siege of London*, *in* THE NOVELS AND TALES OF HENRY JAMES 143, 190 (Charles Scribner's Sons 1908) (1882).

34. *See* WOODY ALLEN, *Examining Psychic Phenomena*, *in* WITHOUT FEATHERS 7, 7 (1975).

SAFETY *AND* FREEDOM: COMMON CONCERNS FOR CONSERVATIVES, LIBERTARIANS, AND CIVIL LIBERTARIANS

NADINE STROSSEN*

I am happy to address this Symposium, and to be a regular speaker at many Federalist Society forums. I always start such presentations by reminding all Federalist Society audiences of your organization's libertarian founding principles, which could come straight from the ACLU's Policy Guide, and which are directly relevant to this panel's topic. Your mission statement's opening words proclaim: "The Federalist Society . . . is founded on the principle[] that the state exists to preserve freedom . . . . The Society seeks . . . reordering priorities within the legal system to place a premium on individual liberty . . . ."[1] Also key to this panel's topic, your mission statement declares "that the separation of governmental powers is central to our Constitution . . . ."[2]

In 1994, I was on a Federalist Society panel with one of your founding gurus, Irving Kristol. As usual, I recited these libertarian tenets of your group, and it sent him into a state of shock. Our discussion was published, so let me read you his exact response:

> I am shocked to discover that the Federalist Society seems to have said somewhere that the State exists to preserve freedom. The Federalist Society should call a meeting immedi-

---

* Professor of Law, New York Law School; President, American Civil Liberties Union; J.D., 1975, A.B., 1972, Harvard University. The author thanks Joshua Davey and other editors of the *Harvard Journal of Law & Public Policy* for having undertaken most of the work in preparing footnotes for this published version of remarks delivered at the Federalist Society National Student Symposium, February 26, 2005. For assistance with several footnotes, the author thanks her ACLU Civil Liberties Fellow, Joseph Griffin (NYLS '05), and her Research Assistant, Jennifer Rogers (NYLS '08).

1. The Federalist Society for Law and Public Policy Studies, Our Purpose, http://www.fed-soc.org/ourpurpose.htm (last visited Nov. 20, 2005).

2. *Id.*

> ately and change that . . . . You say that and you get yourself in the kind of trap that Ms. Strossen has now sprung.[3]

I periodically re-read your brochures and website with trepidation, worrying that you might have heeded Kristol's advice. So far, though, you have not done that. Therefore, you are again "trapped" when it comes to the topic of the present panel. And remember, it is not me who said that, but Irving Kristol!

Even putting aside Kristol's consternation over the libertarian tenets of the Federalist Society, which describes itself as "a group of conservatives and libertarians,"[4] there would still be much common ground between our two organizations. Many conservatives, as well as libertarians, have agreed with the ACLU on this panel's topic: that too many post-9-11 measures have unjustifiably sacrificed our freedom without sufficiently advancing our security. Therefore, the extraordinarily diverse partners in our Safe *and* Free campaign have included individuals who are allied with both the libertarian and the conservative wings of the Federalist Society.[5]

As the Federalist Society's Twenty-Fourth Annual Student Symposium brochure states, members of this panel were invited to discuss our "legal and policy prescriptions for . . . balancing security concerns . . . with the protection of civil liberties."[6] Federalist Society leaders regularly profess fidelity to the Constitution's text.[7] I should think, then, that this text would provide the governing standards for Federalists in assessing how to "balance security concerns . . . with the protection of civil liberties." The Constitution's text has certainly guided civil libertarians' answers to these questions.

Concerning the scope of individual rights in national emergencies, the Constitution contains only one express limitation, on only one right, in only two specified types of national emer-

---

3. Irving Kristol, *Educating the Urban Poor: The (Only) Legitimate Function of the Public Schools*, 1 MICH. L. & POL'Y REV. 325, 325 (1996).

4. The Federalist Society for Law and Public Policy Studies, *supra* note 1.

5. *See* ACLU.org, Conservative Voices Against the USA-PATRIOT Act, Apr. 15, 2004, http://www.aclu.org/SafeandFree/SafeandFree.cfm?ID=12632&c=206.

6. Schedule, Harvard Federalist Society, Twenty-Fourth Annual Federalist Society Student Symposium (Feb. 25–26, 2005) (on file with the Harvard Journal of Law & Public Policy).

7. *See* Charles Krauthammer, *Return to Justice as Founders Saw It*, CHI. TRIB., June 13, 2005, § 1, at 17.

gencies: its provision empowering Congress to suspend the writ of habeas corpus.[8] This "Suspension Clause" imposes a heavy burden of justification before Congress may suspend the writ, limiting the suspension power to "Cases of Rebellion or Invasion," and even in such cases authorizing suspension only "when . . . the public Safety may *require* it."[9] Beyond the limited circumstances in which Congress may suspend the writ of habeas corpus, the Constitution provides no textual warrant for any further limits on rights just because the national security may be in peril. In the post-9-11 context, this crucial point was stressed by the staunch Federalist Society stalwart Justice Antonin Scalia in his *Hamdi* dissent.[10] This key point also has been emphasized by influential judicial opinions arising from various national emergencies throughout U.S. history, from the Civil War[11] to the Great Depression[12] to the Korean War.[13]

---

8. U.S. CONST. art. I, § 9, cl. 2.

9. *Id.* (emphasis added).

10. Hamdi v. Rumsfeld, 542 U.S. 507, 562 (2004) (Scalia, J., dissenting) ("The Suspension Clause was by design a safety valve, the Constitution's only 'express provision for exercise of extraordinary authority because of a crisis . . . .'" (quoting Youngstown Sheet & Tube Co. v. Sawyer, 343 U.S. 579, 650 (1952) (Jackson, J., concurring))).

11. *See Ex parte* Milligan, 71 U.S. (4 Wall.) 2, 120–21 (1866):

> The Constitution of the United States is a law for rulers and people, equally in war and in peace, and covers with the shield of its protection all classes of men, at all times, and under all circumstances. No doctrine, involving more pernicious consequences, was ever invented by the wit of men than that any of its provisions can be suspended during any of the great exigencies of the government.

12. *See* Home Bldg. & Loan Ass'n v. Blaisdell, 290 U.S. 398, 425 (1934):

> [W]e must consider the relation of emergency to constitutional power . . . .
>
> Emergency does not create power. Emergency does not increase granted power or remove or diminish the restrictions imposed upon power granted or reserved. The Constitution was adopted in a period of grave emergency. Its grants of power to the Federal Government and its limitations of the power of the States were determined in the light of emergency and they are not altered by emergency.

13. *See Youngstown*, 343 U.S. at 649–50 (Jackson, J., concurring):

> The appeal . . . that we declare the existence of inherent powers *ex necessitate* to meet an emergency asks us to do what many think would be wise, although it is something the forefathers omitted. They knew what emergencies were, knew the pressures they engender for authoritative action, knew, too, how they afford a ready pretext for usurpation. We may also suspect that they suspected that emergency powers would tend to kindle emergencies. Aside from suspension of the privilege of the writ of habeas corpus in time of rebellion or invasion, when the public safety may require it, they made no express provision for exercise of extraordinary authority because of a crisis. I do not think we rightfully may so amend their work . . . .

In short, apart from the writ of habeas corpus, the Constitution affords the same strong protection to individual rights during national crises as during any other time. For example, the government's power to invade individual privacy and freedom through any type of search and seizure, including any surveillance, should be subject to the Fourth Amendment's key requirements of individualized suspicion and judicial review.[14] Likewise, the government's post-9-11 restrictions on other fundamental rights—including its many restrictions on First Amendment freedoms[15]—should be subject to strict judicial scrutiny. This means that the government bears the burden of demonstrating that the restriction is narrowly tailored to promote a purpose of compelling importance.

The government easily can satisfy the "compelling interest" aspect of strict scrutiny by showing that the purpose of any rights-restricting measure is to protect national security. It is, however, harder for the government to satisfy the appropriately heavy burden of demonstrating that the measure is sufficiently narrowly tailored: specifically, that the measure is necessary and that it is the least restrictive alternative. If the government could promote its national security concerns through alternative means that are less restrictive of fundamental rights, then it must do so.[16] This rights-protective approach to civil liberties post-9-11 is consistent not only with the Constitution's text, but also with the Federalist Society's own tenets of maximizing individual freedom and minimizing government power.

In the ACLU's post-9-11 Safe *and* Free campaign, the ACLU and its ideologically diverse partners have analyzed each touted security measure to ensure that it really does maximize security, with the minimal possible cost to liberty. Not only is this the very analysis the Supreme Court uses as a matter of constitutional law—the "strict judicial scrutiny" approach—but

---

14. *See* U.S. CONST. amend. IV.

15. *See, e.g.*, ACLU.org, Freedom Under Fire: Dissent in Post-9/11 America (Dec. 8, 2003), http://www.aclu.org/safeandfree/safeandFree.cfm?ID=12581&c=207. *See also* Nadine Strossen, *Free Speech in Wartime*, 36 RUTGERS-CAM. L.J. (forthcoming 2005).

16. *See* United States v. Playboy Entm't Group, Inc., 529 U.S. 803, 813 (2000) (articulating the least restrictive alternative test). *See generally* ERWIN CHEMERINSKY, CONSTITUTIONAL LAW: PRINCIPLES AND POLICIES 764–68 (2d ed. 2002) (describing the various levels of scrutiny applied by courts to restrictions on fundamental rights).

it also squares with common sense. After all, why should we give up our freedom if we do not gain security in return, or if we could gain as much security without giving up as much freedom?

Applying this sensible constitutional test to the myriad of post-9-11 policies that officials have implemented or proposed, the ACLU and our allies have concluded that many pass scrutiny. For example, of the approximately 160 provisions in the USA-PATRIOT Act,[17] we have criticized only about twenty.[18] Moreover, even concerning those provisions, we have not advocated repeal, but rather, reform. We advocate reforms that would preserve the core of the powers the government asserts it needs to protect our lives, but we also seek to ensure that such powers are subject to judicial review and other constraints to bring them back in line with constitutional values.[19]

This constrained and constructive criticism by the ACLU and its allies hardly warrants the charge of "hysteria" that was leveled by John Ashcroft while he was Attorney General.[20] If anything, the extreme position is the one that the Bush administration has asserted, refusing to consider any reform to even a single provision in the entire 350-page Act. Rather, the Administration's adamant, unyielding position has been that the whole law must remain in effect without a single amendment, even to the point of not honoring the sunset provisions scheduled to go into effect at the end of 2005.[21]

---

17. Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act of 2001, Pub. L. No. 107-56, 115 Stat. 272 (2001) [hereinafter Patriot Act].

18. *See, e.g.*, ACLU.org, The Sun Also Sets: Understanding the Patriot Act "Sunsets," http://action.aclu.org/reformthepatriotact/sunsets.html (last visited Nov. 20, 2005). The ACLU has urged Congress not to reenact the "sunsetted" provisions of the Patriot Act, which are set to expire in December 2005. More generally, the ACLU has advocated that "Congress should use the debate over the sunsets to highlight these [sunset] provisions in particular, but should also take the opportunity to deliberate more broadly on the state of our freedoms in the 'war on terrorism.'" *Id*.

19. *See generally* ACLU.org, Keep America Safe *and* Free Program, http://www.aclu.org/safeandfree/index.html (last visited Nov. 20, 2005).

20. *See* CBS News.com, Ashcroft Slams Critics' "Hysteria" (Sept. 16, 2003), http://www.cbsnews.com/stories/2003/09/18/national/main573894.shtml.

21. *See, e.g.*, President George W. Bush, Address to the FBI Academy (July 11, 2005), *available at* http://www.whitehouse.gov/news/releases/2005/07/20050711-1.html ("I call on Congress to reauthorize the sixteen critical provisions of this act that are scheduled to expire at the end of this year. The terrorist threats will not

This administration has made other extreme assertions of executive authority since 9-11, such as claiming unilateral power to label any U.S. citizen an "enemy combatant."[22] The Supreme Court rejected this claim in the *Hamdi* case.[23] The Justices who issued the *Hamdi* ruling are generally supportive of presidential power,[24] so it is especially significant that they so resoundingly rejected the administration's claim. In the widely quoted words of Justice Sandra Day O'Connor, hardly a radical civil libertarian, "a state of war is not a blank check for the President when it comes to the rights of the Nation's citizens."[25]

Since the *Hamdi* decision, the ACLU has convinced two other judges to strike down post-9-11 measures because they unjustifiably violated individual rights. In the first court case to rule on any of the Patriot Act's sweeping surveillance powers, a federal judge agreed with us that one such power violated fundamental constitutional rights.[26] Specifically, the court struck down a 1986 statute, as dramatically expanded by the Patriot Act, which empowers the government to issue "National Security Letters," demanding that Internet communications carriers, as well as other businesses and organizations, secretly turn over customers' names and addresses, length of service and billing records, and information concerning the customers' emails and websurfing without any basis for suspecting the customers of any wrongdoing, without any judicial review, and without any notice to the customers that such information has been sought or seized.[27]

More recently, another federal judge agreed with the ACLU[28] that the First Amendment requires a lifting of the automatic

---

expire at the end of this year, and neither should the protections of the Patriot Act.").

22. *See* Hamdi v. Rumsfeld, 542 U.S. 507 (2004).

23. *See id.* at 508.

24. *See, e.g.*, Bernard Schwartz, *A Decade of Administrative Law: 1987–1996*, 32 TULSA L. J. 493, 552 (1997) (noting the Rehnquist Court's deference to presidential power); Harold Hongju Koh, *Reflections on Refoulement and Haitian Centers Council*, 35 HARV. INT'L L.J. 1, 2 (1994) (claiming that the Rehnquist Court has shown a disturbing pattern of reflexive deference to presidential power in foreign affairs).

25. *Hamdi*, 542 U.S. at 536.

26. *See* Doe v. Ashcroft, 334 F. Supp. 2d 471 (S.D.N.Y. 2004), *argued sub nom.* Gonzales v. Doe, No. 05-0570 (2d Cir. Nov. 2, 2005).

27. *See* Patriot Act, *supra* note 17, at 365–66 (codified at 18 U.S.C. § 2709(b)).

28. *See* Doe v. Gonzales, 386 F. Supp. 2d 66 (D. Conn. 2005), *application for vacatur of stay denied*, 126 S. Ct. 1 (2005), *argued*, No. 05-4896 (2d Cir. Nov. 2, 2005).

and permanent gag that accompanies this Patriot Act-expanded National Security Letter power, barring any recipient of such a letter from disclosing its existence to anyone.[29] This latter case has garnered much attention, as the recipient of the National Security Letter is a member of the American Library Association, and there has been special public concern that the Patriot Act's expanded surveillance powers would be used to seize library patrons' records.[30] Moreover, the Department of Justice has made special efforts to repudiate this concern, including then-Attorney General Ashcroft's dismissive characterization of it as "hysterical."[31]

Along with some conservative Supreme Court Justices, some conservative Republican members of Congress also have criticized the administration's post-9-11 excesses. One of the strongest critics of the Patriot Act is Congressman Don Young of Alaska, a member of the House Select Committee on Homeland Security. He has stated that the Patriot Act was the "[w]orst act we ever passed . . . . Everybody voted for it but it was stupid. It was . . . emotional voting."[32]

Another conservative Republican, Congressman Butch Otter of Idaho, has been a leader in Congress to cut back on the Patriot Act's excesses. In 2003, he offered an amendment—the "Otter Amendment"[33]—to the annual spending bill for the Departments of Commerce, Justice, and State, to prohibit funding an overreaching Patriot Act provision that allows the government to conduct so-called "sneak and peek" searches.[34] These are secret government searches of a home or office with a warrant but without any notice to the occupants until potentially long afterward, when it is too late to protect precious privacy rights. The House of Representatives passed the Otter Amend-

---

29. *See* 18 U.S.C. § 2709(c).

30. *See, e.g.*, Alison Leigh Cowan, *At Stake in Court: Using the Patriot Act to Get Library Records*, N.Y. TIMES, Sept. 1, 2005, at B1; Barton Gellman, *The FBI's Secret Scrutiny: In Hunt for Terrorists, Bureau Examines Records of Ordinary Americans*, WASH. POST, Nov. 6, 2005, at A1.

31. *See, e.g.*, Shannon McCaffrey, *Ashcroft Castigates Judges on War Powers*, PITT. POST-GAZETTE, Nov. 13, 2004, at A6.

32. *See* Rick Montgomery, *Federal Patriot Act Meets with Grass-Roots Resistance*, KAN. CITY STAR, May 19, 2003, at A1.

33. *See* H.R. 2799, 108th Cong. (2003) (H. Amdt. 292 by Idaho Representative C.L. ("Butch") Otter).

34. Patriot Act, *supra* note 17, at 285–86 (codified at 18 U.S.C. § 3103(a)).

ment by an overwhelming vote of 308 to 118, with almost half of all House Republicans voting for it.[35]

Meanwhile, we have seen similar activity in the Senate, where a bipartisan group of senators has introduced legislation, the Security and Freedom Enhancement (SAFE) Act, which also would cut back the Patriot Act's excesses.[36] The SAFE Act's co-sponsors include two conservative Republican Senators: Larry Craig of Idaho and John Sununu of New Hampshire.[37] Some Patriot Act provisions were even questioned by the chairman of President Bush's re-election campaign, Mark Racicot, former governor of Montana and former chairman of the Republican National Committee.[38]

As indicated by all of these conservative Republican critics, the Bush Administration is mistaken when it repeatedly tries to dismiss the criticism of its post-9-11 excesses as a mere partisan attack.[39] Worse yet was the accusation by then-Attorney General Ashcroft that those who criticize such civil liberties violations are unpatriotic: He said that we "aid terrorists" and "give ammunition to America's enemies."[40] This reminds me of a recent headline in one of my favorite publications, *The Onion*. This particular headline read: "Bush Asks Congress for $30 Billion to Help Fight War on Criticism."[41] In the same vein, another recent *Onion* headline warned: "Revised Patriot Act Will Make it Illegal to Read Patriot Act."[42] Well, most members of

35. *See* 149 CONG. REC. H7229, 7289–90 (daily ed. July 22, 2003) (statement of Rep. Otter). The Otter Amendment, however, was never signed into law. It was dropped from a year-end spending bill by a joint conference of House and Senate leaders. *See* Jesse J. Holland, *Author Vows Another Assault on the Patriot Act*, GRAND RAPIDS PRESS, Dec. 3, 2003, at A6.

36. *See* Security and Freedom Enhancement Act of 2006, S. 737, 109th Cong. (2005).

37. *See id.*

38. *See* Audrey Hudson, *Kerry Criticized on Patriot Act: Cheney Says Democrat's Original Stance "Was Right,"* WASH. TIMES, June 2, 2004, at A07.

39. *See, e.g.*, Charles Babington, *From GOP, Zero Tolerance for Democratic War Critics*, WASH. POST, May 16, 2004, at A5.

40. Testimony of Attorney General John Ashcroft, Senate Committee on the Judiciary (Dec. 6, 2001), http://www.usdoj.gov/ag/testimony/2001/1206transcript senatejudiciarycommittee.htm.

41. The Onion (July 2, 2003), http://www.theonion.com/content/node/28954.

42. The Onion (Sept. 17, 2003), http://www.theonion.com/content/node/32312.

Congress would not have to worry, since they have admitted that they did not even read the Patriot Act before voting for it![43]

In the ACLU's post-9-11 campaign to keep our great country both safe and free, our allies have included conservative citizens' groups, as well as officials: for example, the American Conservative Union, Americans for Tax Reform, the Free Congress Foundation, Phyllis Schlafly's Eagle Forum, and major gunowners' organizations.[44] Here is how Wayne LaPierre of the National Rifle Association explained to members of the NRA why they should support the ACLU's Safe *and* Free campaign, despite their enthusiastic support for President Bush on gun rights and other issues:

> Maybe you think that with President George W. Bush in the White House, everything is safe. You think you can put aside your principles, just this once, to be a loyal conservative . . . . But if we, as conservatives, don't stand up for these fundamental truths, who will? Never accept the idea that surrendering freedom—any freedom—is the price of feeling safe.[45]

Similarly, we have heard strikingly strong criticisms of post-9-11 executive excesses from the so-called "Religious Right," conservative Christians who campaigned for John Ashcroft's appointment as Attorney General because they agreed with his views on abortion and gay rights. Yet some of these individuals have too decried the new investigative guidelines that then-Attorney General Ashcroft issued after the terrorist attacks, which allow surveillance and infiltration of religious and political groups without requiring any suspicion whatsoever.[46] For example, the then-President of the Family Research Coun-

---

43. *See generally* Declan McCullagh, CNET News.com, *Congress Plans Scrutiny of Patriot Act* (May 9, 2005), http://news.com.com/Congress+plans+scrutiny+of+Patriot+Act/2100-1028_3-5700986.html (noting that many members of Congress did not read the initial enacted version of the Patriot Act).

44. *See* ACLU.org, ACLU Joins Conservatives to Restore Freedoms Lost Under Patriot Act; "Patriots to Restore Checks and Balances" Hopes to Shape National Dialogue (March 22, 2005), http://www.aclu.org/SafeandFree/SafeandFree.cfm?ID=17798&c=206.

45. Wayne LaPierre, Address to the Conservative Political Action Conference, February 2002, Frightened or Free?, http://www.nrahq.org/transcripts/cpac0202.asp (last visited Nov. 20, 2005).

46. *See* The Attorney General's Guidelines for FBI National Security Investigations And Foreign Intelligence Collection (2003), http://www.fas.org/irp/agency/doj/fbi/nsiguidelines.pdf.

cil, Ken Connor, said, "It's important that we [religious] conservatives maintain a high degree of vigilance. We need to ask ourselves, 'How would our groups fare under these new rules?'"[47]

The most recent "strange bedfellows" to join the bipartisan, ideologically diverse critics of the executive branch's sweeping post-9-11 powers are major business organizations, including the U.S. Chamber of Commerce, the National Association of Manufacturers, the National Association of Realtors, the Association of Corporate Counsel, and the Financial Services Roundtable. On October 4, 2005, these organizations wrote to Senator Arlen Specter, Chairman of the Senate Judiciary Committee, to call for cutbacks on the Patriot Act's expansion of the government's power to obtain "voluminous and often sensitive records from American businesses, without judicial oversight or other meaningful checks on the government's power."[48] These organizations objected to the Patriot Act's invasions of the confidentiality rights of business entities themselves, as well as the Act's invasions of the privacy rights of the entities' customers. As they wrote to Senator Specter:

> [W]e are concerned that the rights of businesses to confidential files—records about our customers or our employees, as well as our trade secrets and other proprietary information—can too easily be obtained and disseminated under investigative powers expanded by the Patriot Act. It is our belief that these new powers lack sufficient checks and balances.[49]

As the abolitionist Wendell Phillips said: "Eternal vigilance is the price of liberty."[50] I would add that it is also the price of security. "We the People of the United States"[51]—to quote the opening words of our Constitution—must continue to monitor our government to ensure that it really is effectively protecting our safety and not unjustifiably curbing our freedom.

I would like to conclude by quoting from one of the most re-

---

47. Neil A. Lewis, *Ashcroft's Terrorism Policies Dismay Some Conservatives*, N.Y. TIMES, July 24, 2002, at A14.

48. Letter from Association of Corporate Counsel et al. to Arlen Specter, U.S. Senator (Oct. 4, 2005), *available at* http://www.ala.org/ala/washoff/WOissues/civilliberties/theusapatriotact/BusgrpLtr04oct05.pdf.

49. *Id*.

50. RESPECTFULLY QUOTED 205 (Suzy Platt ed., 1989).

51. U.S. CONST. pmbl.

cent court decisions that agreed with the ACLU—and with our many libertarian and conservative allies—that we Americans need not sacrifice our freedom to ensure our security. In November 2004, the U.S. Court of Appeals for the Eleventh Circuit unanimously upheld the ACLU's challenge to efforts by officials in Columbus, Georgia to invade the privacy and free speech rights of participants in a peaceful demonstration against torture and other human rights abuses.[52] The officials asserted that the so-called War on Terrorism warranted mandatory searches of protestors through the use of a magnetometer.[53] The court resoundingly rejected these claims. As it eloquently declared: "We cannot simply suspend or restrict civil liberties until the War on Terror is over, because the War on Terror is unlikely ever to be truly over. September 11, 2001, already a day of immeasurable tragedy, cannot be the day liberty perished in this country."[54]

52. *See* Bourgeois v. Peters, 387 F.3d 1303, 1316 (11th Cir. 2004).
53. *See id.* at 1307.
54. *Id.* at 1312.

THE VIRTUES OF PREEMPTIVE DETERRENCE

DAVID B. RIVKIN, JR.*

I. INTRODUCTION

Although much about America's twenty-first-century foreign and defense policies seems to irritate its many critics, the so-called Bush Doctrine, announced in a series of speeches following the September 11, 2001, attacks and most articulately expressed in the *National Security Strategy of the United States*,[1] is their favorite target. It vigorously advocates strategic preemption—the use of military force against threats that have not yet fully matured, but nevertheless endanger American security—as a core component of American statecraft in the Twenty-First Century.

The Bush Doctrine was prominently featured in the 2004 presidential campaign. Senator John Kerry argued that the President's formulation was overly zealous and weakened the Atlantic alliance and that preemption, while an option in the case of dire necessity, was not to be bandied about in public and certainly not cast as a centerpiece of America's post-Cold War security strategy. These counter-preemption criticisms, although embraced by many pundits and public officials, both in the United States and abroad, are grounded in fundamental misconceptions about the nature of deterrence and the reasons for the transatlantic discord about core security issues. Even more fundamentally, they manifest a dangerous failure to ap-

---

* Partner, Washington, D.C. office of Baker & Hostetler LLP; Visiting Fellow, Nixon Center; Contributing Editor, *The National Review* and *National Interest* magazines; Member, UN Subcommission on the Promotion and Protection of Human Rights; served in a variety of legal and policy positions at the Departments of Justice and Energy and the White House during the Raegan and H.W. Bush Administrations. This paper, adapted from his remarks during the debate with Professor Alan Dershowitz at the Federalist Society's 2005 National Student Symposium, draws heavily on an article he co-authored for the *Chicago Journal of International Law*. *See* David B. Rivkin, Jr. et al., *Preemption and the Law in the Twenty-First Century*, 5 CHI. J. INT'L L. 467 (2005).

1. *See, e.g.*, THE WHITE HOUSE, THE NATIONAL SECURITY STRATEGY OF THE UNITED STATES OF AMERICA 15 (2002), http://www.whitehouse.gov/nsc/nss.pdf.

preciate the strategic predicament the United States faces in today's world and the corresponding need to adjust the American strategic doctrine and declaratory policy to cope with the new threats.

II. PREEMPTION'S VENERABLE PEDIGREE

The United States has long relied upon preemption to assure its national security; indeed, the Bush Administration's embrace of the doctrine is nothing new.[2] President John F. Kennedy's 1962 decision to forestall Soviet installation of short and intermediate range offensive nuclear missiles in Cuba by declaring a "quarantine" of the island stands as the most prominent modern example of U.S. strategic preemption.[3] Although carefully characterized by the United States as an act of self-defense, the quarantine, directed at Soviet ships carrying nuclear missiles to Cuba, indisputably constituted a belligerent act under international law.[4] Significantly, neither the Soviet Union nor Cuba had attacked the United States, nor was there any imminent threat that the missiles would be launched at the United States once they were deployed. Clearly, President Kennedy acted to preempt an attenuated and distant threat.[5]

---

2. *See generally* JOHN LEWIS GADDIS, SURPRISE, SECURITY, AND THE AMERICAN EXPERIENCE (2004).

3. For an excellent general description of American policymaking during the Cuban missile crisis, see GRAHAM T. ALLISON & PHILIP D. ZELIKOW, ESSENCE OF DECISION: EXPLAINING THE CUBAN MISSILE CRISIS (2d ed. 1999).

4. For a more detailed discussion of this episode, see David B. Rivkin, Jr. et al., *Preemption and Law in the Twenty-First Century*, 5 CHI. J. INT'L L. 467, 481–83 (2005).

5. I have previously articulated this point:

> Kennedy's real objection to Khrushchev's attempt to place short- and intermediate-range ballistic missiles in Cuba was that the missiles would have enabled Moscow to shift the US-Soviet strategic balance precipitously, in a matter of months, rather than gradually, over decades [as Moscow was certain to do in order to implement its military doctrine]. In the context of Iraq, an appropriate analogy [with the set of strategic choices as seen by Kennedy Administration] would have been between accepting the basic proposition that Saddam Hussein could have created (or recreated) a WMD arsenal, provided he did so gradually, and holding that what was really unacceptable was only a precipitous WMD acquisition. The reason it is worthwhile dwelling on these distinctions is to underscore that, as far as the Kennedy Administration was concerned, it could legitimately invoke the anticipatory self-defense doctrine [that is, preemption] when dealing with a comparatively attenuated and highly nuanced threat.

Id. at 482 n.45.

In the post-September 11 geopolitical environment, the imperatives recommending preemption have grown even stronger. Neither toppling the Taliban regime nor killing Osama bin Laden during the summer of 2001 could have reliably prevented the 9-11 attacks. By that point, al Qaeda operatives had already infiltrated the United States and could have proceeded without further assistance. The only way to avert (with a sufficient degree of certainty) clandestine terrorist attacks by pan-national Islamist organizations is to act against them months, or even years, in advance.

Even if preemptive U.S. action cannot destroy the pan-national terrorist threat entirely, taking the battle to the enemy still offers considerable advantages: the U.S. operations against al Qaeda, in addition to killing or capturing numerous operatives, have disrupted the organization's plans and put it on the defensive. In the future, even before the United States actually uses force against a potential adversary, credible preemptive threats can disrupt that adversary's plans, force him to take extra precautions, and diminish his ability to launch new attacks. Moreover, requiring that the United States wait to respond until an attack occurs will often prove ineffectual, given that terrorists routinely relocate their facilities and stand down many of their ongoing activities just before an attack, thereby diminishing their vulnerability to anticipated counterattacks.

Today, preemption's necessity is further reinforced by Islamist terrorists, who, unlike America's Cold War enemies, cannot be reliably deterred. Traditional deterrence fails because of the messianic nature of Islamist terrorists' aspirations, the willingness and even eagerness of their leaders and footsoldiers alike to die for their cause,[6] as well as the more prosaic fact that these groups seldom have valuable assets that may be easily identified and targeted. Even the most robust deterrence "means nothing against shadowy terrorist networks with no nations or citizens to defend."[7]

---

6. Unfortunately, by discounting the impact that enemies' strategic cultures have on their warmaking, the United States often commits a mistake common to most states that practice the "rational" decisionmaking style.

7. President George W. Bush, Remarks at the 2002 Graduation Exercises of the United States Military Academy (June 1, 2002), http://www.whitehouse.gov/news/releases/2002/06/20020601-3.html.

Despite preemption's palpable necessity, its opponents tend to focus on the Bush Administration's candid embrace of preemption as a doctrine,[8] rather than as an acknowledgment of its past record and possible future utility. These same pundits also fret that preemption has taken too prominent a place in the administration's strategic thinking. They prefer to treat preemption like the crazy aunt in the attic: always looming in the background but rarely mentioned.[9] Unfortunately, to extol deterrence sans preemption, or even to claim that preemption ought not to be a component of one's declaratory strategy, is to betray a profound misunderstanding of the logic of deterrence itself.

---

8. While some critics have questioned whether preemption can even be considered a doctrine, as distinct from a combat tactic, history is replete with examples of states choosing preemption as an organizing strategic principle. The Romans employed strategic preemption, perhaps most famously in the Third Punic War (149–146 B.C.). Convinced that a resurgent Carthage threatened Rome, Roman statesman Marcus Porcius Cato (Cato the Elder) tirelessly advocated preemptive war, supposedly ending every speech to the Senate with the same words: "Also, Carthage, methinks, ought utterly to be destroyed." In due course it was. Legend has it that the Romans salted the fields of Carthage so that no living thing, and no rival, might grow there again. *See* 2 PLUTARCH, *Marcus Cato*, *in* LIVES 516, 541 (John Dryden trans., J.M. Dent & Sons 1962) (1683). The British also employed strategic preemption when other countries threatened the naval dominance upon which their empire depended. Perhaps the most famous example of British strategic preemption was Admiral Horatio Nelson's 1801 destruction of the Danish fleet at Copenhagen to forestall the possibility that an as-yet-unrealized Danish alliance with Napoleonic France could tip the balance of power against Britain. *See* ARTHUR HERMAN, TO RULE THE WAVES: HOW THE BRITISH NAVY SHAPED THE MODERN WORLD 366 (2004).

9. To be sure, there is a diversity of views among the critics, with some arguing that the United States should never publicly articulate the preemptive doctrine, and others claiming that the problem is not the public articulation as such, but only the excessive emphasis on the verities of preemption.

III. THE DELICATE LOGIC OF DETERRENCE

A. Balancing the Goals

Deterrence, a venerable concept,[10] was greatly refined at the dawn of the nuclear era by a number of brilliant American strategists.[11] Notably, none of these expositors of modern deterrence theories would have disagreed with the proposition that the American strategic doctrine, together with the associated force structure and the weapons employment policies, has always meant to advance multiple goals. An ideal strategic policy, both at the declaratory and operational levels, should provide the highest possible level of dissuasion of one's enemies and ensure their defeat if deterrence fails, while simultaneously enjoying the warmest reception from one's allies, even when they cannot be overtly identified with the understood policies of their patron.

Unfortunately, the inherent tension between and among these goals often makes them incapable of being achieved simultaneously. We seem to have forgotten this; even during the Cold War, the United States and its European allies had long-standing, contentious disagreements over deterrence requirements. At the risk of oversimplifying several decades' worth of transatlantic debates, it can be said that the United States generally displayed much greater preoccupation with the survivability of its own nuclear forces,[12] targeting and employment flexibility, and continuous refinement of declaratory policy than did its European allies.[13]

---

10. One of the earliest and most concise expositions of deterrence theory was given by the fourth-century Roman strategist Flavius Vegetius: "Igitur qui desierat pacem, praeparet bellum . . . ." ("Therefore, he who desires peace, let him prepare for war . . . .") FLAVIUS VEGETIUS RENATUS, EPITOMA REI MILITARIS 121–23 (Leo F. Stelten ed. & trans., Peter Lang Publ'g, Inc. 1990). Indeed, the gist of Vegetius's maxim, that one is safest when one can frighten one's enemies through demonstrating strength, explains a caveman's brandishing a spear at a threatening Cro-Magnon as easily as it does the deployment of nuclear ballistic missile submarines.

11. *See, e.g.*, BERNARD BRODIE, STRATEGY IN THE MISSILE AGE (1959); HERMAN KAHN, ON THERMONUCLEAR WAR (2d ed., Free Press 1969) (1960); THOMAS C. SCHELLING, ARMS AND INFLUENCE (1966); Albert Wohlstetter, *The Delicate Balance of Terror*, 37 FOREIGN AFF. 211 (1959).

12. *See generally* Wohlstetter, *supra* note 11.

13. For an excellent discussion of intra-allied tensions arising out of diverging theories of deterrence, see HENRY KISSINGER, DIPLOMACY 594–619 (1994).

During the last two decades of the Cold War, a number of U.S. strategists (American civilian and military strategists were also not always in accord) tended to argue that the best way to deter the Soviets, who believed that nuclear war was winnable, was to anchor deterrence within the context of a credible nuclear war-fighting posture and doctrine capable of threatening the military and political control centers of the Soviet Communist Party and its satellites.[14]

In contrast, European strategists clung stubbornly to the "mutual assured destruction" theory of deterrence and rejected war-fighting approaches, especially the use of tactical nuclear weapons, as dangerously destabilizing. They countered American efforts to buttress deterrence through war fighting by arguing that, in all circumstances, the best way to deter was to make nuclear war as horrible and, therefore, unthinkable, as possible. Europeans also tended to find the U.S. government's public emphasis on nuclear war fighting unsettling. Meanwhile, even those European strategists who did not denigrate American advocacy of war-fighting options as inherently destabilizing displayed a kind of doctrinal ennui, arguing that any efforts to refine one's nuclear force posture or employment options were essentially futile.[15]

Thus, even during deterrence's Cold War "golden age," successive American administrations chose to espouse their preferred approach to deterrence, even at the cost of unsettling European allies and diminishing the reassurance value of the U.S. policies. In general, for more than fifty years, transatlantic disagreement on strategic issues—ranging from the proper strategy to relieve Berlin, to the Reagan Administration's deployment of cruise and Pershing II missiles to Western Europe—was the rule rather than the exception.

---

14. *See, e.g.*, Leon Sloss & Marc Dean Millot, *U.S. Nuclear Strategy in Evolution*, STRATEGIC REV., Winter 1984, at 19.

15. Thus, for example, one of Britain's best-known nuclear policy pundits, Lawrence Freedman, opined in the late 1980s:

> [T]he nuclear strategists had still failed to come up with any convincing methods of employing nuclear weapons should deterrence fail that did not wholly offend common sense, nor had they even reached a consensus on whether or not the discovery of such methods was essential if deterrence was to endure.

Lawrence Freedman, *The First Two Generations of Nuclear Strategists*, *in* MAKERS OF MODERN STRATEGY: FROM MACHIAVELLI TO THE NUCLEAR AGE 735, 778 (Peter Paret ed., 1986).

B. Enhancing Deterrence Quality

America's current strategic predicament is that at least some of our enemies both hate us and discount greatly our ability to resist.[16] Indeed, the only tenet of Islamic fascism—whose cargo numbers religious intolerance, theocracy, sexual apartheid, and political repression—that resonates with the world is its belief that an affluent West either cannot or will not defend its culture and is therefore particularly undeserving of the global stature it enjoys.

Presented with these challenging new deterrence requirements, the Bush Administration properly opted to enhance the ability of our declaratory policy to deter our enemies, even at the cost of diminishing its reassurance aspects. As Walter Russell Mead observed: "[T]here is little doubt that the administration believed that it was more important to frighten and deter potential enemies than to reassure friends. If the good guys had to be scared in order to make sure the bad guys knew you were serious, so be it."[17]

Although terrorist groups such as al Qaeda are beyond deterrence, maintaining robust deterrence remains a vital task for American statecraft because state sponsors of terror still have much to lose. Convincing these states' rulers that embarking on a particular course of action will trigger a swift regime change may well dissuade them from supporting terrorism.[18] Even rogue regimes that do not exercise full control over their territory will probably try to cooperate if sufficiently persuaded that the United States will not tolerate their support for terrorism.

Thus, an explicitly preemptive American strategic doctrine is not antithetical to deterrence. Rather, the threat of U.S. preemption broadens the range of conduct that may be deterred. Such broadening is particularly desirable given the nature of mod-

---

16. U.S. deterrence is not as consistently effective or as uniform in quality as it was during the Cold War. While some enemies, such as North Korea, continue to fear the United States and credit it with considerable resolve and staying power, others—most prominently Islamic terrorists—appear to believe that the United States is fundamentally weak.

17. WALTER RUSSELL MEAD, POWER, TERROR, PEACE, AND WAR: AMERICA'S GRAND STRATEGY IN A WORLD AT RISK 115 (2004).

18. One wonders if even the Taliban's Supreme Leader, Mullah Omar, would have played along had he known that indulging Osama bin Laden's homicidal dreams would bring about a regime change in Kabul, Afghanistan.

ern terrorist threat scenarios, in which seemingly benign activities, such as supporting religious charities or pursuing "peaceful" nuclear activities, seamlessly morph into terrorist attacks and Weapons of Mass Destruction (WMD) programs.

Significantly, preemption robs terrorist patrons of their greatest asset: deniability of culpability, or the ability to offer and later deny aid to terrorists. In response, the new salutary strategic arithmetic created by the U.S. preemptive posture is deliberately somewhat hazy to potential enemy states, nudging them toward demonstrating that they do not harbor terrorists, rather than forcing the United States to prove beyond a reasonable doubt that they do.

To be sure, seeking to shift the burdens of production and persuasion onto the enemy is not entirely new in statecraft. Traditional deterrence theory always envisioned a role for so-called confidence-building measures (CBMs), whether practiced in the context of an arms-control regime or unilaterally. For example, the so-called CFE Treaty,[19] dealing with conventional forces in Europe, featured a wide range of CBMs, including restrictions on the numbers of forces that could be deployed in close proximity to the inter-German border, notification of military exercises and maneuvers, and so on, designed to alleviate NATO's perennial Cold War-era concern about a clandestine attack by the Warsaw Pact forces.

Similarly, the Clinton Administration's announcement that it was de-targeting Russian military, political, and economic assets represented a CBM designed to reassure the Russian leadership that it was no longer viewed as an enemy.[20] In the post-September 11 environment, the threat of U.S. preemption serves as a potentially potent CBM-inducing tool, oriented toward rogue states that have flirted with WMD programs or supported terrorist organizations or both. Indeed, given the inherent difficulty, or some would even say impossibility, of devising balanced intelligence analysis procedures that neither

---

19. Treaty on Conventional Armed Forces in Europe, Nov. 19, 1990, S. TREATY DOC. NO. 102-8, 30 I.L.M. 6.

20. The U.S. de-targeting announcement came in the context of a broader U.S.-Russian agreement to cease targeting nuclear missiles at one another. While easily reversible, the agreement was widely viewed as a positive step. *See* Jurek Martin & John Lloyd, *Clinton Signs Three-Way N-Missile Agreement: Ukraine Nuclear Arsenal to Go as US and Russia End Mutual Targeting*, FIN. TIMES, Jan. 15, 1994, at 1.

"under-connect" the dots, as before 9-11, nor "over-connect" them, as with the Iraqi WMD stockpile estimates,[21] any strategy that is capable of prompting the rogue states to assume compliance burdens has much to recommend. The Libyan decision, prompted at least in part by Mohammar Khaddafi's absorption of the "regime change" lessons of Iraq, to go public with its WMD efforts and renounce them, while inviting an intrusive weapons inspection, is a good case in point.

C. Dealing with a "Weak Horse" Problem

The Bush Administration's public articulation of the preemption doctrine has performed another indispensable function: dispelling an impression of American weakness. For most Americans, ingrained with the belief that the United States is the world's strongest military and economic power, the notion of American weakness likely seems laughable. Unfortunately, this has not been the view of our enemies, at least some of whom question our will to fight and, most importantly, our staying power (or even believe that America's very wealth and leisure have robbed it of its old muscularity and tragic acceptance of the occasional utility of war).

This disdain for America's staying power is not particularly surprising. Several decades' worth of American retreats under fire in Vietnam, Lebanon, and Somalia; our failure to respond, except through ineffectual law enforcement methods, to Hezbollah and al Qaeda attacks on American assets in the 1990s; and our seeming toleration of the kidnapping and torture of American personnel in the Middle East have all fostered the impression that, in Osama bin Laden's own words, America is a "weak horse."[22] Properly erasing this impression has become

---

21. For an example of a discussion about this issue, see David B. Rivkin, Jr. & Lee A. Casey, *Intelligence as Usual: The Flawed 9/11 Commission Report*, NAT'L REV. ONLINE, Aug. 24, 2004, http://www.nationalreview.com/comment/rivkin_casey200408240839.asp.

22. Bin Laden's actual words were: "[W]hen people see a strong horse [Islamic fundamentalism] and a weak horse [the United States], by nature, they will like the strong horse." These remarks were part of a tape released by the U.S. Department of Defense shortly after the attacks of September 11, 2001. For a transcript of bin Laden's remarks, see CNN.com, Transcript of Osama bin Laden Videotape (Dec. 13, 2001), http://archives.cnn.com/2001/US/12/13/tape.transcript/. The Islamists' belief that America was a "weak horse" was all the more dangerous because the Soviet defeat in Afghanistan encouraged the conceit that Islamic

a paramount U.S. policy priority, critical to reclaiming a credible deterrence posture.

D. Domestic and Bureaucratic Politics

There are also at least two additional important domestic reasons for the Administration's public embrace of preemption. To begin with, as the 9-11 Commission has pointed out, one key reason why both Democrat and Republican Administrations failed to take the battle to al Qaeda was bureaucratic resistance by the Central Intelligence Agency (CIA) and the Department of Defense against engaging in what they believed to be difficult and uncertain paramilitary operations on foreign soil. This institutional foot dragging was reinforced by the peculiarly debilitating legal regime that grew out of various prior investigations of alleged CIA abuses.[23] A public affirmation that the U.S. government, at its highest levels, embraced the legitimacy of strategic preemption, combined with the well accepted view that wartime operations are subject to a distinctly different legal regime than covert actions, can do much to lift this pernicious bureaucratic fog.

There are also a number of important benefits that would accrue from an explicit articulation by the U.S. government of a preemptive strategy, including the provision of a proper doctrinal context for operational planning and force development, with particular emphasis on the development of both conventional and nuclear long-range strike capabilities. Finally, and perhaps most vitally, comes the problem of public opinion: simply put, in any strategic environment in which preemption can be expected to be used, the U.S. government will have to justify doing so publicly and in advance. Launching preemptive attacks in the face of adverse public opinion might have been an option for a Julius Caesar or even a Benjamin Disraeli,

---

fighters were a "strong horse." *See generally* GILLES KEPEL, JIHAD: THE TRAIL OF POLITICAL ISLAM (Anthony F. Roberts trans., Harvard Univ. Press 2002).

23. *See* NAT'L COMM'N ON TERRORIST ATTACKS UPON THE U.S., THE 9/11 COMMISSION REPORT 90, 113, 131–33 (authorized ed. 2004). The Report details how the CIA was particularly obsessed by the impact, real and exaggerated, of the Executive Order Against Assassinations on the various U.S. government plans to apprehend or eliminate Osama bin Laden, whether carried out by U.S. operatives or third parties. *See* Exec. Order No. 12,333, 46 Fed. Reg. 59,941, 59,952 (Dec. 4, 1981). Excessive concern about collateral damage was another paralyzing factor. *See* NAT'L COMM'N ON TERRORIST ATTACKS UPON THE U.S., *supra*, at 131.

but it is not an option available to an American President of the Twenty-First Century. This proposition is fully validated by the continuing political fallout from the Iraq war, which was justified primarily, albeit not exclusively, as a way to deal with Iraq's WMD threat, rather than as a preemptive response to the grave geopolitical threat posed by the Hussein regime (to say nothing of the regime's barbarity).

E. Preemption's Many Alleged Sins

1. Destabilization: Real or Imagined?

Despite these ample justifications, the Bush Administration's public embrace of preemption has been roundly condemned. To begin, preemption foes opine, in an argument eerily reminiscent of Cold War-era lamentations over deterrence rooted in a viable nuclear war-fighting posture, that preemption is dangerously destabilizing and will result in more, not fewer, conflicts. The reasoning here appears to be that allowing states to use force in advance of an actual attack would make them more prone to use force promiscuously—hypothetically, that the U.S. regime change in Baghdad might induce Argentina to invade the Falklands again.

Regurgitating this old action-reaction argument makes even less sense today than it did fifty years ago.[24] Few foreign policy analysts now believe a country that wishes to go to war with another nation, for whatever domestic ideological or foreign policy reasons, has somehow been "set free" by the Bush Doctrine. That some countries may find it expedient to justify their military exploits by referring to the U.S. strategy, as Vladimir Putin's Russia routinely does when trying to justify its brutal conduct in Chechnya, does not mean that they would not have behaved in exactly the same way in the absence of the Bush Doctrine. Even fewer maintain that, but for the Bush Doctrine, these conflicts would not occur. This criticism ignores the realities of international relations and ascribes to the United States an unrealistic degree of doctrinal influence.

---

24. Even during the Cold War, the argument that some force postures generated great crisis instability was considerably overblown. Moreover, to the extent that the crisis instability logic had any validity at all, it was the vulnerability of one's forces to an enemy's preemptive attack, rather than one's commitment to a war-fighting employment policy, that was most consequential.

There is, of course, no doubt that a "rush" to military action, inherent in a hair-trigger preemptive strategy, can result in a miscalculation. Ultimately, the wisdom of preemption will be judged post hoc. The Anglo-French-Israeli effort in 1956 to retake the Suez Canal is mostly remembered for its failure, which has obscured the original and otherwise defensible reasons for which the effort was launched. These dangers, however, must be weighed against the nature and scope of the threats present in today's world, along with the terrible danger of inaction. The September 11 attacks and the deaths of more than three thousand Americans is the archetypal case in point.

2. Upsetting the Allies

Another set of anti-preemption arguments is based on the claim that the Bush Administration's embrace of preemption has dramatically impaired the reassurance aspect of our policy, by alienating the international community in general, and Europe in particular. Upsetting one's allies does carry a heavy price, both internationally and domestically; if at all possible, it is to be avoided. At its founding, America expressed a desire to respect the "opinions of mankind,"[25] and the United States clearly wants to be viewed favorably by the rest of the world. Given the compelling arguments that support the Administration's policy, however, the United States has no choice but to proceed.

Moreover, European criticism of American policies would not appreciably abate even if the Bush Doctrine was never publicly unveiled. Rather, those criticisms reflect a fundamental dissonance between the Bush Administration and most Western European elites over the root causes of the threat faced in the post-September 11 world and on how to combat this threat. The United States has been advocating the promotion of democracy and freedom as the key Western foreign policy goals in the Greater Middle East, while Western Europe has stressed economic development and stability.[26]

---

25. THE DECLARATION OF INDEPENDENCE para. 1 (U.S. 1776).

26. In a vivid demonstration of this dichotomy, President Bush used the word "freedom" over a dozen times in his second inaugural address, but German Chancellor Schroeder's speech at the 2005 *Werkunde* conference contained half a dozen references to "stability" without a single invocation of freedom. *See* Fred Hiatt, Editorial, *A Few Words Between Friends*, WASH. POST, Feb. 15, 2005, at A17.

Of course, these profound transatlantic differences are not new. They reflect longstanding differences in our origins and values,[27] which were largely obscured by the imperatives of the Cold War.[28] As argued by Max Boot, among others, the carnage of the World Wars produced a generation of myopic pacifists in Europe who feel that nothing is worth fighting and dying for and act as if war itself (rather than the failure to reject appeasement and to provide credible military deterrence against aggressive, undemocratic regimes) accounted for millions of European dead.[29] Not only do Europeans and Americans disagree about the root causes of our current problems, but virtually all Europeans, whether the elites or ordinary citizens, exhibit an enormous aversion to the use of military force. European distaste for Israel's efforts to defend itself during the most recent round of the intifada demonstrates that this aversion extends to all uses of military power, whether preemptive or reactive.[30]

Pacifism aside, Europe's current and projected shortfall in military capabilities is also responsible for shaping European aversion to the use of force in general and preemption in particular.[31] Europe's ability to project military power, or even to

---

27. For an excellent discussion of the differences between the United States and Europe regarding a broad range of international issues, see ROGER REED, UNCOMMON VALUES: THE UNITED STATES AND EUROPE AT THE UNITED NATIONS COMMISSION ON HUMAN RIGHTS (2003).

28. *See* Lee A. Casey & David B. Rivkin, Jr., *Europe in the Balance*, POL'Y REV., June & July 2001, at 41, 41–43; Rivkin, Jr. et al., *supra* note 4, at 493–94. *See generally* FRANCIS FUKUYAMA, STATE-BUILDING: GOVERNANCE AND WORLD ORDER IN THE 21ST CENTURY (2004); ROBERT KAGAN, OF PARADISE AND POWER: AMERICA AND EUROPE IN THE NEW WORLD ORDER (2003).

29. Max Boot, Editorial, *Risky Path for Pacifist Europe*, L.A. TIMES, June 3, 2004, at B11.

30. The *Washington Post's* Jim Hoagland observed that Europeans opine both that Europe's core post-World War II project—the creation of the European Union—"is a peace process in itself" and that Europeans "believe fundamentally that war can never again serve any useful purpose." Jim Hoagland, Editorial, *The Limits of Reconciliation*, WASH. POST, Feb. 20, 2005, at B7.

31. One of the best pieces of evidence of this aversion is the embrace by all European powers, including even Britain, of an extremely crabbed view of the existing international law strictures that govern recourse to force. This point is vividly demonstrated in the recently declassified and released legal opinion by Britain's Attorney General, Lord Goldsmith, prepared to advise Prime Minister Blair on the legality of the Iraq war. The fact that Lord Goldsmith does not countenance the use of preemption is not particularly surprising. However, even when it comes to analyzing the legal status quo in Iraq prior to the initiation of combat operations

make a meaningful contribution to those military ventures of which it does approve, such as the deployments to Bosnia and Kosovo, has decreased to the point of embarrassment. Indeed, Europe has been so strained by the comparatively modest requirements of the Afghan mission, involving the deployment of fewer than nine thousand troops, that it is difficult to envision how Germany or France could have contributed much to the additional military operations, whether in Iraq or elsewhere, even if they supported the U.S. policy.[32]

To be sure, not so long ago, most of the United States' European allies supported the preemptive uses of military force for

---

in 2003 by the Coalition Forces, his observation that "as the [2001] cease-fire conditions were set by the Security Council in resolution 687, it is for the Council to assess whether any such breach of those obligations has occurred" is highly troubling. Memorandum from Lord Goldsmith, Att'y Gen., Gr. Brit., to Tony Blair, Prime Minister, Gr. Brit., (Mar. 7, 2003), *available at* http://image.Guardian.co.uk/sys-files/Guardian/documents/2005/04/28/legal.pdf. The first Gulf War was primarily predicated upon the exercise of individual and collective self-defense rights by the Coalition parties, having been launched in a response to a clear-cut aggression against Kuwait by Saddam Hussein. The Security Council did not authorize the use of force against Iraq under its Chapter VII authority; it merely acknowledged and endorsed the exercise of collective self-defense rights by the Coalition partners. The cease-fire agreement, referred to by Lord Goldsmith, was negotiated by the Coalition military officers and signed by them as well as by representatives of Saddam Hussein. That the Security Council blessed this cease-fire agreement in resolution 687 did not detract from the agreement's independent legal status. In short, to take the view that sovereign states, like the United States and Britain, which have been exercising since 1990 their collective self-defense rights engaging in combat against Saddam Hussein's forces on dozens of occasions between 1991 and 2003, have somehow abdicated to the Security Council their right to reach an indisputable conclusion that Saddam Hussein has repeatedly violated his cease-fire obligations, is nothing short of remarkable. *See* YORAM DINSTEIN, WAR, AGGRESSION AND SELF-DEFENCE 50–53 (3d ed., Cambridge Univ. Press 2001) (1988). To underscore that Lord Goldsmith's opinion reflected more than his own legal idiosyncrasies, it is worth recalling that "Goldsmith's legal opinion was drawn up in part to answer the demands of Adm. Michael Boyce, then chief of the [British] defense staff. He wanted assurances that soldiers could not be prosecuted either by the international court or in the British justice system for taking part in an illegal war." Glenn Frankel, *Blair Releases Memo Questioning Legality of Iraq War*, WASH. POST, Apr. 29, 2005, at A16. John Keegan, one of Britain's preeminent military historians, publicly complained that the legal requirements and procedures imposed on the British military by their political masters are greatly undermining unit cohesion and combat effectiveness of British forces. *See, e.g.*, John Keegan, *Bad Law is Making a Just War So Much Harder to Fight*, DAILY TELEGRAPH, June 2, 2005, at 20. A related problem is that the existing and growing divergence between the United States and other countries concerning the legal imperatives bearing on the rules of engagement renders effective coordination of warfare virtually impossible.

32. *See* Rivkin, Jr. et al., *supra* note 4, at 492.

humanitarian interventions, for example, in Bosnia and Kosovo, provided that the United States was willing to lead the way. Their enthusiasm was shared by much of the international law professoriate and the UN mandarins, who coined the term "duty to prevent" to describe the obligation to use military force to stop genocidal despots—even though President Clinton's deployment of American arms was sanctioned neither by the U.S. Congress nor the UN Security Council.[33] Their enthusiasm for U.S.-led humanitarian interventions faded when it dawned on them that the preemption doctrine was not inherently limited to those ventures that they favored, and that the United States was the most obvious beneficiary and user of the preemption doctrine.

Unfortunately, as most recently manifested by the European Union's apparent interest in lifting the arms embargo against China, despite Beijing's atrocious human rights record and destabilizing policies in Asia, many of our erstwhile allies seem to believe that managing and disciplining American power is the only imperative of their statecraft.[34] At the same time, they discount the threats posed by Iran, North Korea, Syria, and other rogue countries, and display little concern with deterring and defeating the forces of radical Islam. Thus, opposing the Bush Doctrine was something they felt they had to do, especially when even silent opposition to radical Islam can trigger either a terrorist attack or increased unrest among unassimilated Muslim populations of Europe. By contrast, strident invective against the United States resonates with at least a portion of American elites and elicits no meaningful consequences from the American government.

IV. THE LEGACY OF IRAQ

There is, of course, another wrinkle to the preemption debate, and it concerns Iraq. Unfortunately, the broader public

---

33. For a useful discussion of this issue, see Lee Feinstein & Anne-Marie Slaughter, *A Duty to Prevent*, 83 FOREIGN AFF. 136 (2004). *See also* INT'L COMM'N ON INTERVENTION AND STATE SOVEREIGNTY, THE RESPONSIBILITY TO PROTECT (2001), *available at* http://www.iciss.ca/pdf/Commission-Report.pdf.

34. In a recent discussion, the author was told by a mid-level European defense official that some of his colleagues were quite happy with the notion of having China improve its military capabilities to the point where it can "contain" American power, or at least limit Washington's freedom of action in future crises.

debate over preemption has become heavily intertwined with the legal and policy discourse about the Iraq war. These arguments are not concerned with preemption as such, but rather relate to the broader issue of when and how American military power is to be used in the Twenty-First Century.

Not surprisingly, since legal issues these days often dominate policy debates, there are some critics, who, not particularly concerned with the policy merits of the Bush Doctrine, argue that preemption is illegal as a matter of international law and specifically single out U.S.-led military operations in Iraq as an illegal war. The critics' legal arguments reflect an improper parsing of the UN Charter that not only renders nugatory part of the Charter, but also strips legal protections from non-signatories. These legal arguments amount to a strained attempt to proscribe all uses of force, except those which are in response to an armed attack on one's territory or pursuant to explicit Security Council authorization.[35] The notion that the Charter somehow vitiated centuries-old norms of customary international law, which expressly authorized anticipatory defense, borders on the ridiculous. In any case, since the Charter was adopted in 1945, dozens of states, including the United States, Britain, France, Israel, Sweden, Russia, and China, have used force in a preemptive mode while claiming that their behavior was fully consonant with international law.[36]

Legal issues aside, the Iraq war has become a poster child for the preemption foes in two distinct ways. The first is that, given our failure to find WMD stockpiles in Iraq, the United States should not use force preemptively, absent perfect, or near-perfect, information; a frequent companion observation is that any intelligence relating to distant or emerging threats will inevitably be wrong. It is indeed very difficult to penetrate a closed, repressive society, or to estimate reliably the pace of development for complicated and secret weapons. Further, as the United States has learned from its experience with Saddam

35. For an extensive discussion of this issue, see Rivkin, Jr. et al., *supra* note 4, at 474–79.

36. Quite aside from the legality of preemption, the Iraq war was perhaps the most legally justified war of all time. Numerous Security Council Resolutions provided the legal basis for the use of force, as did Saddam's default of his cease-fire obligations. Legally, the 1991 Gulf War was still alive and well in 2003. The past decade has featured repeated U.S. and British enforcement of no-fly zones and intermittent combat.

Hussein, this task becomes virtually impossible in the face of deliberate strategic deception: U.S. analysts underestimated, prior to 1990, the extent of Iraq's nuclear weapons program, while, in 2001, they overestimated the extent to which Iraq had retained its WMD stockpiles.[37]

Significantly, it is often easier to predict, based on an analysis of the long-term trends, what is *eventually* going to happen rather than what will happen tomorrow because the quality of strategic threat forecasts, which involve broad shifts in the balance of power driven by the overarching goals of various geopolitical actors, tends to be better than the assessment of near-term tactical threats.[38] Accordingly, preemption against medium and long-term threats makes more sense than preemption against near-term threats.[39]

Another frequently invoked Iraq-related "lesson" is that it is not easy to bring the preemptive use of force to a successful conclusion. It is true that the Iraq war has proven difficult, protracted, and costly, with currently more than two thousand Americans killed, thousands wounded, reserves stretched, hundreds of billions of dollars spent, and the bulk of U.S. fighting forces tied up in Iraq and not available for deployment elsewhere. To advocate for preemption, however, is not the same as to promise that its execution would lead to easy and painless military campaigns.[40] Whether the costs and the sacrifices involved were and are worth it is something about which reasonable men can disagree; much depends upon the success

---

37. For an example of a discussion about this issue, see David B. Rivkin, Jr. & Lee A. Casey, *Saddam, Nikita and Virtual Weapons of Mass Destruction: A Question of Threat Perception and Intelligence Assessmen*t, IN THE NAT'L INTEREST, June 12, 2003, http://www.inthenationalinterest.com/Articles/Vol2Issue23/Vol2iss23RivkinCasey.html.

38. To demonstrate the validity of this proposition, it would suffice to recall that while U.S. intelligence at first underestimated Iraq's nuclear weapons program and later overestimated it, there is no serious doubt that Saddam's long-term ambitions included the possession of a large diversified WMD arsenal, featuring chemical, biological, and nuclear devices and their delivery systems.

39. Even a threat that is far from certain merits a strong U.S. response if the consequences of inaction could be extremely grave. Conversely, of course, a trivial or minor threat does not merit the use of preemption, unless there is compelling evidence that it is certain to arise.

40. It is unfortunate that a number of erstwhile supporters of the Iraq war fling unjustified criticisms at the way in which the Bush Administration has prosecuted the war, all in an effort to rehabilitate the notion that perfect wars, low in blood and treasure expended, are feasible and therefore, preemption is not that big of a deal.

of the end game in Iraq, which is yet to be fully played.[41] To be sure, preemption, because it can lead to a well planned military campaign prosecuted at the time and in the manner of one's choosing, can be more discriminating and successful than a campaign fought in response to an enemy's first strike.

V. CONCLUSION

Public articulation of strategic preemption doctrine broadens deterrence, updating it to fit the exigencies of American statecraft. And, if deterrence, however buttressed, is unavailing, the execution of a preemptive option is the best way to assure U.S. security. The arguments against preemption are fundamentally flawed. Rarely consistently or uniformly applied, they are hardly directed at preemption as such, but, rather, are driven by underlying differences between the United States and its allies over the question of how and when to use force in today's world. Significantly, the United States' allies' criticism, far from revealing any superior wisdom, demonstrates instead their opposition to U.S. power and their willingness to abdicate responsibility for dealing with the gravest threats of the Twenty-First Century—a convenient stance that will persist until either they themselves fall victim to a terrorist attack similar to the magnitude of September 11, or so alienate the United States that it at last withdraws its military protection from an otherwise mostly defenseless European continent.

41. Survival of a viable and democratic Vietnam, rather than televised 1975 images of helicopters overflowing with American soldiers and Vietnamese refugees, would have given rise to radically different popular wisdom concerning American strategic preemption in the early 1960s.

TOWARD THE FRAMERS' UNDERSTANDING OF "ADVICE AND CONSENT": A HISTORICAL AND TEXTUAL INQUIRY

ADAM J. WHITE*

I. THE FRAMING AND RATIFICATION OF THE ADVICE AND CONSENT CLAUSE.....110
A. The Constitutional Convention.....110
1. Appointment of Judges and Other Officers.....110
2. Matters Related to Senate Advice and Consent.....121
B. Post-Convention Debates.....126
1. The *Federalist Papers*.....127
2. State Ratification Debates.....129
3. Post-Ratification Debate.....131
II. THE MASSACHUSETTS EXPERIENCE.....131
A. The History of the Council.....132
B. The Council in Practice: "Advice" and "Consent".....135
C. Summary.....140
III. MADISON'S ALTERNATIVE, AND COMPLETIONS VERSUS VETOES.....141
A. Madison's Proposal Versus Gorham's Advice and Consent.....142
B. A Textual Framework: Completion Versus Veto.....143
1. Veto or Completion?.....144
2. Discretionary Completion or Mandatory Completion?.....145
C. Summary.....146

---

* Associate, Baker Botts L.L.P.; J.D., Harvard Law School, 2004; B.B.A., University of Iowa, 2001. The author thanks David Barker, Adam Budesheim, Philip Hamburger, Douglas Kmiec, Andrew Oldham, Richard Parker, Nels Peterson, and Kevin White for helpful comments. This Article represents only the views of the author.

IV. CONCLUSION 147

INTRODUCTION

During President George W. Bush's first term, no domestic issue excited the Republican and Democratic bases like judicial nominations.[1] Upon the President's reelection, attention on the war for the courts continued as the President announced the renomination of twenty people whose first-term nominations had not received up-or-down votes in the full U.S. Senate.[2] In 2005, President Bush's first nomination to the Supreme Court—that of Judge John Roberts to replace the late Chief Justice William Rehnquist—was confirmed without delay. As of the writing of this article, the President's second nomination—that of Judge Samuel Alito to replace Justice Sandra Day O'Connor —faced scattered threats of opposition-party delay.[3]

This marks only the most recent battle in the war for the courts. Beginning in earnest with the nomination of then-Justice Abe Fortas to replace outgoing Chief Justice Earl Warren, and continuing through the contentious nomination battle over Judge Robert Bork and the media spectacle of the Clarence Thomas hearings, the fight in the Senate to seat judges had risen to Beltway prominence by the mid-1990s.[4]

During the Clinton presidency, however, Senate refusal to confirm nominations had assumed a markedly different form. No longer was the Senate voting to reject nominees, as it had done with the Bork nomination and many others in the past; instead, the Senate was not voting at all.

President Clinton's nominations to a variety of seats on the courts of appeals failed to result in appointments as the Repub-

---

1. *See, e.g.*, Editorial, *The Choice on Courts*, WASH. POST, Oct. 23, 2004, at A22.

2. *See* Michael A. Fletcher & Helen Dewar, *Bush Will Renominate 20 Judges*, WASH. POST, Dec. 24, 2004, at A1.

3. *See, e.g.*, James Kuhnehn & William Douglas, *Senators on Both Sides Brace for Clash on Alito; Republicans, Democrats Warily Mention Filibuster*, CHARLOTTE OBSERVER, Nov. 1, 2005, at 13A.

4. For excellent histories of judicial nominations battles, see generally HENRY J. ABRAHAM, JUSTICES, PRESIDENTS, AND SENATORS: A HISTORY OF THE U.S. SUPREME COURT APPOINTMENTS FROM WASHINGTON TO CLINTON (rev. ed. 1999), and HENRY B. HOGUE, CONG. RESEARCH SERV., ORDER CODE RL31171, SUPREME COURT NOMINATIONS NOT CONFIRMED, 1789–2004 (updated Mar. 21, 2005).

lican Senate majority left many of the nominations to languish in committee without hearings or full Senate votes. From 1995 to 2000, the President nominated 263 people to seats on the district court benches, and 84 to seats on the courts of appeals.[5] Of the district court nominees, 80% received committee hearings, 78% received committee votes, and 76% received full Senate votes.[6] Of the court of appeals nominees, 65% received committee hearings, 62% received committee votes, and 55% received full Senate votes.[7] These figures represent a marked downturn from the two decades preceding the Clinton mixed-government era.[8] The reduction in hearings and votes became even more significant in the final two years of the Clinton administration: Of the thirty-four nominees to the courts of appeals, only fifteen received committee hearings, committee votes, or full Senate votes.[9] Democrats decried these Republican tactics for being obstructionist and for exacerbating a judicial vacancy "crisis."[10]

George W. Bush's election in 2000 began with five months of a Republican-controlled Congress. But with the Senate divided equally among Republicans and Democrats,[11] the defection of Senator James Jeffords of Vermont from the Republican Party effectively returned Senate control to the Democrats. This change to mixed government quickly led to greater numbers of unconfirmed judicial nominations comparable to those of the last two years of the Clinton administration. In 2001 and 2002, President Bush nominated sixty-one candidates to positions on the courts of appeals. Of those, twenty-four received committee

---

5. DENIS STEVEN RUTKUS & MITCHEL A. SOLLENBERGER, CONG. RESEARCH SERV., ORDER CODE RL31635, JUDICIAL NOMINATION STATISTICS: U.S. DISTRICT AND CIRCUIT COURTS, 1977–2003, at 30 (updated Feb. 23, 2004).

6. *Id.*

7. *Id.*

8. *Id.*

9. *Id.*

10. *See, e.g.*, Patrick J. Leahy, Op-Ed., *Where Are the Judges?*, BOSTON GLOBE, Mar. 1, 1998, at C7 ("Any week in which the Senate does not confirm three judges is a week in which it is failing to address the vacancy crisis. . . . A Republican leader in the House speaks openly about seeking to 'intimidate' federal judges. The confirmation process is not immune to politics, but a particularly virulent strain has infected the Senate's consideration of judicial nominations over the last few years.").

11. The tie-breaking vote of Vice President Cheney gave the Republicans a 51-50 majority.

hearings, nineteen received committee votes, and seventeen received full Senate votes.[12]

The Republicans regained control of the Senate in the 2002 election, but the slim majority left the Democrats with sufficient votes to block certain nominations. Throughout 2003, Democrats focused on six nominations to the courts of appeals and repeatedly prevented successful cloture motions on those nominations. This impasse culminated in a days-long filibuster broadcast nationwide on cable television.[13] Some Republicans suggested using the "constitutional option," alternatively known as the "nuclear option," a parliamentary maneuver capable of circumventing the filibuster.[14] Two factors contributed to the tactic's falling out of the national spotlight: First, President Bush began to use recess appointments temporarily to appoint filibustered nominees to the courts of appeals without a Senate vote;[15] second, a bipartisan group of fourteen Senators agreed to work together to limit the ability of Republicans to invoke the constitutional option, and in exchange, of Democrats to successfully filibuster a nomination.[16] Throughout this standoff, Republicans decried the Democrats' tactics as obstructionist and argued that they exacerbated the "crisis" of judicial vacancies.[17]

---

12. RUTKUS & SOLLENBERGER, *supra* note 5, at 30.

13. *See* Helen Dewar, *Bush, Daschle Trade Charges: Senate's Talkathon on Judicial Nominees Exceeds 30 Hours*, WASH. POST, Nov. 14, 2003, at A6; Helen Dewar, *Senate Filibuster Ends with Talk of Next Stage in Fight*, WASH. POST, Nov. 15, 2003, at A9.

14. *See* Helen Dewar & Mike Allen, *Frist Seeks to End Nominees Impasse: Majority Leader's Plan Would Limit Tactics Used in Senate to Block Judicial Picks*, WASH. POST, May 9, 2003, at A12; *see generally* Martin B. Gold & Dimple Gupta, *The Constitutional Option to Change Senate Rules and Procedures: A Majoritarian Means to Over Come the Filibuster*, 28 HARV. J.L. & PUB. POL'Y 205 (2004).

15. *See* Richard Simon, *Bush, Frustrated by Democrats, Again Bypasses Senate on Judge*, L.A. TIMES, Feb. 21, 2004, at A1; *cf.* Randy Barnett, *Benching Bork*, NAT'L REV. ONLINE, Apr. 29, 2003, http://www.nationalreview.com/comment/comment-barnett042903.asp (recommending that President Bush engage in intentionally controversial recess appointments in order to persuade Democrats to end their filibustering of judicial nominations).

16. *See* Charles Babington & Shailagh Murray, *A Last-Minute Deal on Judicial Nominees*, WASH. POST, May 24, 2005, at A1.

17. *See* Senator Orrin Hatch, Press Conference at Room S-207, The United States Capitol (Apr. 12, 2002) ("We're in the middle of a circuit court of appeals crisis in this country. It's a constitutional crisis . . . . Now the Senate Democrats are trying to create an illusion of movement by creating great media attention concerning a small handful of nominees, in order to make it look like progress, and D. Brooks

President Bush's reelection in 2004 reinvigorated conservative efforts to confirm judicial nominees. In his first post-election announcement on the matter, President Bush challenged the Senate in no uncertain terms: "The Senate has a constitutional obligation to vote up or down on a President's judicial nominees."[18]

But does the Senate actually have such a constitutional *obligation*? The text of the Constitution certainly does not include such a provision; nevertheless, some current Senators from both parties have agreed—both recently and historically—with President Bush's sentiments.[19] Professor Douglas Kmiec, among others, has argued that the Senate would contravene its constitutional duty by not bringing a nomination to the floor after committee consideration.[20] If these government officials and commentators are correct, then the Senate has acted in dereliction of its constitutional duty on a great number of occasions. In addition to its failure to vote on numerous circuit and district court nominations, the Senate has also failed to vote up or down on nominees to the Supreme Court. On 154 occasions the President has nominated someone for appointment to the Supreme Court. Thirty-four of those nominees have not received senatorial approval.[21] Of the thirty-four failed nominations, only eleven were rejected by a full vote of the Senate;[22] the remainder were either withdrawn by the President without Senate action[23] or defeated in the Senate by means other than an up-or-down vote by the full Senate.[24]

Smith appears to be the next one who's going to be up for a smearing.") (transcript available through Federal News Service).

18. Press Release, President George W. Bush, Statement on Judicial Nominations (Dec. 23, 2004), *available at* http://www.whitehouse.gov/news/releases/2004/12/20041223-1.html.

19. *See* John Cornyn, *Our Broken Judicial Confirmation Process and the Need for Filibuster Reform*, 27 HARV. J.L. & PUB. POL'Y 197–99 (2003) ("If the Constitution provides that only a majority is necessary to confirm judges, any Senate rule that purports to prevent a majority of the Senate from exercising that confirmation function directly contradicts and offends our constitutional design."); *see also id.* at 206–11 (citing statements of numerous Senators expressing support for up-or-down votes on judicial nominations).

20. Douglas W. Kmiec & Elliot Mincberg, *The Role of the Senate in Judicial Confirmations*, 7 TEX. REV. L. & POL. 235, 262 (2003).

21. HOGUE, *supra* note 4, at 1.

22. *Id.* at 2.

23. William Paterson, nominated by President Washington, was withdrawn February 28, 1793, one day after his nomination. *Id.* at 18. President Washington

Of course, historical persistence in a certain course of conduct for an extended time does not render that conduct, ipso facto, compliant with the Constitution.[25] A determination of whether such a constitutional obligation exists must begin with the constitutional text.

Nomination and appointment of federal judges is governed by Article II, Section 2, Clause 2 of the U.S. Constitution: "[The President,] by and with the Advice and Consent of the Senate, shall appoint Ambassadors, other public Ministers and Consuls, Judges of the supreme Court, and all other Officers of the United States, whose Appointments are not herein otherwise provided for, and which shall be established by Law . . . ." Although the Constitution lacks an explicit requirement that the Senate act on all nominations, such absence is not necessarily dispositive. If an implicit requirement exists, however, it should be supported by the original meaning of "advice and consent."

The meaning of advice and consent is not self-evident, and the means of its proper application are not obvious. In attempting to ascertain the original meaning of this constitutional provision, we must first look to the debates at the Constitutional Convention and in the subsequent state ratification conventions.

Although the Framers' understanding of a term is not necessarily synonymous with the objective meaning of that term, their understanding can serve as a useful step toward the goal of discerning the original meaning of the term.[26] This Article

---

withdrew his nomination because Washington discovered that four days remained on then-Senator Paterson's term. *Id.* at 2–3. He was confirmed on resubmittal. *Id.* John C. Spencer, nominated by President Tyler, was withdrawn June 17, 1844, the same day as his nomination. *Id.* at 18.

24. *See id.* at 2.

25. *Cf.* E. Enters. v. Apfel, 524 U.S. 498, 538–39 (Thomas, J., concurring) ("In an appropriate case, therefore, I would be willing to reconsider [Calder v. Bull, 3 U.S. (3 Dall.) 386 (1798)] and its progeny to determine whether a retroactive civil law that passes muster under our current Takings Clause jurisprudence is nonetheless unconstitutional under the *Ex Post Facto* Clause.").

26. Justice Scalia is particularly succinct in illustrating the difference between "original meaning" and "Framer intent" or "original understanding": "What I look for in the Constitution is precisely what I look for in a statute: the original meaning of the text, not what the original draftsmen intended." ANTONIN SCALIA, A MATTER OF INTERPRETATION 38 (1997). One year before publishing those words, he explained the difference more fully in a speech at The Catholic University of America:

does not attempt to authoritatively define the original meaning of advice and consent. Rather, its narrow aim is to further the understanding of what the Framers believed "advice and consent" meant by looking at evidence long unappreciated by scholars of constitutional history.

The first unappreciated consideration is the actual practice of advice and consent in the Commonwealth of Massachusetts, the example after which constitutional advice and consent was explicitly created. The second unappreciated consideration is James Madison's alternative to advice and consent: a majority veto of presidential nominations by the Senate, to be used at its discretion.

Part I of this Article traces the framing of the Constitution's provision for Senate advice and consent of judicial appointments from the Constitutional Convention through the ratification debates. Part II explores the history of advice and consent in Massachusetts, outlining the historical relationship between the Governor and the Executive Council, and detailing long-ignored evidence contained in the official eighteenth-century records of the council. Part III focuses on Madison's alternative to advice and consent—a discretionary Senate veto—and offers a framework for analyzing the textual differences between the veto and advice and consent. Marshalling these findings, this Article concludes that this evidence does not support an assertion that the constitutional provision for advice and consent contains an implicit obligation to act on the President's nominations.[27]

> The theory of originalism treats a constitution like a statute, and gives it the meaning that its words were understood to bear at the time they were promulgated. You will sometimes hear it described as the theory of original intent. You will never hear me refer to original intent, because as I say I am first of all a textualist, and secondly an originalist. If you are a textualist, you don't care about the intent, and I don't care if the framers of the Constitution had some secret meaning in mind when they adopted its words. I take the words as they were promulgated to the people of the United States, and what is the fairly understood meaning of those words.

Associate Justice Antonin Scalia, Address at the Catholic University of America (Oct. 18, 1996), *available at* http://www.courttv.com/archive/legaldocs/rights/scalia.html.

27. This analysis focuses on the duties and obligations of the Senate as a singular body. As Gold and Gupta so thoroughly demonstrated in their recent article in this Journal's pages, *see* Gold & Gupta, *supra* note 14, the will of the Senate—by its own rules and precedents—can be effectively determined by a simple majority vote of its members. For that reason, this Article should in no way

I. THE FRAMING AND RATIFICATION OF THE ADVICE AND CONSENT CLAUSE

The constitutional provision for Senate advice and consent of presidential nominations for judicial appointments was developed in the debates at the Constitutional Convention of 1787. During the debates, a variety of other modes of appointment—including presidential appointment, senatorial appointment, and presidential appointment subject to a discretionary Senate veto—were discussed and sometimes initially accepted. Eventually, all were rejected. According to Madison's record of the debates, the advice and consent model was initially presented with reference to the Massachusetts model of executive appointment, in which the Governor appointed officers subject to the advice and consent of an executive council. Upon the conclusion of the Convention, discussion of Senate involvement in executive appointments continued, albeit briefly, in the *Federalist Papers* and in state ratification debates.

A. The Constitutional Convention

At the 1787 Federal Constitutional Convention, the delegates were divided—often sharply—over the proper allocation of power among the various branches of federal government and over the mode of selecting officeholders. The appointment of judges was specifically discussed in the course of these deliberations, but these discussions were interspersed in the discussion of the larger issues of the powers and election of the President and Senate. As such, the debate over appointment of judges was shaped and directed by these larger structural debates.

1. Appointment of Judges and Other Officers

Toward the conclusion of the Convention, the delegates were clearly divided on the question of appointments. One group of delegates, led by James Wilson, Nathaniel Gorham, Alexander Hamilton, and Gouverneur Morris, favored control of appointments by a strong executive. The opposing camp, led by

---

be taken as an endorsement of the efforts of vocal members of the Senate's current minority party to dictate the agenda of the full body through the threatened use of filibusters. That is an intrabranch debate, not an interbranch debate relevant to the Advice and Consent Clause.

Charles Pinckney, Luther Martin, George Mason, Roger Sherman, Oliver Ellsworth, and John Rutledge, favored legislative control of the appointment process.

a. May 29: Appointment by the Legislature

The power to appoint judges initially was assigned to the legislature. This arrangement was proposed on May 29 by Edmund Randolph in what came to be known as the Virginia Plan. In his fifteen-point plan for a national government, Randolph recommended that "a national judiciary be established . . . to be chosen by the national legislature . . . ."[28] The Virginia Plan, which also provided for a bicameral legislature[29] and a national executive chosen by the legislature,[30] initially contained no provision for the appointment of nonjudicial officers.[31] In a later session, however, the delegates voted overwhelmingly, and with only minimal debate, to allocate the power of officer appointment to the Executive.[32]

On June 5, the Committee of the Whole considered the Virginia Plan's method of judicial appointment. That Plan's first critic was James Wilson of Pennsylvania, who supported executive control over judicial appointments on the ground that "[e]xperience shewed the impropriety of such appointmts. by numerous bodies. Intrigue, partiality, and concealment were the necessary consequences."[33] Wilson's call for "unity in the executive" would continue throughout the debates.

Benjamin Franklin suggested that federal judges might be chosen, in the Scottish tradition, by lawyers.[34] This suggestion was not extensively considered in the course of the convention.

---

28. 1 THE RECORDS OF THE FEDERAL CONVENTION OF 1787, at 21 (Max Farrand ed., rev. ed. 1937) [hereinafter CONVENTION RECORDS].

29. *Id.* at 20.

30. *Id.* at 21.

31. *See id.* at 20–23. The Virginia Plan did make reference to the impeachment of "any National officers." *Id.* at 22.

32. *See id.* at 63. The States, voting on June 1, supported the proposal nine-to-zero. *Id.* at 67. Connecticut's delegates were divided and thus did not vote. *Id.*

33. *See id.* at 119.

34. *Id.* at 119–20. According to Franklin, who related the Scottish tradition "in a brief and entertaining manner," lawyers would tend to select the best among them for a judgeship "in order to get rid of him, and share his practice among themselves." *Id.* at 120.

James Madison, a critic of appointment by the legislature, recommended appointment by the Senate alone. Those of "Legislative talents" would, in "intrigue and partiality," support candidates on the basis of membership in the legislature.[35] But Madison also disagreed with Wilson's suggestion of appointment by the Executive—although he provided little substantive argument against such appointment. Madison favored appointment by the Senate as a matter of pragmatic compromise: The Senate was "numerous eno' to be confided in—as not so numerous as to be governed by the motives of the [House of Representatives]; and as being sufficiently stable and independent to follow their deliberate judgments."[36] Madison did not submit his proposal for a vote; rather, he submitted that "*appointment by the Legislature*" be struck from the Plan and replaced by blank space, to be filled upon "maturer reflection."[37] The motion passed, nine states to two.[38]

b. June 13: Appointment by the Senate

The issue of judicial appointments was reconsidered on June 13, when the delegates reconvened to fill the blank left by the vote of June 5. Madison resumed his argument in favor of appointment of judges by the Senate. He argued that appointment by the whole legislature would be biased in favor of nominees who had served in the legislature and who may have provided fellow legislators with political favors.[39] Madison supported appointment by the Senate, "a less numerous & more select body" that would be a better judge of a candidate's qualifications and that would be "sufficiently numerous to justify such a confidence in them."[40] Madison's argument was successful; Roger Sherman of Connecticut and Charles Pinckney[41] of South Carolina withdrew their motion for appointment

35. *See id.* at 120.

36. *Id.*

37. *Id.*

38. *See id.*

39. *See id.* at 232–33.

40. *Id.* at 233.

41. Throughout his account, Madison incorrectly spelled the delegate's name "Pinkney." The correct spelling is "Pinckney." *See* Library of Congress, Documents from the Continental Congress and the Constitutional Convention, 1774–1789, Charles Pinckney (1757–1824), http://memory.loc.gov/ammem/collections/continental/pinckney.html (last visited Nov. 20, 2005).

by the legislature, and senatorial appointment was agreed upon without opposition.[42]

On June 15, William Paterson[43] of New Jersey presented the "New Jersey Plan," a structural alternative to the Virginia Plan that was much more favorable to the interests of small states.[44] Among its provisions was an allocation of the power to appoint judges to the Executive,[45] who himself was selected by the then-unicameral legislature.[46]

In the following days the delegates discussed the relative merits of the New Jersey and Virginia Plans. On June 18, in response to the New Jersey Plan, Alexander Hamilton presented his own plan of government, under which the Executive would be empowered "to have the sole appointment of the heads or chief officers of the departments of Finance, War and Foreign Affairs; to have the nomination of all other officers (Ambassadors to foreign Nations included) subject to the approbation or rejection of the Senate."[47] Hamilton's plan did not explicitly refer to the appointment of judges, but "all other officers" appears to include judges.[48] Hamilton's choice of language describing the relationship between the Executive and Senate on appointments—to wit, "subject to the *approbation or rejection* of the Senate"—is noteworthy: It represents the first usage of language comparable to the advice and consent clause ultimately adopted by the Convention.[49] The Hamilton Plan was not put to a vote; the convention was adjourned for the day following Hamilton's speech.

On June 19, without debate on the specific issue, the delegates reaffirmed their earlier support of Senate appointment of

---

42. *See* 1 CONVENTION RECORDS, *supra* note 28, at 233.

43. Throughout his account, Madison incorrectly spelled the delegate's name "Patterson." The correct spelling is "Paterson." *See* William Paterson University, Biography of William Paterson, Patriot and Statesman 1745–1806, http://www.wpunj.edu/aboutus/wpbio.htm (last visited Oct. 19, 2005).

44. *See* 1 CONVENTION RECORDS, *supra* note 28, at 242–45.

45. *Id.* at 244.

46. *See id.* The New Jersey Plan's legislature was the same as that under the Articles of Confederation: It was to consist of equal representation among the States (like the eventual U.S. Senate). *See id.* at 286.

47. *Id.* at 292.

48. *See id.*

49. *See id.* (emphasis added). Hamilton also noted the "advice and approbation of the Senate" relationship with respect to the President's treaty power. *See id.*

judicial nominees,[50] the Paterson and Hamilton proposals notwithstanding. The delegates' resolution also provided that the Executive would appoint all other officers unless otherwise provided by law.[51]

c. July 18: Gorham Proposes Advice and Consent

The question of judicial appointment lay dormant for a month following the June 19 vote. In that month, the delegates enacted major changes to the mode of election of the President and the Senate.[52] On July 18, the delegates returned to the matter for extended debate on a day that would prove to be pivotal in the delegates' consideration of the question of judicial appointments. Nathaniel Gorham[53] of Massachusetts initiated the debate by urging an alternative to Senate appointment of judges:

> Mr. Ghorum, wd. prefer an appointment by the 2d branch to an appointmt. by the whole Legislature; but he thought even that branch too numerous, and too little personally responsible, to ensure a good choice. He suggested that the Judges be appointed by the Execuve. with the advice & consent of the 2d branch, in the mode prescribed by the constitution of Masts. This mode had been long practised in that country, & was found to answer perfectly well.[54]

Gorham argued that Senate appointment would be inefficacious because Senators would be partial to nominees hailing from their own home States; the President, on the other hand, would "be careful to look through all the states for proper characters" in making appointments.[55]

---

50. *See id.* at 322.

51. *See id.*; *see also id.* at 233.

52. *See infra* Part I.A.2.

53. Throughout his account, Madison misspelled the delegate's name as "Ghorum." The correct spelling is "Gorham." *See* John O. McGinnis, *The President, the Senate, the Constitution, and the Confirmation Process: A Reply to Professors Strauss and Sunstein*, 71 TEX. L. REV. 633, 640 n.29 (1993) (citing 7 DICTIONARY OF AMERICAN BIOGRAPHY 433 (Dumas Malone ed., 1931)).

Gorham was personally acquainted with the advice and consent procedures of the Massachusetts Constitution. At least once, on March 19, 1781, the Governor nominated him to appointment (in this case, to the office of Justice of the Peace & Quorum), and his nomination was confirmed by the Council on March 27, 1781. *See* 26 COUNCIL RECORDS 202 (March 19, 1781); *see id.* at 221 (March 27, 1781).

54. 2 CONVENTION RECORDS, *supra* note 28, at 41.

55. *Id.* at 42.

Gorham's suggestion sparked a renewed debate among the now-familiar advocates of senatorial and presidential control. Wilson warmed to it as a matter of political pragmatism: Although he still preferred executive appointment of judges, he would accept Gorham's proposal as a second-best option.[56] Luther Martin of Maryland supported Senate appointment, as the Senate "would be best informed of characters & most capable of making a fit choice."[57]

Sherman, who had supported legislative appointment on June 13, eventually returned to support of Senate appointment.[58] He argued that the collected Senators would have "more wisdom" than the President alone; they would bring "a more diffusive knowledge of characters," and they would be less likely to suffer the effects of intrigue than the solitary President.[59] Sherman did concede, however, that the advice and consent model was preferable to pure executive appointment.[60] Randolph, originally a proponent of vesting the appointment power in the whole legislature, concurred with Sherman's support of senatorial selection,[61] and recommended that the confirmation votes of each Senator be recorded in the Senate's Journal.[62]

In response to Gorham's advice and consent proposal, Madison proposed a noteworthy alternative. Although his proposal was, on its face, similar to Gorham's, Madison specified the level of Senate support required for successful approval of the nomination:

> Mr. M<adison>, suggested that the Judges might be appointed by the Executives with the concurrence of <⅓ at least>, of the 2d. branch. This would unite the advantage of responsibility in the Executive with the security afforded in

56. *See id.* at 41.

57. *Id.*

58. *Id.*

59. *Id.* at 43.

60. *Id.* at 44.

61. While Randolph's support for Senate appointment of judges was unqualified on July 18, he abruptly reversed position, without explanation, mere days later. *See infra* notes 74–77 and accompanying text.

62. 2 CONVENTION RECORDS, *supra* note 28, at 43.

> the 2d. branch agst. any incautious or corrupt nomination by the Executive.[63]

At this point in the debate, therefore, Madison was supporting a position to the contrary of his Senate appointment model, the arrangement at that time endorsed by the Convention.

Following the debate over the merits of the Gorham proposal, the delegates proceeded to vote. First, they voted on vesting the appointment power solely in the Executive. The motion failed, two states to six.[64] Next, they voted on Gorham's advice and consent proposal as seconded by Gouverneur Morris.[65] The motion failed, with the delegations deadlocked, four states to four.[66] Two states that had voted against appointment by the Executive alone, Maryland and Virginia, voted in favor of the advice and consent plan.

Upon the failure of the Gorham motion, Madison proposed a variation of his earlier plan. Under his revised proposal, the Executive did *not* need Senate *concurrence* for his appointment; rather, he merely needed the Senate to *decline* to *oppose* explicitly the nomination: "<Mr.> M<adison> moved that the Judges should be nominated by the Executive, & such nomination should become an appointment <if not> disagreed to within [blank] days by ⅔ of the 2d. branch."[67] Gouverneur Morris again seconded this motion, but it was postponed without a vote.[68]

On July 21 the delegates resumed the debate, with Madison offering a more comprehensive justification for his Senate-veto plan: It would place primary responsibility on the President, who would be the most "capable & likely" to choose men of

---

63. *Id.* at 42–43.

64. *Id.* at 44. The motion was supported by Massachusetts and Pennsylvania. It was opposed by Connecticut, Delaware, Maryland, Virginia, North Carolina, and South Carolina. Georgia voted "absent." *Id.*

65. *Id.* ("Mr. Gorham moved 'that the Judges be <nominated and appointed> by the Executive, by & with the advice & consent of the 2d branch <& every such nomination shall be made at least [blank] days prior to such appointment'>. This mode he said had been ratified by the experience of 140 years in Massachustts. If the appt. should be left to either branch of the Legislature, it will be a mere piece of jobbing.").

66. *Id.* The motion was supported by Massachusetts, Pennsylvania, Maryland, and Virginia. It was opposed by Connecticut, Delaware, North Carolina, and South Carolina. Again, Georgia voted "absent." *Id.*

67. *Id.*

68. *Id.*

good character for the nomination. It would also be harder for the President to hide "selfish motives" behind collective action. To the extent that a President did choose men out of selfish motives, the Senate veto would provide a sufficient check.[69] Madison also pointed out that when the Senate was originally entrusted with the appointment power, its seats were apportioned among the States according to population; under later compromises, however, the Senate's seats were now apportioned equally among the States.[70] Entrusting the appointment power to the Senate would enable representatives of a national popular minority to control appointments; the President, "considered as a national officer, acting for and equally sympathising with every part of the U. States," would not be as likely to be dominated by a popular minority.[71]

The delegates responded with a variety of counterproposals. Pinckney persisted in his support of Senate appointment.[72] Elbridge Gerry of advice-and-consent Massachusetts agreed with Pinckney.[73] But Randolph—who previously had allied with Pinckney to support Senate appointment,[74] and who was the author of the Virginia Plan, which would have assigned appointment power to the national legislature as a whole[75]—supported the Madison proposal. He argued that he would have preferred Gorham's advice and consent model, but, in its absence, Madison's discretionary Senate veto was better than Senate appointment.[76] Echoing Madison's argument, he set forth his enthusiasm in no uncertain terms: He "thought the motion . . . so great an improvement of the clause as it stands, that he anxiously wished it success."[77] Morris also echoed

---

69. *Id.* at 80.

70. Equal representation of the States in the Senate was approved on July 16, by a vote of five states to four. *Id.* at 15.

71. *Id.* at 81.

72. *Id.*

73. *Id.* at 82.

74. *See supra* notes 61–62.

75. *See supra* note 28 and accompanying text.

76. 2 CONVENTION RECORDS, *supra* note 28, at 81. Randolph's explicit differentiation between Gorham's advice and consent model and Madison's Senate-veto model is particularly probative in this Article's analysis of the original meaning of advice and consent. *See infra* Part III.A.

77. 2 CONVENTION RECORDS, *supra* note 28, at 81. James Gauch attributes Randolph's abrupt about-face to belated appreciation of the advantages the Madison model would provide to such a large state as Virginia, whose power in

Madison's arguments.[78] Oliver Ellsworth[79] argued in favor of Senate nomination of appointments, subject to an executive veto, with the Senate empowered to override such a veto by a two-thirds vote; Ellsworth thought that the Executive, located in one place, could not be as knowledgeable of nominees as could the Senators of the several States.[80] George Mason of Virginia was thoroughly unimpressed by the differences between any of the plans of joint President-Senate appointment, arguing that "[n]otwithstanding the form of the proposition . . . the appointment was substantially vested in the [Executive] alone."[81] He saw executive appointment to be "a dangerous prerogative. It might even give him an influence over the Judiciary department itself."[82]

Elbridge Gerry of Massachusetts criticized Madison's plan on, among other things, the ground that the one-third Senate vote needed to prevent a Senate veto was too low. Madison's response was very revealing with respect to his understanding of the differences between his plan and Gorham's:

> Mr. <Madison>, observed that he was not anxious that ⅔ should be necessary to disagree to a nomination. He had given this form to his motion chiefly *to vary it the more clearly from one which had just been rejected*.[83] He was content to obviate the objection last made, and accordingly so varied the motion *as to let a majority reject*.[84]

---

the Senate was reduced by the equal-representation allocation of seats. James E. Gauch, Comment, *The Intended Role of the Senate in Supreme Court Appointments*, 56 U. CHI. L. REV. 337, 349–50 (1989). Gauch argues that the Convention's debate on appointments was driven in part by rivalry between large and small states. *See id.* at 340, 347–50.

78. 2 CONVENTION RECORDS, *supra* note 28, at 82.

79. Throughout his account, Madison misspelled the delegate's name "Elseworth." The correct spelling is "Ellsworth." *See, e.g.*, Members of the Supreme Court of the United States, http://www.supremecourtus.gov/about/members.pdf (last visited Nov. 20, 2005).

80. 2 CONVENTION RECORDS, *supra* note 28, at 81.

81. *Id.* at 83.

82. *Id.*

83. The full import of this distinction is discussed in greater detail further below. *See infra* Part III.A.

84. 2 CONVENTION RECORDS, *supra* note 28, at 82 (emphases added).

The delegates first put Madison's proposal to a vote. It was defeated, three states to six.[85] Maryland, which voted in favor of Gorham's advice and consent plan days earlier, voted against Madison's plan. Georgia, which had voted "absent" on both the Executive-only proposal and on the Gorham advice and consent proposal, also voted against Madison's Senate-veto plan. The delegates immediately voted to reaffirm senatorial appointment of judges. The motion carried, six states to three.[86]

On July 26, the delegates referred the matter, along with its progress on other issues, to the Committee of Detail.[87] Among the provisions sent to the committee was a provision for appointment of judges by the Senate.[88] The Executive retained full power for appointment of officers "in cases not otherwise provided for."[89] The Committee on Detail returned its report on August 6.[90] Its text provided for appointment of Supreme Court judges and ambassadors by the Senate,[91] and appointment of "officers in all cases not otherwise provided for by this Constitution" by the President.[92]

On August 23, Gouverneur Morris resumed his criticism of appointment of officers by the Senate. He relied primarily on the threat of cabal present in such a "numerous" body; Wilson concurred.[93] But their criticisms did not result in a formal motion and vote.

---

85. *Id.* at 83. Voting in favor: Massachusetts, Pennsylvania, and Virginia. Voting against: Connecticut, Delaware, Maryland, North Carolina, South Carolina, Georgia. *Id.*

86. *Id.* Voting in favor: Connecticut, Delaware, Maryland, North Carolina, South Carolina, and Georgia. Voting against: Massachusetts, Pennsylvania, and Virginia. *Id.*

87. *Id.* at 128.

88. *See id.* at 83. The July 21 resolution that judges be appointed by the Senate was sent to the Committee. *Id.* at 128.

89. *Id.* at 120–21. The July 26 resolution that the Executive appoint officers in cases "not otherwise provided for" was sent to the Committee. *Id.* at 128.

90. *Id.* at 177.

91. *Id.* at 183.

92. *Id.* at 185. The Committee on Detail's report was the first to bestow the title of "President" upon the national Executive. *Id.*

93. *Id.* at 389.

d. September 4: Advice and Consent

On September 4, the convention broke from its agreement on Senate appointment without explanation, resurrecting appointment by the President with the advice and consent of the Senate. David Brearley of New Jersey presented the report of the Committee on Compromise, which provided for appointment of judges, ambassadors, and other officers by the President and Senate together:

> The President by and with the advice and Consent of the Senate, shall have power to make Treaties; and he shall nominate and by and with the advice and consent of the Senate shall appoint ambassadors, and other public Ministers, Judges of the Supreme Court, and all other Officers of the U—S—, whose appointments are not otherwise herein provided for.[94]

In debate on the provision three days later, the delegates once more considered this commingling of Executive and Senate involvement. Wilson, who had accepted this model as a second-best option when Gorham first proposed it in July,[95] objected to the "blending [of] a branch of the Legislature with the Executive."[96] He proposed instead that the President's appointments be assisted by the nonbinding advice of a "privy council."[97] The involvement of a privy council was opposed by Rufus King of Massachusetts, a state that used such a council in its own appointments.[98] King supported the Senate's advice and consent role.[99] Morris continued in his support for the advice and consent plan.[100]

Pinckney, who initially had supported appointment by the legislature,[101] and who later supported Senate appointment,[102] now opposed Senate involvement in appointments by the

94. *Id.* at 498–99.

95. *See supra* notes 53–56 and accompanying text.

96. 2 CONVENTION RECORDS, *supra* note 28, at 538.

97. *Id.* at 539; *see also infra* Part I.A.2.b (discussing the proposal of a "privy council").

98. *See infra* Part II.

99. 2 CONVENTION RECORDS, *supra* note 28, at 539.

100. *Id.*

101. *See supra* Part I.A.1.b.

102. *See supra* Part I.A.1.c.

President. (He did, however, support appointment of ambassadors by the Senate alone.)[103]

The delegates approved the Committee on Compromise's language with only slight modification.[104] The motion, which had been opposed so strenuously in earlier votes, passed without opposition.[105] On September 12, the Committee of Style presented what would become the final version of the Appointments Clause,[106] which the delegates finally accepted on September 17.[107]

2. *Matters Related to Senate Advice and Consent*

a. *Election of and Representation in the Senate*

The idea of the Senate—a second house of the national legislature—was first proposed in Randolph's Virginia Plan, to serve as a compromise in the apportionment battles between large and small states.[108] As detailed on May 29, the Virginia Plan called for the second house of the bicameral national legislature to consist of members nominated by the legislatures of the several States and elected by the other house of the national legislature.[109] On June 7, Dickinson and Sherman proposed that members of the Senate be chosen by the individual state legislatures.[110] Madison was a vigorous opponent of this mode of election; he warned that such a plan would require the Convention to "depart from the doctrine of proportional representation; or admit into the Senate a very large number of members."[111] The former was rejected out of hand as "evidently

---

103. 2 CONVENTION RECORDS, *supra* note 28, at 539.

104. *Id.* (adding the parenthetical term "(and Consuls)" between "other public ministers" and "Judges of the Supreme Court").

105. *Id.*

106. *Cf. id.* at 599, *with* U.S. CONST. art. II, § 2, cl. 2. Three days later, following minimal debate, the Convention approved an additional clause allowing Congress to vest in the President, courts of law, and heads of departments the appointment of inferior officers. 2 CONVENTION RECORDS, *supra* note 28, at 627–28.

107. 2 CONVENTION RECORDS, *supra* note 28, at 648–49.

108. *See* JACK N. RAKOVE, ORIGINAL MEANINGS: POLITICS AND IDEAS IN THE MAKING OF THE CONSTITUTION 59–65 (1996).

109. 1 CONVENTION RECORDS, *supra* note 28, at 20.

110. *Id.* at 150.

111. *Id.* at 151. Election by the State legislatures was assumed to require an enlargement of the number of Senators because, at this point in the debate, no one had suggested that the States should be represented equally. Under a system of

unjust." But his critique of the latter point, which he deemed "inexpedient," revealed his vision for the second branch of the Legislature:

> The use of the Senate is to consist in its proceeding with more coolness, with more system, & with more wisdom, than the popular branch. Enlarge their number and you communicate to them the vices which they are meant to correct.[112]

The arguments of Madison and others notwithstanding, the delegates approved election of Senators by state legislatures unanimously.[113]

The apportionment of seats in the Senate was a contentious point. On June 11, the closely divided committee of the whole voted to apportion votes in the Senate proportionally among the States according to their respective populations, the same manner in which suffrage was to be based in the House of Representatives.[114]

Mere days later, the New Jersey Plan was introduced; it proposed a unicameral legislature wherein each state would have one vote.[115] This proposal incited a number of proponents of the Virginia Plan to present critiques of Paterson's plan. Hamilton denounced the unicameral legislature of the New Jersey plan in his extensive remarks on June 18.[116] The following day, Madison himself extensively criticized the New Jersey Plan.[117] Following his remarks, the New Jersey Plan was rejected, seven states to three, with Maryland divided.[118]

Following the defeat of the New Jersey plan, the delegates offered a variety of compromise proposals for representation in

---

proportional representation, giving small states a Senator would mean that each Senator could only represent a population as big as the smallest state. Thus, as Pinckney warned, "If the small States should be allowed one Senator only, the number will be too great, there will be 80 at least." *Id.* at 150.

112. *Id.* at 151.

113. *Id.* at 156.

114. *Id.* at 202.

115. *Id.* at 243. The New Jersey Plan would have retained the unicameral, equal representation of the Congress of the Articles of Confederation.

116. *Id.* at 286–87.

117. *Id.* at 314–22.

118. *Id.* at 322. Voting in favor: Massachusetts, Connecticut, Pennsylvania, Virginia, North Carolina, South Carolina and Georgia. Voting against: New York, New Jersey, and Delaware. *Id.*

the two houses of the legislature. On June 28, John Lansing of New York and Jonathan Dayton of New Jersey proposed that the States enjoy equal representation in the House of Representatives.[119] The motion was defeated the next day, four states to six, with Maryland divided.[120] The same day, Ellsworth proposed that the States enjoy equal representation in the Senate.[121] It was defeated by a tie vote on July 2.[122] Having reached apparent deadlock, the delegates—over strenuous objection by Madison—assigned the issue to a Committee on Compromise.[123]

In committee, it was Benjamin Franklin who brokered the compromise: in the House the States would be represented in proportion to their population; in the Senate the States would be represented equally. This compromise "was barely acquiesced in by the members from the states opposed to an equality of votes in the 2d. branch and was evidently considered by the members on the other side, as a gaining on their point."[124] From July 5 to July 16, the delegates debated the compromise proposal generally. On July 16, they voted in favor of equal representation of States in the Senate.[125]

As intimated above, the July 16 change in Senate representation—from a proportional-representation scheme to one of equal representation—caused change in the Convention's allocation of appointment power.[126] On July 21, Madison noted the change in his defense of his Senate-veto plan for appointments, which would (according to him) prevent control of the ap-

---

119. *Id.* at 445. Lansing and Dayton proposed that representation in the House of Representatives be "according to the rule established by the Confederation," which granted one vote to each state. *Id.*

120. *Id.* at 468. Voting in favor: Connecticut, New York, New Jersey, and Delaware. Voting against: Massachusetts, Pennsylvania, Virginia, North Carolina, South Carolina, and Georgia. *Id.*

121. *Id.*

122. *Id.* at 510. Voting in favor: Connecticut, New York, New Jersey, Delaware, and Maryland. Voting against: Massachusetts, Pennsylvania, Virginia, North Carolina, and South Carolina. Georgia's vote was divided. *Id.*

123. *Id.* at 515–16.

124. 1 CONVENTION RECORDS, *supra* note 28, at 526 n*.

125. 2 CONVENTION RECORDS, *supra* note 28, at 15. Voting in favor: Connecticut, New Jersey, Delaware, Maryland, and North Carolina. Voting against: Pennsylvania, Virginia, South Carolina, and Georgia. Massachusetts's vote was divided. *Id.*

126. *See supra* Part I.A.1.c.

pointments by representatives of a national popular minority. In light of this change, he advocated entrusting primary responsibility for appointments in the President, a national officer, with a discretionary veto held by the Senate.

b. The President and an Executive Council

Although the Senate eventually came to be seen as an advisory body to the President in appointments and other matters, some participants in the Constitutional Convention urged the creation of bodies within the executive branch for the purpose of advising and counterbalancing the singular Executive. Such a council was no mere abstraction to the delegates: In the twelve years between the Declaration of Independence and the Constitutional Convention, at least nine states utilized an executive council in the making of appointments.[127]

The first delegate to propose an executive advisory council was Elbridge Gerry of Massachusetts, a state that long had utilized such a council. On June 1, Gerry suggested that a council to the Executive would "give weight & inspire confidence."[128]

Gerry's proposal came amidst a larger discussion of whether the national Executive should be a unitary office or a multiperson body. The major proponent of a multiperson Executive was Randolph, who regarded a unitary Executive as "the fœtus of monarchy."[129] The question of the form of the Executive was set aside until June 4. Wilson favored the unitary Executive, citing the experience of the several States.[130] Sherman noted that all thirteen states featured not only a unitary executive, but also "a Council of advice, without which the first magistrate could not act."[131] Wilson immediately disclaimed any appearance that he supported such a council.[132] Gerry, an apparent supporter of

---

127. SYDNEY GEORGE FISHER, THE EVOLUTION OF THE CONSTITUTION OF THE UNITED STATES 172–73 (1910) (naming South Carolina, Virginia, New Jersey, Delaware, Pennsylvania, Maryland, Vermont, Massachusetts, and New Hampshire). In addition, New York utilized a council selected out of the state senate. *Id.* at 173. Alexander Hamilton cited New York's appointment process as an example of a dangerous cabal. THE FEDERALIST NO. 77, at 461 (Alexander Hamilton) (Clinton Rossiter ed., 1961).

128. 1 CONVENTION RECORDS, *supra* note 28, at 66.

129. *Id.*

130. *Id.* at 96.

131. *Id.* at 97.

132. *Id.*

such a council as a check on the Executive, then specifically disclaimed the notion of a multiperson Executive. The proposed unitary Executive was accepted by the convention, seven states to three.[133]

The question of whether the unitary Executive should be advised by a privy council was again raised on August 18.[134] Ellsworth suggested that such a council consist of the "President of the Senate—the Chief-Justice, and the Ministers as they might be estab[lished] for the departments of foreign & domestic affairs, war finance, and marine, who should advise but not conclude the President."[135] Pinckney warned that an executive council would be an instrument of pure presidential discretion, for him to utilize as he saw fit: "Give him an able Council and it will thwart him; a weak one and he will shelter himself under their sanction."[136]

On August 20, Pinckney and Gouverneur Morris proposed a "Council of State" consisting of the Chief Justice and the heads of various departments.[137] The President would be free to submit matters for their discussion, and to require the written opinions of any of them,[138] but he was not to be bound by their views. This proposal was submitted to the Committee on Detail, which, on August 22, proposed that the President should have a privy council composed in a manner comparable to the Ellsworth and Morris proposals.[139] The proposal did not go to vote.

George Mason persisted in his support for a privy council as the delegates debated the Compromise Committee's proposal that appointments be made by the President with the advice and consent of the Senate.[140] Opposed to vesting appointment power in the Senate, but also wary of vesting such absolute

---

133. *Id.* at 97. Voting in favor: Massachusetts, Connecticut, Pennsylvania, Virginia, North Carolina, South Carolina, and Georgia. Voting against: Delaware, Maryland, and New York. *Id.*

134. On August 18, the Convention still had assigned the power of appointment to the Senate.

135. 2 CONVENTION RECORDS, *supra* note 28, at 329.

136. *Id.*

137. *Id.* at 342–43.

138. *Cf.* U.S. CONST. art. II, § 2, cl. 1 ("The President . . . may require the Opinion, in writing, of the principal Officer in each of the executive Departments . . . .").

139. 2 CONVENTION RECORDS, *supra* note 28, at 367–68.

140. *See supra* Part I.A.1.d.

power in the Executive, he proposed a six-member privy council to assist the President in making appointments.[141] Wilson agreed with Mason, as he saw Senate advice and consent as representing inappropriate blending of the executive and legislative branches.[142] But Rufus King of council-experienced Massachusetts opposed the creation of such a council, arguing that Senate involvement would be no more costly than council involvement and would require no "unnecessary creation of New Corps which must increase the expence as well as influence of the Government."[143] The delegates then voted in favor of appointments by the President with the advice and consent of the Senate.[144]

Mason tried once more to persuade the delegates to create a council, warning that a President lacking such restraint would be "an experiment on which the most despotic Governments had never ventured."[145] Franklin agreed, suggesting that a council "would not only be a check on a bad President but be a relief to a good one."[146] Gouverneur Morris argued that the council had been considered and rejected by the Council on Compromise in order to focus responsibility upon the President. Wilson, Dickinson, and even Madison voiced support for the council, but Mason's motion was defeated, eight states to three.[147] The Convention ended days later, leaving the Senate alone to serve as the advisory council for the President on appointments and treaties.

B. Post-Convention Debates

Upon the close of the convention, discussion of the proposed Constitution continued throughout the thirteen states. This dis-

---

141. 2 CONVENTION RECORDS, *supra* note 28, at 537–38. But not in the making of treaties or the appointment of ambassadors, which he deemed to be a "legislative" endeavor. *Id.* at 537.

142. *Id.* at 538–39.

143. *Id.* at 539.

144. *Id.* at 539–40.

145. *Id.* at 541.

146. *Id.* at 542.

147. *Id.* Voting in favor: Maryland, South Carolina, and Georgia. Voting against: New Hampshire, Massachusetts, New Jersey, Pennsylvania, Delaware, Virginia, Connecticut, and North Carolina. *Id.* Of the States voting in favor, Maryland and South Carolina used councils. *See* FISHER, *supra* note 127, at 172–73. Of the States voting against, New Hampshire, Massachusetts, New Jersey, Pennsylvania, Delaware, and Virginia used councils. *Id.*

cussion, which in a few instances touched on the issue of appointments, was conducted primarily in two ways: in written advocacy—both public (most notably, the *Federalist Papers*) and private—and in state ratification conventions.

1. The Federalist Papers

In three of his contributions to the famous series of essays to the People of the State of New York, Alexander Hamilton discussed the advice and consent model of appointment. In *Federalist Papers Nos. 66*, *76*, and *77*, he presented arguments that described the operation of the Senate in executive appointments. His arguments reflected in many ways the arguments regarding appointments presented by delegates at the Convention, although his arguments regarding appointments were not wholly reflective of the Convention debate or of the appointment process in practice.

Hamilton's *Federalist No. 66*, the "leading defense as well as the most thorough exposition of the provisions in the Constitution concerning the appointing power,"[148] gives a clear description of Hamilton's overall view of the advice and consent process. He did not expect the Senate to play any role in choosing the initial nominee:

> It will be the office of the President to *nominate*, and, with the advice and consent of the Senate, to *appoint*. There will, of course, be no exertion of *choice* on the part of the Senate. They may defeat one choice of the Executive, and oblige him to make another; but they cannot themselves *choose*—they can only ratify or reject the choice he may have made.[149]

Hamilton's defense of joining the President and the Senate in appointments echoed a major argument from the Convention: Vesting primary responsibility in the President would provide for responsibility and freedom from regional bias, while Senate consent would provide security from personal attachment. Hamilton's language implies that Senate rejection of the President's nominee would not be a regular occurrence: "But his nomination may be overruled: this it certainly may, yet it can only be to make place for another nomination by himself . . . . It

148. JOSEPH P. HARRIS, THE ADVICE AND CONSENT OF THE SENATE 27 (1953).
149. THE FEDERALIST NO. 66 (Alexander Hamilton), *supra* note 127, at 405.

is also not very probable that his nomination would often be overruled."[150]

Hamilton's *Federalist* remarks provide insight into *his* perception of the mechanics of advice and consent, but, on closer examination, they do not appear to be a reliable indicator of what the delegates endorsed in advice and consent. First, as noted above, Hamilton described the Senate's consent action in binary terms: "they can only *ratify or reject* the [Executive's] choice."[151] Later, he describes the Senate's "consent" power in "veto" terms: "their right of negative upon his nominations";[152] "[t]he censure of rejecting a good [nomination] . . . ."[153] This characterization—that the Senate approves or rejects—harkens back not so much to the Gorham advice and consent proposal, which provided for consent but left implicit the Senate's duty to affirmatively reject, but to Hamilton's own proposal of June 18, preceding the Gorham proposal, in which the President nominated officers "subject to the *approbation or rejection* of the Senate."[154] Indeed, Hamilton was not himself present at the Convention July 18 through 21 when Gorham proposed, described, and defended the advice and consent model.[155]

Hamilton's *Federalist* argument used the appointment process of New York—home of the *New York Packet*, an original publisher of the *Federalist* essays—as an example of a poor appointment process. New York relied on a Council of Appointments, composed of the Governor and four others.[156] Hamilton criticized the council's lack of public visibility—a "small body, shut up in a private apartment, impenetrable to the public eye"—as contributing to an environment of "cabal."[157]

Hamilton's argument that the Senate is the appropriate size for advising on appointments does not just focus on the advantages of the Senate relative to a smaller council; rather, he also

---

150. THE FEDERALIST NO. 76 (Alexander Hamilton), *supra* note 127, at 457.

151. *Id.* at 405 (emphasis added).

152. THE FEDERALIST NO. 77 (Alexander Hamilton), *supra* note 127, at 460.

153. *Id.* at 461.

154. 1 CONVENTION RECORDS, *supra* note 28, at 292 (emphasis added); *see supra* Part I.A.1.c.

155. RON CHERNOW, ALEXANDER HAMILTON 235, 238 (2004). Hamilton left the Convention after offering some remarks on June 29 and did not return until around August 13. *Id.*

156. THE FEDERALIST NO. 77 (Alexander Hamilton), *supra* note 127, at 461.

157. *Id.*

argues its advantages over the much larger legislature as a whole: "A body so fluctuating and at the same time so numerous can never be deemed proper for the exercise of that power."[158] Under the original Constitution, the Senate consisted of twenty-six members, the House sixty-five: in total, nine members smaller than the current Senate alone.[159]

In sum, a close reading of Hamilton's discussion provides much insight into his perception of the Senate's advice and consent role: The Senate could not exercise its own choice of nominee; the Senate would generally accept the President's nominee; the Senate would act in a transparent manner, preventing intrigue or cabal; the Senate would act in a majoritarian fashion; and the size of the 1789 Senate was pragmatically superior to a small council of five men or a larger Congress of ninety-one men. But most importantly, Hamilton clearly equated his "approbation or rejection" proposal with the mechanics of the final advice and consent proposal. Hamilton's arguments may be indicative of the opinions of one of the Federalist contingent's leaders, but they more plausibly represent not so much the appointments framework accepted by the Framers, but rather the appointments framework proposed by Hamilton but never voted upon at the Convention.[160]

2. State Ratification Debates

The appointment power did not receive extensive scrutiny in most of the state ratification conventions, as the Ratifiers focused more closely on the larger, more prominent issues of the federal government's structure and powers. Furthermore, in the course of the rare discussions actually conducted on the issue of appointments, speakers did not illuminate the Ratifiers' nuanced impressions of the specific roles of President and Senate so much as criticize the mix, variously favoring vesting absolute appointment power in either the President or the legislature.

An exception to this trend is found in the remarks of James Wilson before the Pennsylvania Convention. Although he ad-

158. *Id.* at 463.

159. *See* U.S. CONST. art. I, § 2, cl. 3.

160. *See generally* THE FEDERALIST NO. 77 (Alexander Hamilton), *supra* note 127, at 457. New York's Ratifiers, however, may have shared Hamilton's understanding, other states' Ratifiers notwithstanding.

mitted that "the senate was not a favorite of mine,"[161] he endeavored to defend its form. To those who feared that the Senate would exert too much power over the President in the appointment process, he argued that "the Senate stands controlled; if it is that monster which it said to be, it can only show its teeth, it is unable to bite or devour."[162] His remarks may or may not imply an obligation upon of the Senate to vote on nominations: "[T]he President must nominate before they can vote."[163] He clearly saw a vote involved at least in the process of consenting to a nomination. But he said "can vote," not "would vote," or "do vote." This description falls short of an explicit note of awareness of obligation.

Likewise, in North Carolina, delegate James Iredell spoke in terms that may or may not imply an affirmative obligation upon the Senate to act on the nomination: "The President proposes [a nominee]. The Senate has to consider upon it. If they think him improper, the President must nominate another . . . ."[164] According to Iredell, the Senate must "consider" it; but Iredell did not explicitly say whether "consideration" must culminate in a vote. Moreover, Iredell was not a party to the initial Convention. While his statements may have had an impact on the North Carolina ratification, they are of limited probative value beyond the extent that they indicate that the "reasonable North Carolina Ratifier" may have shared his impressions.

Beyond these meager pieces of evidence, little insight into the intended mechanics of Senate involvement in appointments can be gleaned from the state ratification debates.[165] The great majority of discussions on the matter of appointments focused

---

161. *James Wilson in the Pennsylvania Convention*, *in* PENNSYLVANIA AND THE FEDERALIST CONSTITUTION, 1787–1788 (John Bach McMaster & Frederick Stone eds., 1888), *reprinted in* 3 CONVENTION RECORDS, *supra* note 28, at 162. As shown in the discussion of the Federal Convention above, Wilson had been a staunch supporter of vesting appointment power in the President alone. *See supra* Part I.A.

162. *Id.*

163. *Id.*

164. Speech of James Iredell at the North Carolina Ratifying Convention, *in* 4 THE DEBATES IN THE SEVERAL STATE CONVENTIONS ON THE ADOPTION OF THE FEDERAL CONSTITUTION AS RECOMMENDED BY THE GENERAL CONVENTION AT PHILADELPHIA 134 (Jonathan Elliot ed., 2d ed. 1881) [hereinafter ELLIOT'S DEBATES].

165. *See generally* RATIFYING THE CONSTITUTION (Michael Allen Gillespie & Michael Lienesch eds., 1989).

on broader questions of policy surrounding the perceived "dangerous[] blending [of] the executive and legislative powers,"[166] on the threat of Senate control of the Executive through appointments,[167] or on fears that the Senate's check on the President in appointments will be chimerical, allowing the President, "when he pleases, to become a king in *name*, as well as in substance."[168]

3. Post-Ratification Debate

Following the ratification of the Constitution, John Adams, a man quite familiar with the Massachusetts Constitution, continued to debate the efficacy of the new Constitution with Roger Sherman. In a 1789 letter, Adams used language implying that he perceived a Senate duty to act on nominations: "The whole senate must now deliberate on every appointment . . . ."[169] Adams argued that the President should be advised on appointments by a privy council, not by the Senate, because, among other things, the burden of deliberating on every nomination for appointment would be intolerable. In his response, Sherman did not refute Adams's contention that the Senate would have to deliberate on every nomination.[170] But neither party suggested that "deliberation" would have to culminate with a formal vote.

II. THE MASSACHUSETTS EXPERIENCE

As noted above, when Nathaniel Gorham introduced his advice and consent model of executive appointments, he pointed to Massachusetts as his inspiration: "This mode had been long practised in that country, & was found to answer perfectly

---

166. George Mason, An Address Before the Virginia Ratifying Convention (1788), *in* PAMPHLETS ON THE CONSTITUTION OF THE UNITED STATES; PUBLISHED DURING ITS DISCUSSION BY THE PEOPLE, 1787–1788, at 330 (Paul L. Ford ed., 1888), *reprinted in* HARRIS, *supra* note 148, at 25.

167. 4 ELLIOT'S DEBATES, *supra* note 164, at 116–18, *quoted in* HARRIS, *supra* note 148, at 26 (argument of Samuel Spencer, delegate to the North Carolina Ratifying Convention).

168. Luther Martin, *Genuine Information*, MD. GAZETTE & BALT. ADVERTISER, Dec. 28, 1787, *reprinted in* 3 CONVENTION RECORDS, *supra* note 28, at 218.

169. John Adams, Third Letter to Roger Sherman on the Constitution of the United States, *in* 6 THE WORKS OF JOHN ADAMS 436 (Charles Francis ed., 1851).

170. *See* Roger Sherman, Reply to John Adams' Third Letter on the Constitution of the United States, *in* 6 THE WORKS OF JOHN ADAMS, *supra* note 169, at 440–42.

well."[171] That the delegates to the Federal Convention in Philadelphia used Massachusetts as their paradigm for advice and consent has hardly gone unnoticed by contemporary commentators.[172] But although commentators often have highlighted Gorham's reference to Massachusetts as a model for advice and consent, none have ventured to shed light on how advice and consent actually operated in Massachusetts in the years preceding the Convention. Investigation into records of executive appointments in 1780s Massachusetts demonstrates that, although the overwhelming majority of the Governor's nominations were approved by the constitutionally mandated advising Council, when nominations were not approved, no formal rejection of the nominees was announced. Moreover, the advising Council announced its decisions in a single voice; no records were kept of vote tallies in council decisionmaking.

The terms of the Massachusetts Constitution of 1780 obligated the Governor to secure the advice and consent of the Council, not just in the appointment of judges and officers, but also in the spending of treasury monies.[173] In other matters, the Governor had to secure the *advice* of the Council before acting, but he was not required to seek its *consent*.[174]

A. The History of the Council

Although the Council today plays only a limited role in Massachusetts's government,[175] its roots in Massachusetts's history stretch back long before the Constitution of 1780. Privy councils advised the Governor as early as the Seventeenth Century. The

---

171. 2 CONVENTION RECORDS, *supra* note 28, at 41; *see supra* Part I.A.1.c.

172. *See, e.g.*, Gauch, *supra* note 77, at 348; Matthew D. Marcotte, Note, *Advice and Consent: A Historical Argument for Substantive Senatorial Involvement in Judicial Nominations*, 5 N.Y.U. J. LEGIS. & PUB. POL'Y 519, 530 (2001); William G. Ross, *The Functions, Roles, and Duties of the Senate in the Supreme Court Appointment Process*, 28 WM. & MARY L. REV. 633, 637–38 (1987).

173. MASS. CONST. pt. 2, ch. 2, § 1, art. IX (1780); *id.* pt. 2, ch. 2, § 1, art. XI.

174. *Id.* pt. 2, ch. 2, § 1, art. V (to adjourn the General Court); *id.* ch. 2, § 1, art. VIII (to pardon offenses); *id.* pt. 2, ch. 2, § 1, art. X, cl. 4 (to appoint various military officers); *id.* pt. 2, ch. 2, § 1, cl. 7 (to appoint officers of the Continental Army); *see also* Op. of the Justices to the Governor and Council, 190 Mass. 6, 617–20 (1906).

175. Robert L. Turner, *A Council Whose Time Has Passed*, BOSTON GLOBE, Feb. 27, 1992, at 19 ("For decades, the council has proved itself to be close to useless, or worse. Its duties have gradually been trimmed to a point at which confirmation of judges and clerks is its only real job.").

Charter of 1691 was the first to vest in the Governor the power to appoint judges, justices of the peace, and sheriffs, with the advice and consent of the Council.[176] That Council, the upper house of the General Court,[177] consisted of twenty-eight members selected initially by the King, later by the Governor, and thereafter by the Governor upon nomination by the General Court.[178] Colonial strife led the King to reassert control over the Council beginning in August 1774, although, under that British legislation, the Governor no longer needed council consent in the appointment of judges and other officials.[179] In 1775, as the Continental Congress steeled itself against the British, the Congress authorized the Provincial Congress (a shadow government outside of British-controlled Boston) to hold elections and to select a new Council to assume partial control of executive responsibilities (the Governor's and Lieutenant Governor's offices having been declared vacant by the Provincial Congress).[180]

The Council proceeded to exert substantial control over local government on the reasoning that, under the Charter of 1691, the Council was to exercise the executive power in the absence of the Governor and Lieutenant Governor. Although the House did question whether the Council or a House majority should be considered the "Governor," exigent circumstances soon forced their hand. On December 10, 1776, with Rhode Island occupied by British troops, the House resolved that "the whole power of the General Court, be during the recess of the Court devolved on the Council of this State, so far forth as are necessary for the purpose of protecting and defending this State . . . and any seven shall be considered as a quorum."[181]

When both the General Court and the Council were in session, they managed to coexist relatively peacefully. From 1775 to 1779, civil officers were appointed "without difficulty" by

---

176. 2 COMMONWEALTH HISTORY OF MASSACHUSETTS 8, 11 (Albert Bushnell Hart ed., 1928).

177. The "General Court" is the name of the Massachusetts state legislature. *Id.* at 11–12.

178. *Id.* at 12.

179. 3 COMMONWEALTH HISTORY OF MASSACHUSETTS 64 (Albert Bushnell Hart ed., 1929).

180. *Id.* at 67–68.

181. *Id.* at 70.

the Council and the House of Representatives together.[182] But after the Declaration of Independence, the ill-fitting colonial Charter of 1691 came under question. Local leaders began to call for a new constitution.

In 1776 and 1777, the General Court called on towns to send representatives for a constitutional convention, but received no effective response. Thus, the General Court and Council gathered themselves into a constitutional convention and produced a draft constitution. This constitution established a Senate and a House of Representatives,[183] a Governor and Lieutenant Governor,[184] but no Council. All civil officers were to be selected annually by vote of the General Court. All other officers, military and civil, including judges, were appointed by the Governor and Senate.[185]

The proposed constitution was rejected overwhelmingly by statewide popular vote, 2,083 to 9,972.[186] The towns of Essex County gathered in convention to publicly enumerate their reasons for rejecting the Constitution of 1778.[187] Among their grievances were the improper intermingling of executive and legislative powers through Governor-Senate cooperation;[188] the impropriety of vesting appointment power in a numerous body prone to vote-trading and cabal;[189] and, most noteworthy

---

182. *Id.* at 73.

183. A CONSTITUTION AND FORM OF GOVERNMENT FOR THE STATE OF MASSACHUSETTS-BAY art. I (1778), *microformed at* the Harvard Law School Library [hereinafter CONSTITUTION OF 1778]; *see also* REPORT OF A COMMITTEE OF CONVENTION OF A FORM OF GOVERNMENT FOR THE STATE OF MASSACHUSETTS-BAY (1777), *microformed at* the Harvard Law School Library [hereinafter CONSTITUTIONAL REPORT OF 1777].

184. CONSTITUTION OF 1778, *supra* note 183, art. II.

185. *Id.* art. XIX. The Constitution did not provide for the mode of cooperation between the Governor and Senate. *See id.*

186. 2 COMMONWEALTH HISTORY OF MASSACHUSETTS, *supra* note 176, at 184.

187. *See* RESULT OF THE CONVENTION OF DELEGATES HOLDEN AT IPSWICH IN THE COUNTY OF ESSEX, WHO WERE DEPUTED INTO CONSIDERATION THE CONSTITUTION AND FORM OF GOVERNMENT PROPOSED BY THE CONVENTION OF THE STATE OF MASSACHUSETTS BAY (1778), *microformed at* the Harvard Law School Library [hereinafter THE ESSEX RESULT].

188. *See id.* at 4–5. More specifically, the document laid out that Article XIX of the 1778 constitution was "exceptional, because a due independence is not kept up in between the supreme legislative, judicial, and executive powers, nor between any two of them." *Id.* This language, which was kept in the Massachusetts Constitution of 1780, was cited with approval by James Madison in his seminal defense of separation of powers. *See* THE FEDERALIST NO. 47 (James Madison), *supra* note 127, at 304–05.

189. THE ESSEX RESULT, *supra* note 187, at 37–38, 46–47, 65–66.

for present purposes, the lack of an executive privy council selected by the legislative branch.[190]

In response to *The Essex Result*, a second constitutional convention was called, resulting in the Constitution of 1780. In his address at the close of the convention, President James Bowdoin highlighted the retention of the Privy Council and its role in the appointment of civil officers, which would "prevent the governor from abusing the Power which is necessary to be put into his hands."[191] Thus, the new constitution retained a strong privy council as a check on the Governor.

B. The Council in Practice: "Advice" and "Consent"

The Constitution of 1780 required that the nine-member Council record all of its resolutions and advice in a register, signed by the members present and available at all times for inspection by either house of the legislature.[192] Moreover, "any member of the council may insert his opinion contrary to the resolution of the majority."[193] In the years immediately preceding and following the effective date of the Constitution of 1780, these records were maintained in bound, handwritten volumes entitled *Council Records*.[194] This author investigated the records of two separate year-long periods between the effective date of the Constitution and the beginning of the Federal Constitutional Convention of 1787: First, November 1780 to November 1781 (the first year under the new Constitution), and second, May 1786 to May 1787 (the final year leading up to the Federal Constitutional Convention). According to the *Council Records*,

---

190. *See id.* at 58–60.

191. AN ADDRESS OF THE CONVENTION FOR FRAMING A CONSTITUTION OF GOVERNMENT, FOR THE STATE OF MASSACHUSETTS-BAY, TO THEIR CONSTITUENTS 15 (1780), *microformed at* the Harvard Law School Library [hereinafter BOWDOIN ADDRESS].

192. MASS. CONST. pt. 2, ch. 2, § 3, art. V (1780).

193. *Id.*

194. The *Council Records* are available on microfiche at the Massachusetts State Archives at the University of Massachusetts at Boston. Having been handwritten rather than printed, individual records vary in legibility. Their legibility is further reduced because the documents were reproduced in negative form (white text on a black page). But given the imaging tools of the microfiche readers, and the subsequent double-checking of all entries, the author is confident that his account of the *Council Records* is highly accurate. An Appendix detailing the entirety of his data is on file with the author.

these two years involved approximately 900 nominations for appointments.

The form of nomination and consent varied little throughout the span of records. In 1780 and 1781, nominations were announced with, "His Excellency was this day pleased to nominate [name(s)]."[195] Council response generally was phrased as, "His Excellency was pleased this day agreeable to his Nomination of [date] to appoint with the Advice and Consent of the Council [name(s)]."[196] By 1786 and 1787, the wording had changed slightly. Nominations were still announced with, "His Excellency nominated to the council the following Gentlemen, viz. [name(s)],"[197] but council response was announced with, "Agreeable to His Excellency's nomination of [date], The Council advised and consented to the appointment of [name(s)]."[198] In the latter cases, no specific record of the Governor's appointment was included in the Council's records, following the records of advice and consent.

The swift and consistent manner with which the Council dispatched its advice and consent is apparent. The constitution itself required that the Governor submit his nominations at least seven days before the planned appointment.[199] In the vast majority of cases, the Council responded to the Governor's nominations within one to three weeks. In most cases the Council announced its advice and consent in the same batches in which the nominations were originally received.

The results of investigation into the *Council Records* were further illuminating. Most notably, of the approximately nine hundred nominations presented to the Council by the Governor, only fifty-seven of them did not appear to receive the advice and consent of the Council—more than half of which were

195. *See, e.g.*, 26 COUNCIL RECORDS 202 (Mar. 19, 1781) (noting the nominations of Nathaniel Gorham et al.).

196. *See, e.g.*, 26 COUNCIL RECORDS 221 (Mar. 27, 1781) (noting the appointments, upon the Council's advice and consent, of Nathaniel Gorham et al.).

197. *See, e.g.*, 30 COUNCIL RECORDS 130 (Apr. 7, 1787) (noting the nominations of James Bowdoin, Jr. et al.).

198. *See, e.g.*, 30 COUNCIL RECORDS 149 (Apr. 26, 1787) (noting the appointments, upon the Council's advice and consent, of the nominations of James Bowdoin, Jr. et al.).

199. MASS CONST. ch. II, § I, art. IX.

nominated en masse for a single office.[200] Of these fifty-seven, not a single rejection was explicitly recorded in the *Council Records*; only nominations advised and consented to by the Council and appointed by the Governor can be found.

This finding appears at first blush to reveal a practice inconsistent with the Massachusetts constitution's requirement that "[t]he resolutions and *advice* of the Council *shall* be recorded in a register."[201] This leads to three possible conclusions:

> (1) the Council acted contrary to the explicit requirement of the constitution in the first year of the constitution's existence, by not recording its advice that the Governor not appoint a nominee; or
>
> (2) the term "advice," standing alone in the constitutional council records requirement, was not synonymous with the advice component of constitutional "advice and consent" of nominees (that is, "advice and consent" was a term of art), such that the council's "advice" not to appoint a nominee was not constitutionally required to be recorded in the *Council Records*; or
>
> (3) the Council's consent to a nomination was a matter of advice that was constitutionally required to be recorded in a register, but the Council's *lack* of consent to a nomination was *not* a matter of advice required to be recorded in a register.

The first of the three possible conclusions seems to be the least plausible, because the Council made extensive records of even the smallest spending resolutions.[202] To suggest that the Council followed the requirement of the constitution in cases involving such small expenditures while ignoring the same requirement with respect to the constitutional appointment process at the center of the constitutional debates of 1777 to 1780 is a dubious proposal.

The second possible conclusion is at least more plausible than the first. In each of the *Council Records* entries announcing the Council's approval of the nomination, the Council used a variation of the phrase "advised and consented to" as a whole;

---

200. 26 COUNCIL RECORDS 149–50 (Feb. 17, 1781) (thirty-three unconfirmed nominations to the post of "Justice for the Taking of Dangerous Persons").

201. MASS. CONST. pt. 2, ch. 2, § 3, art. V (1780) (emphasis added).

202. *See, e.g.*, 26 COUNCIL RECORDS 151 (Feb. 19, 1781) (authorizing expenditure of twenty pounds).

in no case did it specify any added "advice" beyond the mere approval of the candidate, coupled with its consenting to the nomination. This precise repetition of the full phrase suggests that "advice and consent" may have been a term of art. But constitutional provisions before and after the appointments provision also require the Governor to obtain the "advice" of the Council but not the consent. For example, the article preceding the appointment provision requires that the Governor obtain the "advice" of the Council before granting pardons.[203] In addition, the article following the appointment provision requires the Governor to seek the "advice" of the Council before appointing military officers and Continental Army officers;[204] the article following those provisions requires council "advice and consent" before spending Treasury monies.[205] Further refuting the term-of-art reading of "advice" and "advice and consent," James Bowdoin, in announcing the then-new constitution, used the term "advise" in the context of appointments, in the very place of "advice and consent": "[W]e have provided that he shall have a Council to *advise* him at all times and upon all important Occasions, and he with the *advice* of his Council is to have the Appointment of Civil Officers."[206] In light of these considerations, it appears unlikely that "advice" within "advice and consent" was intended to be a term of art not synonymous with other uses of "advice" at the time of Nathaniel Gorham's proposal at the Federal Convention.[207]

The most plausible conclusion, then, is that the Council conveyed dissatisfaction with executive nominations for appointment in a manner not qualifying as official advice for the

---

203. MASS. CONST. pt. 2, ch. 2, § 1, art. VIII (1780).

204. *Id.* pt. 2, ch. 2, § 1, art. X.

205. *Id.* pt. 2, ch. 2, § 1, art. XI.

206. BOWDOIN ADDRESS, *supra* note 191, at 15 (emphasis added).

207. The Massachusetts Supreme Judicial Court has agreed with this conclusion in an advisory opinion, interpreting "advice" alone and "advice" in "advice and consent" to be synonymous:

> We do not think that these different phrases, used in different parts of the Constitution, namely, "by and with the advice of council," "by and with the advice and consent of the council," "with the advice and consent of the council," "with advice of council," and "with advice of the council," differ at all in legal effect. They all recognize the fact that the act, first of all, and afterwards for all time, is to be the act of the Governor. The only connection that the council can have with it is advisory. Whether the Governor takes advice or not, his conclusion must rest finally upon his own judgment.

Op. of the Justices to the Governor and Council, 190 Mass. 616, 619–20 (1906).

purposes of the constitution. In essence, under this theory, "advice" would have been a term of art not just with respect to "advice and consent" situations, but all of the constitutional advice situations: "advice" would have referred only to certain types of responses of the Council to the Governor. At first blush, this seems as implausible as the preceding conclusion; the framers of the Constitution of 1780 clearly wanted a general advisory board for the Governor, one that would meet with the Governor nearly every day[208] and provide him with its opinions on executive matters. But the absence of any negative, nonappointment advice in the spans of *Council Records* investigated for this Article implies that either the Council gave virtually no advice, or that its advice was limited to rare occasions, and only to occasions in which it affirmatively *condoned* action, never when it advised *against* action.[209]

This can be illustrated by example: The overwhelming majority of *Council Records* entries relate to warrants on spending. *In not a single case* do the *Council Records* note the council advising *against* spending; all entries involve the *allowance* of spending. It appears that Massachusetts's Council, in the execution of its constitutional powers and duties, did not find its "advice" to extend beyond limited matters of *affirmative* espousal of proposals set before it.

This interpretation of advice and consent also reflects the traditional use of the phrase in England, which, long before 1787, customarily noted that the King's law was enacted "by and with the advice and consent of the lords spiritual and temporal and commons in parliament assembled" (or a functionally equivalent reference to "advice and consent") in pronouncements of law that received the concurrence of Parliament.[210] "Advice and consent" did not refer to parliamentary

---

208. *See generally* 26 COUNCIL RECORDS (1780–1781).

209. *See, e.g.*, 26 COUNCIL RECORDS 24 (Nov. 16, 1780) ("Advised that Deacon Thomas Foster, Keeper of the Powder House, be & hereby is directed to deliver to Messers. Otis and Henley four Tons of powder lodged in the Magazine in Boston being the property of the United States."); *id.* at 42 (Dec. 2, 1780) ("Advised that a warrant be drawn on Treasury for thirty three hundred & four pounds seven shillings . . . ."). In the overwhelming majority of cases, warrants of spending were phrased without the term "advice." *See, e.g.*, *id.* at 50 (Dec. 13, 1780) ("Warrant drawn on the Treasury for . . . .").

210. *See* Arthur Bestor, *Separation of Powers in the Domain of Foreign Affairs: The Intent of the Constitution Historically Examined*, 5 SETON HALL L. REV. 527, 541–47 (1974).

rejection of the King's proposed course of action, but only its endorsement of that proposal. In this respect, "advice" was not a synonym for "opinion," but rather, it was a synonym for "approval."[211]

Outside of the above conclusions regarding the advice of the council, investigation into the *Council Records* illuminates one further feature of the Council's norms: In no apparent case did the Council record a vote tally or include dissenting opinions of any Councilors. The Council consisted of nine members, each with a constitutional right to include his dissenting opinions in the *Council Records*.[212] Neither in the year following the effective date of the constitution, nor in the year preceding the 1787 Federal Constitutional Convention, did a single member include a dissenting opinion. Furthermore, in no apparent case did the *Council Records* record any vote tallies. It recorded only the ultimate resolution of issues and spoke in apparently unanimous terms.

Finally, it should be noted that, according to the plain text of the constitution, the council was a majoritarian body.[213] Whether the council followed this rule in practice is not known, especially given that the *Council Records* did not record vote tallies.

C. Summary

The Massachusetts process of council advice and consent for executive nominations served as Nathaniel Gorham's model in proposing an advice and consent model of judicial appointment in the Federal Constitutional Convention. As seen in the colonial history of the council, and in the Essex controversy, the people of Massachusetts deemed a strong Council essential to the effective restraint of the Governor; it was a core feature of the revised Massachusetts Constitution of 1780. It was originally envisioned as a majoritarian body, per the text of its con-

---

211. *See id.* at 545. Bestor notes that, although the U.S. Constitution uses the term "advice" with respect to relationship between the President and the Senate on matters of appointment and ratification, it uses the term "opinion" with respect to the relationship between the President and his Cabinet, U.S. CONST. art. II, § 2, cl. 1, suggesting that the terms were not synonyms. But note that, in describing the Massachusetts Constitution of 1780, James Bowdoin appeared to use the term "advise" as a synonym for "opine." *See supra* note 206 and accompanying text.

212. MASS. CONST. pt. 2, ch. 2, § 3, art. V (1780).

213. *Id.*

stitutional mandates. But as seen in the investigation of its records, its advice appears to have been limited to formal pronouncements on matters of spending and appointment, not including any general advice it offered the Governor in their near-daily meetings. It appears that its formal advice and "consent" did not include rejections of nomination or of other gubernatorial requests; rather, it was limited to the affirmative, explicit approval of matters before it. The Massachusetts Council never evidenced obedience to any obligation to actively and publicly reject those nominees who failed to receive its consent, at least not in any way that comported with the explicit mandate of the then-new state constitution, in the brief interim between ratification of the state constitution of 1780 and the Federal Constitutional Convention of 1787. This stands as strong evidence against the suggestion that the Federal Constitution's advice and consent provision includes an implicit obligation that the Senate actively and publicly reject those nominations that fail to secure the Senate's consent.

III. MADISON'S ALTERNATIVE, AND COMPLETIONS VERSUS VETOES

As noted above, Nathaniel Gorham's advice and consent proposal was met by a variety of competing plans for appointment of judges: presidential appointment, senatorial appointment, legislative appointment, and presidential appointment subject to a discretionary senatorial veto.[214] The last of these plans, proposed by James Madison with specific reference to the Gorham plan, provides an interesting counterpoint to the Gorham plan, because in its final version, it would have allowed the Senate to veto the nomination with merely a majority vote.

Because James Madison explicitly differentiated his Senate veto from Gorham's advice and consent, his plan serves as an invaluable piece of evidence in the search for the meaning of Gorham's plan. Nevertheless, Madison's plan has been wholly overlooked in scholarship on the matter. Although Madison's proposal does not completely resolve the questions of advice and consent, it does contribute to the debate, highlighting cru-

214. *See supra* Part I.A.1.c.

cial textual differences between interbranch relationships that call for obligations to act and relationships where no such obligations exist.

A. Madison's Proposal Versus Gorham's Advice and Consent

On July 18, in response to the Gorham advice and consent proposal,[215] James Madison recommended an alternative means of appointment of judges and other officers: presidential appointment subject to a discretionary Senate veto.[216] In its final form on July 21, Madison's proposal ultimately allowed for the Senate to veto the President's appointment by *majority* vote.[217]

Undoubtedly, Madison saw a qualitative difference between the Gorham plan, by which the President's nominations could be appointed subject to Senate advice and consent via majority vote, and his own plan, by which the President's nominations were effective subject to Senate veto by majority vote. Madison was not the only one to recognize a difference between the plans. Randolph, for example, noted that he preferred advice and consent to the Senate veto.[218] Moreover, the Maryland delegation, earlier a supporter of Senate control of appointments, supported Gorham's advice and consent proposal, but voted *against* the Madison plan. And Georgia, which had earlier supported Senate control of appointments, and which voted "absent" on Gorham's plan, also voted *against* the Madison plan.

Madison, a proponent of executive control over appointments; Randolph, the pro-Senate-appointment Maryland delegates; and the pro-Senate-appointment Georgia delegates recognized that the Gorham plan provided the Senate with more power in appointments than the Madison plan. Appointment by the President upon the Senate's advice and consent would empower the Senate more than would appointment by the President subject to the Senate's discretionary veto. In the former, the burden is on the President and supporters of his nominee to engage in organization of Senate support; in the latter, the burden is on the Senate's opponents of the nomina-

---

215. *See supra* note 54 and accompanying text.
216. *See supra* note 63 and accompanying text.
217. *See supra* note 84 and accompanying text.
218. *See supra* note 28 and accompanying text.

tion to engage in organization of Senate action, because the President wins merely by maintaining the status quo. Moreover, the Senate's relative power is increased further when advice and consent includes no obligation to vote either up or down, as compared to when advice and consent includes an obligation to vote either up or down, because the obligation to vote eliminates the President's burden of persuading the Senators to vote, leaving only the burden of persuading them to vote affirmatively. In sum, then, the Senate's relative power increases as the framework moves from presidential appointment subject to discretionary Senate veto, to presidential appointment subject to Senate advice and consent with the Senate obligated to vote, to presidential appointment subject to Senate advice and consent with the Senate not obligated to vote.

But while the relative mechanics of Madison's and Gorham's proposals explain which of the two plans empowered the President and which empowered the Senate, they do not answer the question of whether the relative differences were merely large (that is, discretionary veto versus advice and consent with an obligation to vote) or yet larger (that is, discretionary veto versus advice and consent *without* an obligation to vote). Fortunately, a textual analysis provides further information on this point.

B. A Textual Framework: Completion Versus Veto

The intersection of the legislative and executive branches in appointments, with the responsibility of the President reinforced by the security of the Senate, constitutes a core interbranch "check and balance." It is not, of course, the only one. Other interbranch checks and balances include presidential veto of legislation,[219] presidential "pocket veto" of legislation,[220] congressional review of presidential vetoes,[221] Senate ratification of treaties,[222] congressional impeachment of judges, Presidents, and executive officers,[223] and vice-presidential service as President of the Senate.[224] While generally known as checks

219. U.S. CONST. art. I, § 7, cl. 2.
220. *Id.*
221. *Id.*
222. *Id.* art. II, § 2.
223. *Id.* art. I, §§ 2–3.
224. *Id.* art. I, § 3, cl. 4.

and balances, these powers in fact break down into two categories, which can be referred to as "veto power" and "completion power." Although these categorizations may not necessarily be self-evident when looking at the Constitution, they are quite obvious upon comparing the two plans found to be so different at the Constitutional Convention: Gorham's advice and consent and Madison's veto. But before considering the respective proposals, further development of the analytic framework is appropriate.

1. Veto or Completion?

In matters of veto power, the act of one branch becomes complete without participation by another branch, but another branch may step in to negate the first branch at its discretion. For example, a bill having passed both houses of Congress shall become law if not returned by the President within ten days after being presented to him.[225] If the President disagrees with the bill, then he may in his discretion return it to Congress with his objections, but his failure to act results in enactment of a law.[226] Likewise, impeachments and subsequent removals constitute a veto of general executive or judicial action: The President may conduct himself in perceived execution of the law, but Congress may, in its discretion, veto his general execution through impeachment and removal.

Matters of completion power, on the other hand, consist of checks and balances in which the act of one branch does not realize its desired effect unless and until completion by another branch. An impeachment trial of the President cannot occur until the Chief Justice presides over it.[227] Counterintuitively, the presidential pocket veto represents completion power—or, more precisely, the nonexecution of completion power—because when a bill is presented to the President within ten days of a congressional adjournment, it cannot become a law

225. Except in the cases where the ten-day time limit will lapse after congressional adjournment. *Id.* art. I, § 7, cl. 2.

226. Such a veto constitutes, in its own right, an action subject to veto by another branch: The Congress can overturn a Presidential veto by a two-third vote in each house. U.S. CONST. art. I, § 7, cl. 2.

227. *Id.* art. I, § 3, cl. 6.

unless the President completes Congress's work by signing it.[228]

Madison's discretionary Senate veto clearly falls into the former category, veto power. Just as the President, for ten days, enjoys the discretion to veto legislation presented to him upon bicameral passage, Madison's proposal would have given the Senate a span of time during which to veto, at its discretion, the President's nomination. In contrast, Senate advice and consent with respect to nominations and treaties clearly falls into the latter category, completion power. By comparing advice and consent to other examples of completion power, one can see more clearly whether there exists an obligation to act on judicial nominations.

2. Discretionary Completion or Mandatory Completion?

In cases of veto power there is, by definition, no obligation to act. The power is wholly discretionary. But in cases of completion power, there is no obvious answer to the question of whether an obligation exists. In some cases, such as the Chief Justice's constitutional role as president of presidential impeachment trials, there is obvious completion power, and an obligation to execute it: "the Chief Justice shall preside."[229] This may be referred to as "mandatory completion." In the pocket veto context, where the President has received a bill within ten days of congressional adjournment, there is obviously no obligation to complete; on myriad occasions Presidents have thwarted legislation by not acting on it. This may be referred to as "discretionary completion."[230]

In which of these cases does the advice and consent to presidential nominations for judicial office fall? The most obvious difference between the Senate's advice and consent power and clear examples of mandatory completion power is the absence of an unqualified "shall" directive. The Constitution explicitly states that, upon conditions precedent, one branch will function in completion of another branch. Upon commencement of a trial of presidential impeachment, the Chief Justice "shall"

---

228. *Id.* art. I, § 7, cl. 2.

229. *Id.* art. I, § 3, cl. 6.

230. Like "Veto Power" and "Completion Power," the labels "Discretionary Completion" and "Mandatory Completion" are the invention of the present author.

preside.[231] The Vice President "shall" preside over the Senate generally.[232] In clear cases of discretionary completion, the constitutional directive lacks an unqualified "shall." For example, the Pocket Veto Clause describes the conditions in which the President may allow a bill to expire uncompleted; it provides no "shall" with respect to the President's decision to act.[233]

In light of these differences, senatorial advice and consent to treaties and nominations appears to fall into the discretionary category. The structure of the clause provides no specific directive for the Senate to act; indeed, the language of the clause makes Senate advice and consent a condition precedent to the President's power to appoint.[234]

Although this first point may seem trivial, its importance can be illustrated by looking to the full text of the Constitution. In matters of mandatory action, in interbranch activities or otherwise, the Constitution uses "shall" prescriptions to ensure action: In the context of the Constitution's Appointment Clause, the document prescribes presidential actions using the "shall" directive; it makes no "shall" directive for the condition precedent, Senate consent. The directive "shall" is nowhere to be found in matters of discretionary vetoes or discretionary completion, including the Advice and Consent Clause. The Framers' choice of words on that matter is compelling evidence that the Senate is under no constitutional duty to act on nominations.

C. Summary

Just as the Constitution contains no explicit requirement that the President act in the pocket veto context, it contains no explicit obligation that the Senate act to demonstrate its lack of consent to a judicial nomination. Moreover, the structure of Madison's alternative proposal does little to suggest that the Advice and Consent Clause contains any implicit obligation. As seen in the discussions of Madison's plans, the Framers knew how to create an explicit majority-vote requirement if they wanted. Moreover, they knew how to require the Senate

231. U.S. CONST. art. I, § 3, cl. 6.
232. *Id.* art. I, § 3, cl. 4.
233. *Id.* art. I, § 7.
234. *Id.* art. II, § 2, cl. 2.

to complete an interbranch action when they wanted it so to act.[235]

IV. CONCLUSION

Despite suggestions by the President, various Senators, and numerous commentators that the Senate has a constitutional obligation to act on judicial nominations, the text of the Constitution contains no such obligation. Moreover, the suggestion that the obligation is implicit in the Advice and Consent Clause does not appear to comport with the Framers' understanding of the term. This analysis is limited, of course, to a consideration of (1) the Framers' reliance on the example of the Commonwealth of Massachusetts's advice and consent model; and (2) the structure of their choice of constitutional text, particularly by contrast with James Madison's failed alternative to advice and consent.

President Bush suggests that the Senate bears a "constitutional obligation" to act on judicial nominations,[236] but it is difficult, if not impossible, to demonstrate that the Framers would have agreed with such an assertion. The Framers' understanding of the terms of the Constitution is not necessarily synonymous with the reasonable meaning of the Constitution's terms

---

235. The Department of Justice's Office of Legal Council recently came to a similar conclusion in answering the question whether the President "has a legal duty to appoint and commission a nominee once the Senate has given its advice and consent to the nomination." The OLC, looking to the Appointments Clause, noted:

> The Constitution thus calls for three steps before a presidential appointment is complete: first, the President's submission of a nomination to the Senate; second, the Senate's advice and consent; third, the President's appointment of the officer, evidenced by the signing of the commission. All three of these steps are discretionary. Even after the Senate's advice and consent, up to the moment that the President signs the commission, he can grant or withhold an appointment in accordance with his will and judgment.

Appointment of a Senate-Confirmed Nominee, Op. Off. Legal Counsel (Oct. 12, 1999), http://www.usdoj.gov/olc/marbury_ltr.htm. Although this opinion letter rightfully states that the President is under no duty to nominate someone to fill a vacant office—despite the Constitution's instruction that he "shall" so nominate—this does not contradict this Article's theory of "mandatory" and "discretionary" completion. When the President offers a nomination, he does not "complete" the efforts of another branch. Of course, the President's appointment does complete the work of the Senate, and that discretionary action is not preceded by a "shall" term in its constitutional mandate.

236. Press Release, President George W. Bush, *supra* note 18.

as written, but understanding what the Framers may have understood a term to mean certainly is a useful starting point for further analysis.

ENDING THE WAR ON TERRORISM ONE TERRORIST AT A TIME: A NONCRIMINAL DETENTION MODEL FOR HOLDING AND RELEASING GUANTANAMO BAY DETAINEES

TUNG YIN*

I. DETENTION AT GUANTANAMO: WHO, WHY, AND IN WHAT CONDITIONS? 156
A. Al Qaeda.....157
B. The United States's Response to the 9-11 Attacks.....158
C. Camp X-Ray and Camp Delta.....160
II. CRIMINAL LAW VERSUS THE LAW OF ARMED CONFLICT.....163
A. Different Focuses and Purposes.....164
1. Focus 165
2. Purpose of Confinement.....165
3. Length of Confinement.....170
4. Stigma of Having Been Convicted Versus Having Been Held as a POW...171
5. Summary 172
B. Criminal Law and the War on Terrorism..173
1. Murder, Conspiracy, and Other Common Law-Based Crimes.....174
2. Recent Antiterrorism Statutes: 18 U.S.C. §§ 2339A, 2339B 179
3. War Crimes.....181

* Associate Professor, University of Iowa College of Law; J.D., University of California, Berkeley (Boalt Hall), 1995. Thanks to Eric Andersen, Randy Bezanson, Stephanos Bibas, Christina Bohannan, Arthur Bonfield, Bill Buss, Erin Buzuvis, Jill Gaulding, Kevin Heller, Nicholas Johnson, Amy Liu, Peter Oh, Margaret Raymond, Wendy Schneider, Mark Sidel, Alexander Somek, Peter Spiro, Ethan Stone, and participants at a faculty workshop at the University of Georgia School of Law for helpful feedback, and to Mark Liu ('05), James McHugh ('06), and Jeffrey Stone ('06) for research assistance.

III. NONCRIMINAL DETENTION 182
A. Pretrial Detention 183
B. Civil Commitment of the Mentally Ill 184
C. Civil Commitment of Sex Offenders 185
D. Medical Quarantines 186
IV. NONCRIMINAL DETENTION IN THE WAR ON TERRORISM 188
A. War or Military Force as the Limiting Factor 189
1. Defining the Class of Persons to be Detained 189
2. Application to the War on Terrorism 196
B. Proving Dangerousness 197
1. Counsel 202
2. Periodic Evaluations 204
3. Choice of Forum 206
C. A Note on the Conditions of Detention 207
D. Comparison to the Bush Administration's Administrative Review Hearings 208
V. CONCLUSION 210

Several years after the September 11, 2001 terrorist attacks, the United States continues to detain as "enemy combatants" hundreds of persons captured in Afghanistan or Pakistan during the war on terrorism.[1] Many of those detainees have languished in the detention facility at the U.S. naval base at Guantanamo Bay, with no apparent prospect for release, particularly because this military conflict does not involve an enemy sovereign with whom peace can be negotiated. In addressing this problem, there are two dangers. Continuing to detain persons who are no longer threats to the United States is undesirable and is unlikely to persuade the rest of the world of our good intentions. On the other hand, releasing persons who in fact intend to commit mass violence against the United

1. *See* News Release No. 935-04, Office of the Asst. Sec'y of Def. (Pub. Aff.), Transfer of Detainees Completed (Sept. 22, 2004), *available at* http://www.dod.gov/releases/2004/nr20040922-1310.html [hereinafter News Release No. 935-04].

States or to rejoin the ranks of those fighting our military is also undesirable.[2]

The Bush Administration has not given consistent signals about the future of the Guantanamo detainees. Early on, it seemed as if the detainees would be held indefinitely. Then-Deputy Assistant Attorney General John Yoo said forthrightly in a speech in late 2002, "Does it make sense to ever release them if you think they are going to continue to be dangerous, even though you can't convict them of a crime?"[3] The administration also suggested, drawing upon the rhetoric of the "war on terrorism," that the detainees would be released when the "war" was over.[4] In a sense, the administration's position that the detainees will be released when the war on terrorism is over is tantamount, as critics have charged, to saying that we will hold them for the rest of their lives.[5]

Then, in October 2004, the deputy commander of the detention facility said that "the majority of [the detainees] will either be released or transferred to their own countries."[6] The deputy commander's superior quickly repudiated that statement.[7] And at the beginning of 2005, the *Washington Post* cited unnamed administration officials as pondering "long-term solutions"

---

2. According to the U.S. government, at least seven persons who were released from detention at Guantanamo Bay have returned to fight against the United States in Afghanistan. Vanessa Blum, *U.S. Building New Prisons for Terrorists*, LEGAL TIMES, Oct. 4, 2004, at 1; Matt Moore & John J. Lumpkin, *Some Detainees Have Returned to Terrorism*, HOUSTON CHRON., Oct. 18, 2004, at A3.

3. Henry Weinstein, *Prisoners May Face "Legal Black Hole,"* L.A. TIMES, Dec. 1, 2002, at A1.

4. *See* Hamdi v. Rumsfeld, 542 U.S. 507, 540 (2004) (Souter, J., concurring) ("The Government asserts a right to hold Hamdi under these conditions indefinitely, that is, until the government determines that the United States is no longer threatened by the terrorism exemplified in the attacks of September 11, 2001.").

5. *See, e.g.*, Joseph Lelyveld, *"The Least Worst Place": Life in Guantanamo*, *in* THE WAR ON OUR FREEDOMS: CIVIL LIBERTIES IN AN AGE OF TERRORISM 100, 101–02 (Richard C. Leone & Greg Anrig, Jr. eds., 2003) (interpreting President Bush's statement regarding the detainees as a willingness "to hold the lot of them at Guantanamo until the distant day, if it ever comes, when Islamic terrorist networks have been universally uprooted"); David Luban, *The War on Terrorism and the End of Human Rights*, *in* THE CONSTITUTION IN WARTIME: BEYOND ALARMISM AND COMPLACENCY 219, 228 (Mark Tushnet ed., 2005) ("It follows, then, that the war on terrorism will be a permanent war, at least until the United States decides to abandon it.").

6. John Mintz, *Most at Guantanamo to Be Freed or Sent Home, Officer Says*, WASH. POST, Oct. 6, 2004, at A16.

7. *Id.*

that included "indefinitely imprisoning suspected terrorists whom they do not want to set free or turn over to courts in the United States or other countries."[8] The administration has already begun to lay the groundwork for such long-term solutions by building new prison facilities at Guantanamo Bay.[9]

The war on terrorism has exposed some cracks in the foundation of the law of armed conflict and, as a result, in U.S. domestic law as well. The problem arises because terrorism, even on the scale of the 9-11 attacks, can either be treated as a criminal matter to be handled by civilian courts, or an armed attack warranting response under the law of armed conflict. Prior to 9-11, the United States treated terrorism, including that sponsored by al Qaeda, as a criminal matter; thus, the United States criminally prosecuted the terrorists who bombed the World Trade Center in 1993,[10] the Murrah Federal Building in Oklahoma City in 1995,[11] and the U.S. embassies in Kenya and Tanzania in 1998.[12] In contrast, Congress and the President chose to treat the 9-11 attacks as an armed attack, and Congress authorized the President to use military force against those persons responsible for the attacks.[13]

Neither of these approaches, however, fits the 9-11 attacks perfectly. Meaningful criminal prosecution of al Qaeda terrorists would have required that the United States gain physical custody of them. At the time of the 9-11 attacks, most of the al Qaeda leadership was living in Afghanistan, sheltered by that country's Taliban leaders. The United States, like most other countries, had no diplomatic relations with the Taliban, and therefore extradition of suspected terrorists was not an option.

---

8. Dana Priest, *Long-Term Plan Sought for Terror Suspects*, WASH. POST, Jan. 2, 2005, at A1. Senators from both political parties immediately condemned the plan to hold detainees indefinitely. *Lugar Condemns Plan to Jail Detainees for Life*, WASH. POST, Jan. 3, 2005, at A2 (quoting Senators Richard Lugar (R., Ind.) and Carl Levin (D., Mich.)).

9. *See infra* Part.I.C.

10. *See* United States v. Salameh, 261 F.3d 271 (2d Cir. 2001); United States v. Salameh, 152 F.3d 88 (2d Cir. 1998).

11. *See* United States v. McVeigh, 153 F.3d 1166 (10th Cir. 1998); *see generally* LOU MICHEL & DAN HERBECK, AMERICAN TERRORIST: TIMOTHY MCVEIGH & THE OKLAHOMA CITY BOMBING (2001) (detailing the indictments stemming from the Oklahoma City bombing).

12. *See* United States v. bin Laden, 92 F. Supp. 2d 225 (S.D.N.Y. 2000).

13. Authorization for Use of Military Force, Pub. L. No. 107-40, § 2(a), 115 Stat. 224, 224 (2001).

In fact, military force may have been the only possible approach for capturing the would-be defendants. The law of armed conflict, on the other hand, does not appear to have anticipated the use of military force against nonstate actors, making it difficult to determine when to release detainees captured during such a military conflict. Thus, much of the post-9-11 debate has concerned whether the terrorist attacks constituted criminal acts to be addressed through international cooperation and the criminal justice system, or acts of war by nonstate actors to be met with military force and direct application of the laws of armed conflict, with al Qaeda substituted as the equivalent of the enemy nation-state.[14]

What is needed in this age of military action against nonstate terrorist groups such as al Qaeda is a new subcategory of the law of armed conflict. For the purposes of American law, this new subcategory would be triggered by the appropriate invocation of military force, whether by congressional authorization or by the President's commander-in-chief power to repel sudden attacks.[15]

In developing this new subcategory, there are several issues that, while relevant, are beyond the scope of this Article. First, this Article assumes that the detainees either have been properly accorded prisoner-of-war (POW) status or have been properly denied that status. As of this writing, the appropriate status of the detainees remains a hotly contested matter;[16] the

---

14. The war-versus-law-enforcement dilemma presented by al Qaeda has prompted much scholarship. *See, e.g.*, M. Cherif Bassiouni, *Legal Control of International Terrorism: A Policy-Oriented Assessment*, 43 HARV. INT'L L.J. 83 (2002) (arguing for development of international cooperation in criminal law enforcement of terrorism); Derek Jinks, *September 11 and the Laws of War*, 28 YALE J. INT'L L. 1 (2003) (arguing for armed conflict); Ronald J. Sievert, *War on Terrorism or Global Law Enforcement Operation?*, 78 NOTRE DAME L. REV. 307 (2003); Greg Travalio & John Altenburg, *State Responsibility for Sponsorship of Terrorist and Insurgent Groups: Terrorism, State Responsibility, and the Use of Military Force*, 4 CHI. J. INT'L L. 97, 98 (2003) (noting the two approaches of criminal law and law of armed conflict); John C. Yoo & James C. Ho, *The Status of Terrorists*, 44 VA. J. INT'L L. 207, 209–15 (2003) (arguing for armed conflict); Harold Hongju Koh, Op-Ed., *We Have the Right Courts for Bin Laden*, N.Y. TIMES, Nov. 23, 2001, at A39 (arguing for law enforcement).

15. *See, e.g.*, The Prize Cases, 67 U.S. (2 Black) 635 (1862).

16. *See, e.g.*, Omar Akbar, Note, *Losing Geneva in Guantanamo Bay*, 89 IOWA L. REV. 195, 215–20 (2003) (analyzing POW status of both al Qaeda and Taliban detainees under the Geneva Convention); Lawrence Azubuike, *Status of Taliban and Al Qaeda Soldiers: Another Viewpoint*, 19 CONN. J. INT'L L. 127, 143–154 (2003) (examining the POW status of Taliban and al Qaeda detainees).

Bush Administration contends that neither al Qaeda nor Taliban detainees are entitled to POW status,[17] but various human rights groups have argued that the Taliban detainees do qualify for such status.[18] This Article also assumes that the detainees have been properly classified as combatants, as opposed to noncombatant civilians;[19] that assumption, however, also may be a contested matter at present.[20] These assumptions are relevant because they relate to matters of *status*, which are not insurmountable; the United States could address these objections and still assert authority to continue indefinite detention of the detainees until "the end of the war on terrorism."

Analysis of the potentially indefinite detention of persons in the course of the war on terrorism raises three related questions. First, has military force been properly authorized against an ascertainable group of nonstate actors? If there is no ascertainable group of nonstate actors against whom military force can be authorized, then detention cannot be justified as incident to military action. Second, is the detainee in question actually a member of that ascertainable group? In the current war on terrorism, this question asks whether a given detainee is indeed a combatant under the terms of the military force authorization. If a particular detainee is demonstrably not a combatant, then detention cannot be justified, even if military force has been properly authorized. Arguably, the Combatant Status Review Hearings established by the Department of Defense shortly after the Supreme Court issued its decision in *Rasul v. Bush*[21] resolve this question by providing an annual

17. *See, e.g.*, Donald Rumsfeld, U.S. Sec'y of Def. & Gen. Richard Myers, Chairman, Joint Chiefs of Staff, Pentagon Briefing (Feb. 8, 2002), *available at* http://usinfo.state.gov/regional/nea/sasia/afghan/text/0208dod.htm [hereinafter Rumsfeld Pentagon Briefing].

18. *See, e.g.*, Amnesty Int'l, Memorandum to the US Government on the Rights of People in US Custody in Afghanistan and Guantanamo Bay (Apr. 15, 2002), *available at* http://web.amnesty.org/library/index/engamr510532002; Human Rights Watch, Background Paper on Geneva Convention and Persons Held by U.S. Forces (Jan. 29, 2002), *available at* http://hrw.org/backgrounder/usa/pow-bck.htm.

19. *Cf.* Tung Yin, *Procedural Due Process to Determine "Enemy Combatant" Status in the War on Terrorism*, 73 TENN. L. REV. (forthcoming 2006) (proposing procedures to properly classify combatants).

20. *See, e.g.*, *In re* Guantanamo Detainee Cases, 355 F. Supp. 2d 443 (D.D.C. 2005).

21. 542 U.S. 466 (2004).

forum in which to reevaluate the combatant status of each detainee.[22] Finally, even if the individual's initial detention as a combatant was proper, is continued detention still justified? Put another way, when military force is employed against nonstate actors, particularly ones for whom there may be no equivalent to a nation-state with which to negotiate an armistice or peace treaty, is the detaining nation-state entitled to hold such prisoners for the rest of their lives? If not, when should the detaining nation-state be obligated to release the prisoners?

The focus of this Article is on the third question. Part I provides a review of the detainees at Guantanamo Bay: who are they, why were they captured, and in what conditions are they being held? Part II contrasts criminal law with the law of armed conflict. While there is some overlap between the two areas of law (in a broad sense, both aim to protect society), they serve distinct purposes. Criminal law sets forth the substantive elements required for society to *punish* an individual, whereas the law of armed conflict regulates the *detention* of enemy soldiers.[23] Because criminal law is concerned with punishment, it sets a high bar for detention. It is too high a bar to reach many al Qaeda members, even the high-level architects of the 9-11 attacks who are in U.S. custody in undisclosed locations.

Part III surveys a number of situations in which the Constitution permits noncriminal detention, where society is allowed to detain a person involuntarily for reasons other than proof beyond a reasonable doubt of culpable conduct. There are four such situations: pretrial detention of persons facing indictment, civil commitment of mentally ill persons, civil commitment of violent sexual predators who have completed their criminal sentences, and quarantine or isolation of persons infected with dangerous, contagious diseases.

Part IV synthesizes common elements from the examples of noncriminal detention identified in Part III to come up with a model for noncriminal detention of persons captured during

---

22. For more analysis of the Combatant Status Review Hearings and *Rasul*, see Tung Yin, *The Role of Article III Courts in the War on Terrorism*, 13 WM. & MARY BILL RTS. J. 1035, 1046–56, 1074–89 (2005).

23. Punishment under the law of armed conflict exists when soldiers violate specific rules of conduct governing warfare, but even absent such violations, a nation-state is entitled to incapacitate captured soldiers via detention.

military conflicts against nonstate actors. The goal of this model is to create a procedure for determining when the need to detain individual nonstate actors ends. There is such a need in military conflicts with nonstate actors precisely because there is no enemy sovereign with which to negotiate an end to the conflict. The model would be used only where dealing with the use of military force against nonstate actors; in the case of a traditional nation-state conflict, the Geneva Convention rules would apply. The key features of this new model are that individual nonstate actors may be detained where the government can demonstrate (a) that they fall within the class of persons identified by Congress as the appropriate targets of military force; and (b) that they continue to pose a danger to the United States. The government would bear the burden of proving that the detainee remains dangerous, and in such proceedings, which would be held annually (circumstances permitting), the detainee would be entitled to counsel. Finally, this Article compares this model to the Administrative Review Hearings currently in use at Guantanamo Bay and concludes that the Administrative Review Hearings are flawed in a number of respects, and that they should be reconstituted to address those flaws.

I. DETENTION AT GUANTANAMO: WHO, WHY, AND IN WHAT CONDITIONS?

On September 11, 2001, nineteen terrorists hijacked four airplanes: two departing from Boston, one from Washington, D.C. (Dulles), and one from Newark, New Jersey.[24] The hijackers, foreign nationals from Saudi Arabia (fifteen), the United Arab Emirates (two), Egypt (one), and Lebanon (one), killed the pilots and crashed two of the planes into the World Trade Center towers and the third into the Pentagon; the fourth flight was intended for either the U.S. Capitol or the White House, but the passengers and flight attendants fought back and forced the terrorists to crash the plane in a remote area in Pennsylvania.[25] Less than a week later, Congress authorized the President to

24. THE 9/11 COMMISSION REPORT: FINAL REPORT OF THE NATIONAL COMMISSION ON TERRORIST ATTACKS UPON THE UNITED STATES 1–14 (2003) [hereinafter 9-11 REPORT].

25. *Id.* at 1–14, 433–38.

use military force "against those nations, organizations, or persons he determines planned, authorized, committed, or aided the terrorist attacks that occurred on September 11, 2001, or harbored such organizations or persons."[26] By then, President Bush had already publicly blamed the terrorist group al Qaeda for the attacks.[27]

A. Al Qaeda

The 9-11 attacks were far from the first efforts by al Qaeda to harm the United States. By 1997, the Central Intelligence Agency found evidence that al Qaeda was linked to "attacks on U.S. troops in Aden and Somalia in 1992 and 1993."[28] There was also evidence linking al Qaeda to the so-called 1994–1995 "Manila air plot," in which Ramzi Yousef, one of the masterminds of the 1993 World Trade Center bombing, planned to blow up a dozen jumbo jets over the Pacific Ocean;[29] Philippine authorities learned of the plot and foiled it after torturing a suspected terrorist for sixty-seven days.[30] Al Qaeda may also have had links to the 1996 truck bomb attack on the Khobar Towers housing complex in Saudi Arabia, which killed 19 Americans and wounded 372 others.[31]

---

26. Authorization for Use of Military Force, Pub. L. No. 107-40, § 2(a), 115 Stat. 224, 224 (2001).

27. *See* David S. Cloud et al., *Faint Trail: It's Surprisingly Tough to Pin Terror Attacks on the "Prime Suspect,"* WALL ST. J., Sept. 19, 2001, at A1 ("By the standards of the Old West, President Bush has good reason to declare Osama bin Laden wanted in last week's terrorist attacks, 'dead or alive.'"); David S. Cloud & Neil King, *Terrorists Destroy World Trade Center, Hit Pentagon in Raid With Hijacked Jets*, WALL ST. J., Sept. 12, 2001, at A1 ("One U.S. official said intelligence agencies already had gathered 'strong information' linking Mr. bin Laden to the attacks"); Doyle McManus & Robin Wright, *Broad New U.S. Strategy Emerging to Fight Terror*, L.A. TIMES, Sept. 16, 2001, at A1 ("The administration has not yet compiled conclusive proof that Bin Laden ordered Tuesday's attack . . . . But they believe they are assembling a circumstantial case that will justify U.S. action to capture the terrorist chief . . . ."); Robin Wright & John Daniszewski, *U.S. Asks Pakistan for Help to Track Down Bin Laden*, L.A. TIMES, Sept. 14, 2001, at A26 ("Bin Laden has emerged as the primary focus of the U.S. investigation . . . .").

28. 9-11 REPORT, *supra* note 24, at 109.

29. *Id.* For more information on the Manila air plot, see SIMON REEVE, THE NEW JACKALS: RAMZI YOUSEF, OSAMA BIN LADEN AND THE FUTURE OF TERRORISM 77, 90–91 (2001).

30. *See* ALAN M. DERSHOWITZ, WHY TERRORISM WORKS 137, 249 n.10 (2002) (citing newspaper accounts).

31. 9-11 REPORT, *supra* note 24, at 60.

In early 1998, Osama bin Laden issued a *fatwa* calling "for the murder of every American, anywhere on earth, as the 'individual duty for every Muslim who can do it in any country in which it is possible to do it.'"[32] Al Qaeda's involvement became more apparent in ensuing attacks. On August 7, 1998, al Qaeda operatives set off massive truck bombs nearly simultaneously against the U.S. embassies in Kenya and Tanzania.[33] On October 12, 2000, al Qaeda operatives rammed a small boat loaded with explosives into the *U.S.S. Cole* while that destroyer was docked in Yemen; the blast killed seventeen American sailors and wounded forty.[34] Bin Laden directly supervised this attack by picking the target and the suicide attackers, and by paying for the explosives and other equipment.[35]

During this time, al Qaeda established camps in Afghanistan for the purpose of training fighters and terrorists. In mid-2002, CNN aired footage from videotapes recovered from a suspected al Qaeda location in Afghanistan; some footage showed "mock kidnappings by masked men swiftly intercepting a vehicle on a dusty road, assassinations from the back of a motorcycle and instructions in the use of a surface-to-air missile launcher capable of blasting an airliner out of the skies."[36] Another videotape showed a white dog exposed to a lethal biological or chemical agent; within minutes, the dog was on its back, howling in agony.[37] One antiterrorism report in 2003 estimated that approximately 20,000 men passed through the training camps, only 2,000 of whom have been captured or killed, leaving 18,000 at large, ready to be mobilized against the United States.[38]

B. The United States's Response to the 9-11 Attacks

The World Trade Center towers were still smoldering when President Bush demanded that the Taliban rulers of Afghani-

---

32. *Id.* at 47.

33. *Id.* at 115.

34. *Id.* at 190.

35. *Id.*

36. Mark Jurkowitz, *Tapes Show Array of Terrorist Activities*, BOSTON GLOBE, Aug. 20, 2002, at A12.

37. Robert Russo, *Al-Qaeda Tapes Show Gassing of Dogs*, EDMONTON J., Aug. 20, 2002, at A4.

38. *See* Barry Renfrew, *Al Qaeda Ranks Boosted by War on Terror, Iraq*, WASH. TIMES, May 26, 2004, http://washingtontimes.com/world/20040525-0937488557r.htm.

stan stop protecting al Qaeda and turn the group's leaders over to the United States.[39] Whether because of the intransigence of the Taliban or poor negotiating on the part of the United States,[40] the Taliban refused to turn over the al Qaeda leaders. As a result, on October 7, 2001, the United States, aided by the United Kingdom, launched missile and air strikes against the Taliban and al Qaeda. By October 19, a small number of U.S. soldiers began fighting on the ground with the Northern Alliance. By November, the Taliban forces were in retreat, often withdrawing from major cities without a fight.

Over the course of the fighting, the United States and its allies captured about 10,000 persons believed to be either al Qaeda or Taliban fighters.[41] According to the Department of Defense, most of the 10,000 were either detained in Afghanistan or were released outright.[42] Approximately 1,000 were sent to the U.S. naval base on Guantanamo Bay, Cuba,[43] to be detained as enemy combatants. These 1,000 persons were identified by individual interviews conducted by teams composed of

---

39. Jeanne Cummings & David Rogers, *Bush Says "Freedom and Fear Are at War*," WALL ST. J., Sept. 21, 2001, at A3 ("In an address before a joint session of Congress, Mr. Bush explicitly demanded that the Taliban regime that runs Afghanistan turn over all members of the terrorist group led by Osama bin Laden, close its camps, and release all foreign nationals 'unjustly imprisoned.'").

40. Taliban leaders might have been prepared to turn over Osama bin Laden and his lieutenants, but needed some concessions to save face. On September 20, 2001, a meeting of one thousand high level Muslim clerics in Afghanistan resulted in "an edict that said Osama bin Laden should be persuaded to leave the country," though the United States quickly rejected the suggestion, which may have prompted Taliban leaders also to reject the decree. *See* John F. Burns, *Afghans Coaxing bin Laden, but U.S. Rejects Clerics' Bid*, N.Y. TIMES, Sept. 21, 2001, at A1.

41. Bill Dedman, *U.S. to Hold Most Detainees at Guantanamo Indefinitely*, BOSTON GLOBE, Apr. 25, 2004, at A1. Some of the 10,000 were probably wrongly identified. There are reports that the Northern Alliance was collecting cash bounties from the United States for each suspected al Qaeda or Taliban fighter turned over. *See, e.g.*, *id.*

42. *Id.* at A24.

43. The United States has maintained a naval base on Cuba at Guantanamo Bay pursuant to a lease entered in 1903 and ratified in 1934. *See* Agreement Between the United States and Cuba for the Lease of Lands for Coaling or Naval Stations, U.S.-Cuba, Feb. 16–23, 1903, T.S. 418; *see also* Treaty Between the United States of America and Cuba Defining Their Relations, U.S.-Cuba, May 29, 1934, 48 Stat. 1682; Agreement Between the United States of America and Cuba for the Lease of Coaling or Naval Stations, U.S.-Cuba, July 2, 1903, T.S. No. 426.

Justice Department attorneys, Army personnel, Central Intelligence Agency officers, or a combination thereof.[44]

The government selected Guantanamo Bay for at least two reasons. Located on the other side of the globe from Afghanistan and surrounded by water, it provided a secure location to hold combatants, with a low risk of escape or rescue.[45] Perhaps more importantly, lower federal courts had almost uniformly held that they lacked territorial jurisdiction to hear claims arising from the military base, as it was deemed "outside" U.S. territory.[46] Based on such precedent, the government no doubt believed that its actions there would be free from judicial review.[47]

C. Camp X-Ray and Camp Delta

The first detainees transported to Guantanamo Bay were kept in a facility known as Camp X-Ray, a hastily assembled collection of wire cages, which a visiting human rights attorney described as "dog-run-like cages."[48] Detainees were kept inside these wire cages for as many as twenty-three and a half hours a

---

44. *See* Rumsfeld Pentagon Briefing, *supra* note 17; *see also* Ruth Wedgwood, Op-Ed., *War Comes to Court*, WALL ST. J., Apr. 20, 2004, at A20 (noting that these decisions followed "an intricate multilevel review process involving lawyers, intelligence officers and field commanders").

45. *See, e.g.*, Scott Higham et al., *A Holding Cell in War on Terror*, WASH. POST, May 2, 2004, at A23 (noting that Guantanamo Bay was more attractive than detention facilities in Asia, which were deemed vulnerable); Tania Branigan & Vikram Dodd, *The Bitterest Betrayal*, GUARDIAN WEEKEND, July 19, 2003, at 27 ("Guantanamo Bay, located on the south-eastern tip of Cuba, is reachable only by a US military flight: its remoteness adds to its security.").

46. *See* Cuban Am. Bar Ass'n v. Christopher, 43 F.3d 1412, 1425 (11th Cir. 1995) ("The district court here erred in concluding that Guantanamo Bay was a 'United States territory.'"); Haitian Refugee Ctr., Inc. v. Baker, 953 F.2d 1498, 1513–15 (11th Cir. 1992) (noting in statement of facts that Haitian refugees were "detained outside United States" on naval ships and at Guantanamo Bay); Bird v. United States, 923 F. Supp. 338, 340–42 (D. Conn. 1996) (holding that Guantanamo Bay fell within the "foreign country" exception to the Federal Tort Claims Act); Fed'n for Am. Immigr. Reform, Inc. v. Reno, 897 F. Supp. 595, 598 (D.D.C. 1995) (noting in statement of facts that Haitian refugees were denied entry to the United States and instead sent to Guantanamo Bay or Panama).

47. *See* Rasul v. Bush, 542 U.S. 466, 497–98 (2004) (Scalia, J., dissenting) ("Today, the Court springs a trap on the Executive, subjecting Guantanamo Bay to the oversight of the federal courts even though it has never before been thought to be within their jurisdiction—and thus making it a foolish place to have housed alien wartime detainees.").

48. MICHAEL RATNER & ELLEN RAY, GUANTANAMO: WHAT THE WORLD SHOULD KNOW 36 (2004).

day, leaving them exposed to the hot Cuban sun during the days and to the elements during the nights.

On April 28, 2002, the Administration opened Camp Delta, a more permanent detention facility. Constructed for approximately $10 million, Delta uses prefabricated steel containers, originally designed for international shipping, to create individual cells measuring six feet, eight inches by eight feet.[49] Three of the steel sides are cut out and replaced with steel mesh, allowing a detainee to see and communicate with detainees to either side and across.[50] The cells at Camp Delta provide floor-level toilets—an improvement over Camp X-Ray, where detainees had to use buckets.[51] In a typical week, each detainee is allowed out of his cell for a total of one hour, absent "rewards" or interrogation.[52] Camp Delta is divided into separate camps: 1, 2, 3, 4, Echo, Iguana, and others. Camps 1, 2, and 3 are rated maximum security; Camp 4 is rated medium security. Camp Echo is the facility largely designated to hold detainees who will be tried in military tribunals; it too is maximum security, and the detainees are kept in solitary confinement.[53]

While the similarity between much of Camp Delta and maximum-security prisons (both state and federal) is striking, "in many ways, Camp Delta is harsher."[54] At just over fifty-three square feet, the cells at Camp Delta are smaller than the typical cells for death row inmates.[55] In terms of recreational time, with the exception of detainees at Camp 4, the one hour every other day that Guantanamo detainees are allowed pales

---

49. Lelyveld, *supra* note 5, at 110.

50. *See id.* at 110–11; *see also* Ted Conover, *In the Land of Guantanamo*, N.Y. TIMES MAG., June 29, 2003, at 43.

51. Lelyveld, *supra* note 5, at 111.

52. *Id.*; Conover, *supra* note 50, at 43.

53. RATNER & RAY, *supra* note 48, at 37; Kathleen T. Rhem, *Detainees Living in Varied Conditions at Guantanamo*, ARMED FORCES PRESS SERVICE, Feb. 16, 2005, *available at* http://www.defenselink.mil/news/Feb2005/n02162005_2005021604.html; Carol Rosenberg, *War Crimes Sessions Are Shut Down*, MIAMI HERALD, Nov. 13, 2004, at 3A.

54. Conover, *supra* note 50, at 43; *see also* Carlotta Gall & Neil A. Lewis, *Tales of Despair From Guantanamo*, N.Y. TIMES, June 17, 2003, at A1.

55. *See, e.g.*, Soering v. United Kingdom, 11 Eur. H.R. Rep. 439, 459 (1989) (describing Mecklenburg Correctional Facility in North Carolina as providing a three meter by three meter (ninety-seven square feet) cell); Lelyveld, *supra* note 5, at 111 (noting that Camp Delta cells are "slightly smaller" than the death row cells in Texas).

in comparison to the six to seven hours a week allowed to death row inmates.[56] At Camp 4, so-called "Level 1 detainees," those who are "compliant and willing to follow camp rules," are allowed seven to nine hours a day of access to exercise yards, are provided board games and playing cards, are allotted white uniforms instead of the orange jumpsuits given to other detainees, and are allowed to eat meals together.[57] To help Level 1 detainees endure the heat, Camp 4 provides electric fans and twenty-four-hour ice water.[58]

In addition to putting up with the harsh conditions at Camp Delta, detainees also endure severe abuse, if allegations by released detainees are accurate. Numerous former detainees have claimed that they suffered beatings, interrogations under gunpoint or while chained "like dogs," sexual humiliations, dousing with chemical solvent, and other abuses.[59] Of course, the complainants might be exaggerating, if not fabricating, their allegations, out of anger at the United States.[60] On the other hand, after Camp Delta's commander, General Geoffrey Miller, took over the Abu Ghraib prison facility in Iraq, similar allegations were documented at Abu Ghraib,[61] providing some reason to believe that abuses occurred in Guantanamo Bay.

Camp Delta came with a lifespan projection of five years, which expires in 2007. During Delta's existence, the Pentagon

---

56. *See* Vanessa Blum, *U.S. Building New Prisons for Terrorists: Construction of Guantanamo Jails Signals Long-Term Plans for Base*, LEGAL TIMES, Oct. 14, 2004, at 1; Rhem, *supra* note 53.

57. Rhem, *supra* note 53.

58. *Id.*

59. *See* RATNER & RAY, *supra* note 48, at 41; Matthew Hay Brown, *Freed U.S. Detainees Claim Abuse, The Allegations Could Stir up Anti-American Sentiment, Already at a Fever Pitch in Islamic Nations*, ORLANDO SENTINEL, Mar. 17, 2004, at A1; Neil A. Lewis, *Frequent Prisoner Coercion Alleged; Guantanamo Personnel Detail Regular Abuse of Those Held at Base*, HOUSTON CHRON., Oct. 17, 2004, at 23; Carol Rosenberg, *3 GIs Punished for Detainee Abuse, Mistreated Inmates at U.S. Military Prison at Guantanamo Bay*, PITTSBURGH POST-GAZETTE, Nov. 12, 2004, at A11; Charlie Savage, *Ex-Detainees Detail Alleged Abuse at US Base*, BOSTON GLOBE, Aug. 5, 2004, at A1; Paul Waugh, *My Torture Hell in Camp Delta Cage*, EVENING STD., Aug. 3, 2004, *available at* http://www.thisislondon.com/news/articles/12320351.

60. For example, one news story described a former detainee's complaint of suffering through forty-five days of sleep deprivation. *Freed Afghans Say U.S. Captors Abusive; Beatings Alleged at Guantanamo*, CHICAGO TRIB., Mar. 17, 2004, § 1, at 12.

61. *See* ANTHONY R. JONES, AR-15 INVESTIGATION OF THE ABU GHRAIB PRISON AND 205TH MILITARY INTELLIGENCE BRIGADE, Aug. 23, 2004, *available at* http://www.globalsecurity.org/intell/library/reports/2004/intell-abu-ghraib_ar15-6.pdf.

has constructed a new camp, known as Camp 5, and has also asked for $25 million to construct Camp 6.[62] Camp 5 is another "supermax" facility designed to hold one hundred inmates and intended to be a permanent structure. Camp 6 will be a medium security facility designed to hold more cooperative detainees.[63]

Not all detainees face indefinite detention; the United States has periodically announced releases of detainees, totaling 256 as of November 2005.[64] At least fifty-six of those detainees were repatriated for further detention, while the remainder were simply released to their home nations.[65] Nevertheless, as of November 2005, the United States continued to hold five hundred detainees at Guantanamo Bay.[66] Although these remaining detainees might be released in the future as a result of the Administrative Detainee Hearings, the existence of Camp 5 and the plans for Camp 6 suggest that the government intends to detain *some* prisoners indefinitely.

II. CRIMINAL LAW VERSUS THE LAW OF ARMED CONFLICT

One of the main criticisms of the Bush Administration's detention of prisoners at Guantanamo Bay is that the detainees are kept in a lawless environment where they face indefinite detention with no prospect of release. The establishment of Combatant Status Review Hearings, charged with evaluating whether each detainee is in fact an enemy combatant, has done little to blunt this criticism.[67] Similarly, the administration's Administrative Review Hearings, which examine whether each detainee continues to pose a threat to the United States, have also failed to satisfy critics. Instead, these critics call for either criminal prosecution of the detainees or outright release.[68] This

---

62. *See* Carol Rosenberg, *Permanent Jail Set for Guantanamo*, MIAMI HERALD, Dec. 9, 2004, at 1; Blum, *supra* note 56 at 1.

63. *See* Blum, *supra* note 56, at 1.

64. *Id. See* News Release No. 1151-05, Office of the Asst. Sec'y of Def. (Pub. Aff.), Detainee Transfer Announced, Nov. 5, 2005, *available at* http://www.defenselink.mil/releases/2005/nr20051105-5073.html [hereinafter News Release No. 1151-05].

65. Blum, *supra* note 56, at 1.

66. *See* News Release No. 1151-05, *supra* note 64; News Release No. 935-04, *supra* note 1.

67. *See* Yin, *supra* note 23, at 1100.

68. *See, e.g.*, Duncan Campbell, *Isolated, Strictly Supervised and Deprived of Rights: How Camp Delta Treats Inmates*, GUARDIAN, July 5, 2003, at 4; Press Release,

Part of the Article considers the different goals of criminal law and the law of armed conflict and analyzes the adequacy, or lack thereof, of the former with regard to protecting the nation from the threat posed by terrorism suspects detained at Guantanamo Bay.

A. Different Focuses and Purposes

At a broad level of generality, criminal law and the law of armed conflict serve the same purpose: to protect society from persons who would do it harm. In the case of criminal law, the harm may have already occurred (though not necessarily, given inchoate crimes such as conspiracy and attempted murder), but punishing the criminal nevertheless is justified, in part to protect society by incapacitating the offender.[69] In the case of the law of armed conflict, which is governed largely by the 1949 Geneva Convention relative to the Treatment of Prisoners of War,[70] the harm justifying intervention includes potential future harm; detention of captured enemy soldiers is justified purely on preventative incapacitation grounds.

Beyond the broad generality of protecting society, however, there are significant differences between criminal law and the law of armed conflict.

	Criminal Law	**Law of Armed Conflict**
Focus	Individual culpability	Group status
Purpose of Confinement	Retribution (punishment)	Prevention (incapacitation)
Length of Confinement	Set by judicial sentence	Duration of the armed conflict
Stigma Imposed	Yes	No

Amnesty Int'l, Detainees From Afghan Conflict Should Be Released or Tried (Nov. 1, 2002), http://web.amnesty.org/library/index/ENGAMR511642002.

69. *See infra* Part II.A.2.

70. Geneva Convention Relative to the Treatment of Prisoners of War art. 12, Aug. 12, 1949, 6 U.S.T. 3316, 75 U.N.T.S. 135 [hereinafter Geneva Convention (III)].

1. Focus

The first difference identified above is that criminal law focuses on individual culpability, whereas the law of armed conflict focuses on group status. Put simply, criminal law punishes the defendant because of what he has done in the past; the law of war detains an enemy soldier because of what that person may do in the future, because of that person's membership in the armed forces of an enemy nation.

To justify incarceration of an individual pursuant to criminal law, the government must prove beyond a reasonable doubt that the person violated some substantive law for which the prescribed punishment includes imprisonment.[71] Additionally, the crime must require that the defendant act with a culpable mental state. (The rare crimes that do not require any particular mental state[72] are regulatory crimes that do not carry the threat of imprisonment.) On the other hand, under the law of armed conflict, it is the detainee's status as a member of a group, not his past conduct, that justifies detention as a prisoner of war. The key concept here is "combatant." Combatants are defined primarily as persons who are "[m]embers of the armed forces of a Party to the conflict" or "other militias or volunteer corps" who follow a chain of command, openly carry arms, have a fixed and distinctive sign (for example, a uniform), and conform to the laws of war.[73] Once captured by an enemy state, a combatant can be detained lawfully for the duration of active hostilities, even if the combatant has committed no crimes under either the law of the enemy state or international law, and even if the combatant in fact harbors no intention or desire to fight the detaining state. The laws of war allow such detention because of a presumption of inherent dangerousness of enemy soldiers.

2. Purpose of Confinement

A second difference between criminal law and the law of armed conflict is the theoretical justification for confinement of

---

71. *See, e.g.*, *In re* Winship, 397 U.S. 358, 364 (1970).

72. *See, e.g.*, United States v. Park, 421 U.S. 658, 670–71 (1975).

73. Geneva Convention (III), *supra* note 70, art. 4. Other combatants entitled to prisoner of war status when captured include those who "spontaneously take up arms" when invaded, provided they carry arms openly and comport with the laws of war. *Id.*

prisoners. Imprisonment typically serves one or more classic purposes: retribution, deterrence, rehabilitation, and incapacitation.[74] The relative significance of these purposes has certainly shifted over time,[75] but almost all justifications of punishment involve one or more of these concepts. Detention of enemy prisoners of war, however, satisfies only the goal of incapacitation: By detaining enemy prisoners of war, a country protects itself and its soldiers from being attacked by those prisoners, who presumably would be pressed back into service if released.[76]

It is difficult to see how detention of prisoners of war could satisfy any other purpose of imprisonment. Consider, for example, H.L.A. Hart's definition of retribution: "the application of the pains of punishment to an offender who is morally guilty."[77] Under the law of armed conflict, an enemy soldier comporting with the laws of war is not morally guilty of anything, because that soldier is entitled to combatant immunity.[78] An act that would ordinarily be considered murder—for example, deliberately shooting another person with intent to kill—is not a crime when the shooter and target are soldiers

---

74. *See, e.g.*, MICHAEL S. MOORE, LAW AND PSYCHIATRY: RETHINKING THE RELATIONSHIP 233–38 (1984).

75. Before the Revolutionary War, rehabilitation was not a recognized goal, which may explain why death was the usual punishment for felonies. *See, e.g.*, United States v. Scroggins, 880 F.2d 1204, 1206 (11th Cir. 1989). By the 1800s, however, rehabilitation came into favor with the belief that offenders could be "cured." *Id.* (citing ARTHUR W. CAMPBELL, LAW OF SENTENCING § 2 (1978)). The desire to rehabilitate offenders brought indeterminate sentences, where the amount of prison time served by an offender was based on a parole board's determination of when the offender was ready to return to society. *See id.* at 1206–07. By the 1970s, rehabilitation had lost significant support, largely because rehabilitating criminals proved difficult and often unsuccessful. *See* Mistretta v. United States, 488 U.S. 361, 365–66 (1989); ANDREW VON HIRSCH ET AL., THE SENTENCING COMMISSION AND ITS GUIDELINES 3–4 (1987).

76. *See, e.g.*, *In re* Territo, 156 F.2d 142, 145 (9th Cir. 1946).

77. H.L.A. HART, PUNISHMENT AND RESPONSIBILITY: ESSAYS IN THE PHILOSOPHY OF LAW 9 (1968).

78. *See, e.g.*, Waldemar A. Solf & Edward R. Cummings, A *Survey of Penal Sanctions Under Protocol I to the Geneva Conventions of August 12, 1949*, 9 CASE W. RES. J. INT'L L. 205, 212 (1977); *see also* Johnson v. Eisentrager, 339 U.S. 763, 793 (1950) (Black, J., dissenting) ("[L]egitimate 'acts of warfare,' however murderous, do not justify criminal conviction . . . . [I]t is no 'crime' to be a soldier." (citing *Ex parte* Quirin, 317 U.S. 1, 30–31 (1942))); United States v. Valentine, 288 F. Supp. 957, 987 (D.P.R. 1968) ("Mere membership in the armed forces could not under any circumstances create criminal liability." (citing Ford v. Surget, 97 U.S. 594, 605–06 (1878))).

during a military conflict. This is true even if the war itself is deemed illegal under international law (for example, a war of aggression), because the law of armed conflict has evolved to draw a distinction between *jus ad bellum* ("justice of war," or the reason for military action) and *jus in bello* ("justice in war," or the conduct of military action in a war).[79] Thus, even though Germany and Japan were deemed to have started an aggressive war (World War II), earning many of the top political and military leaders death sentences at the International Military Tribunals at Nuremburg and Tokyo, the only individual soldiers prosecuted for war crimes were those who committed individual atrocities. No German or Japanese soldiers were prosecuted solely on the basis of having fought in an aggressive war.

Of course, there are instances where a nation seeks to exact retribution against an enemy soldier by prosecuting that soldier in a military trial.[80] Some notorious examples from World War II include at least eighteen Japanese soldiers convicted for mutilating dead Allied soldiers by bayoneting or eating them, as well as Max Schmid, a German medical officer convicted of maltreating the corpse of an unknown U.S. soldier by severing the head, boiling away the flesh, bleaching the skull, and keeping it on his desk for months.[81] Conviction of enemy soldiers results in punishment ranging from incarceration for a term of years to death. These instances of punishment, however, involve war crimes in violation of established international law.[82]

---

79. MICHAEL WALZER, JUST AND UNJUST WARS: A MORAL ARGUMENT WITH HISTORICAL ILLUSTRATIONS 21 (2d ed. 1992).

80. Recent war crimes prosecutions are those in the International Military Tribunals for the Former Yugoslavia, Rwanda, and Sierra Leone. *See* Statute of the International Tribunal for Rwanda, Nov. 8, 1994, 33 I.L.M. 1598; Statute of the International Tribunal for Yugoslavia, May 25, 1993, 32 I.L.M. 1203; Statute of the Special Court for Sierra Leone (Aug. 14, 2000), http://www.sc-sl.org/scsl-statute.html; *see also* Tara Lee, *American Courts-Martial for Enemy War Crimes*, 33 U. BALT. L. REV. 49, 66 (2003) (arguing that the United States should exercise the power given by Congress to court-martial enemy soldiers for war crimes against American soldiers, with specific reference to war crimes committed by Iraqi soldiers during the 1991 Gulf War).

81. *See* PHILIP R. PICCIGALLO, THE JAPANESE ON TRIAL: ALLIED WAR CRIMES OPERATIONS IN THE EAST, 1945–1951, at 78, 128–29, 190 (1979); TRIAL OF MAX SCHMID, XIII LAW REPORTS OF THE TRIALS OF WAR CRIMINALS 151, 151–52 (1949).

82. For example, the mutilation crimes discussed above violated Article 3 of the 1929 Geneva Convention and Article 3 of the 1906 Geneva Convention, both predecessors to the 1949 Geneva Convention.

Detention of soldiers probably will not serve a deterrence function. The individual soldiers who face the possibility of being captured and detained as prisoners of war are unlikely to be specifically deterred. As Michael Walzer notes, the laws of war "assume that [soldiers] all fight unwillingly."[83] Even in the United States military today, which is comprised exclusively of volunteers, soldiers fight because they are so ordered, not because they so choose.[84] Furthermore, the fact that some U.S. soldiers seek injunctive relief in an attempt to avoid being sent on military missions suggests an unwillingness of soldiers to fight.[85] Thus, the prospect of being detained as a prisoner of war if captured is unlikely to have any significant deterrent effect upon an individual soldier's decision to fight.[86] As for general deterrence, the only relevant parties who can be deterred are the political and military leaders who decide to engage in military conflict. Such leaders are unlikely to be deterred by the prospect of detention of captured soldiers, however, because the leaders are themselves generally not in danger of being detained as prisoners of war.

---

83. WALZER, *supra* note 79, at 35. As further support for Walzer's contention that soldiers fight unwillingly, consider that American soldiers who fire their weapons without having been ordered into conflict (and not in self-defense) would generally violate rules of engagement and face punishment, including possible court-martial. *See generally* Mark S. Martins, *Rules of Engagement for Land Forces: A Matter of Training, Not Lawyering*, 143 MIL. L. REV. 3, 36–37, 50, 57 (1994).

84. The increasing use of professional contractors to perform duties for the Defense Department may call this assumption into question. One scholar estimates that there are up to 20,000 civilians working as contractors in Iraq, earning up to $200,000 per year. *See* Jonathan Turley, *Shadow Army Casts a Pall*, FT. WORTH STAR-TELEGRAM, Sept. 20, 2004, at B11; *see also* Mark McDonald, *Afghans to Try U.S. Mercenary Operated Freelance Prison, Torture Chamber*, PITTSBURGH POST-GAZETTE, July 21, 2004, at A5 (discussing trial of Jonathan Idema, an American who went to Afghanistan seeking to capture bin Laden for the $50 million bounty).

85. *See, e.g.*, Doe v. Rumsfeld, No. Cir. 5-04-2080 FDC K, 2004 WL 2753125, at *1 (E.D. Cal. Nov. 5, 2004) (denying injunction sought by U.S. military reservist to keep him from being sent back to Iraq). In other instances, American soldiers have simply refused to be deployed in an effort to avoid being sent off to fight. *See, e.g.*, Russ Bynum, *War Veteran Refuses 2nd Iraq Deployment*, ASSOCIATED PRESS, Jan. 13, 2005, *available at* http://www.truthout.org/docs_05/011405W.shtml (discussing examples).

86. There might be some lesser specific deterrent effect on the soldier in the sense of making the soldier more reluctant to expose himself to the possibility of battlefield capture, which might result in less effective fighting. Another possibility is that soldiers who face capture may, depending on the reputation of prisoner of war camps, decide to fight to the death rather than face capture.

Finally, rehabilitation is not a goal when a nation detains enemy prisoners of war. Rehabilitation presupposes that the defendant has a flaw or defect of some sort to be remedied; a violent criminal, for example, would be "cured" of his violent tendencies. Under the law of armed conflict, however, an enemy prisoner of war has no defect or flaw of which to be cured save his national allegiance.

Article 110 of the Geneva Convention supports the conclusion that the law of armed conflict allows detention of enemy prisoners of war solely to prevent them from engaging in further conflict against the detaining state. Article 110 requires that the following class of prisoners be repatriated directly: (1) "Incurably wounded and sick whose mental or physical fitness seems to have been gravely diminished"; (2) "Wounded and sick who, according to medical opinion, are not likely to recover within one year, whose condition requires treatment and whose mental or physical fitness seems to have been gravely diminished"; and (3) "Wounded and sick who have recovered, but whose mental or physical fitness seems to have [been] gravely and permanently diminished."[87] The underlying rationale of article 110 is that enemy prisoners of war who no longer pose any threat to the detaining State—whether because of incurable wounds, long-term injuries, or other mental or physical impairment—should be repatriated. During the Korean War, almost 7,000 North Korean and Chinese prisoners of war were exchanged for 684 United Nations soldiers under this provision.[88]

Under criminal law, on the other hand, there is no such general release of prisoners whose physical or mental conditions have deteriorated to the point where they no longer pose any future threat to society. On occasion, prisoners falling into such categories have been granted compassionate release, but there has been no obligation on the part of the government to do so.[89]

Notice also the difference in treatment of prisoners who are recaptured after successfully escaping confinement. In the case of prisoners of war, the Geneva Convention holds that such

87. Geneva Convention (III), *supra* note 70, art. 110.

88. LAW OF WAR WORKSHOP DESKBOOK 88 (Brian J. Bill ed., 2004).

89. *See generally* Marjorie P. Russell, *Too Little, Too Late, Too Slow: Compassionate Release of Terminally Ill Prisoners: Is the Cure Worse than the Disease?*, 3 WIDENER J. PUB. L. 799, 819–32 (1994) (surveying state practices).

escapees "shall not be liable to any punishment in respect of their previous escape."[90] Even a prisoner of war who unsuccessfully attempts to escape from the prisoner of war camp may be subject only to "disciplinary punishment."[91] Such disciplinary punishment is limited to fines taken out of earnings (for no more than thirty days), loss of discretionary privileges (those not required by the Geneva Convention), "fatigue duties," or confinement.[92] A civilian prisoner who escapes or attempts to escape, on the other hand, may generally be prosecuted for having committed a criminal act.[93] Thus, the law of armed conflict recognizes a norm whereby prisoners of war may try to escape from POW camps with minimal consequences if they fail, and none if they succeed and are subsequently recaptured. Criminal law, however, does not recognize such a norm among civilian prisoners.

3. Length of Confinement

Under the law of armed conflict, prisoners of war may be detained until the "cessation of active hostilities"; once active hostilities have ended, the prisoners of war "shall be released and repatriated without delay."[94] Prisoners convicted under criminal law, on the other hand, serve a sentence that is set by the judge (though in some instances, the sentence might be shortened through parole or commutation).

Of course, the term "cessation of active hostilities" does not submit to any categorical definition. Obviously, a formal peace treaty that resolves the underlying dispute and reinstates peace would constitute the cessation of active hostilities. However, there need not be a formal peace treaty in place for active hostilities to be deemed over. The state of war between the United States and Germany during World War II did not end formally until 1951,[95] but the obligation to repatriate arose much sooner,

---

90. Geneva Convention (III), *supra* note 70, art. 91.

91. *Id.* art. 92.

92. *Id.* art. 89.

93. *See, e.g.*, 18 U.S.C. § 751 (1994); ARIZ. REV. STAT. ANN. § 13-2503 (2001); CAL. PENAL CODE § 4530 (2000).

94. Geneva Convention (III), *supra* note 70, art. 118.

95. Pub. L. No. 82-181, 65 Stat. 451 (1951).

given Germany's surrender in 1945.[96] One commentator notes, "It is the practical application of the Parties and the way they *construct* a particular agreement on the cessation of hostilities which will operate toward the termination of hostilities . . . ."[97] Still, unlike the prisoner imprisoned under a criminal sentence, the prisoner of war detained under the law of armed conflict faces a term of confinement that is of uncertain duration.

4. Stigma of Having Been Convicted Versus Having Been Held as a POW

A final difference between criminal law and the law of armed conflict is the perception that the general public may have about a person convicted of a crime, compared to that of a person captured and held as a prisoner of war, as well as the loss of privileges and rights stemming from such conviction.

Convicted persons, in addition to being subjected to the usual penalties of incarceration, are often subjected to public shame. One stigma that accompanies a conviction is that in subsequent civil or criminal court proceedings, a convict who testifies can be impeached by the mere fact of his prior conviction,[98] including convictions in foreign jurisdictions.[99] This is an exception to the general rule of inadmissibility of so-called "character evidence."[100] The underlying rationale for such an exception is grounded in the idea that a person who has been previously convicted of a felony (including ones that do not involve dishonesty) will be disposed to lie under oath.[101] Also, because there is a common reluctance to hire felons, a con-

---

96. Still, the United States held some German POWs until the middle of 1947, almost two years after Germany had surrendered to the Allied Powers. *See* ARNOLD KRAMMER, NAZI PRISONERS OF WAR IN AMERICA 249 (1979).

97. CHRISTIANE SHIELDS DELESSERT, RELEASE AND REPATRIATION OF PRISONERS OF WAR AT THE END OF ACTIVE HOSTILITIES: A STUDY OF ARTICLE 118, PARAGRAPH 1 OF THE THIRD GENEVA CONVENTION RELATIVE TO THE TREATMENT OF PRISONERS OF WAR 97 (1978).

98. *See* FED. R. EVID. 609.

99. *See, e.g.*, United States v. Wilson, 556 F.2d 1177, 1178 (4th Cir. 1977) (per curiam) (rejecting appeal based on the use of a rape conviction in Germany to impeach the defendant).

100. *See* FED. R. EVID. 404(a).

101. The logic of this rationale has not gone unchallenged. *See, e.g.*, H. Richard Uviller, *Credence, Character, and the Rules of Evidence: Seeing Through the Liar's Tale*, 42 DUKE L.J. 776, 813 (1993).

victed felon is likely to face difficulty in finding employment.[102] Federal law also prohibits convicted felons from owning firearms,[103] evincing a societal belief that such persons are not to be trusted with deadly weapons, even after they have served their terms of imprisonment and even if the original crime of conviction was not a crime of violence.[104] Felons are not the only ones who suffer from this disability; the others—who include fugitives from justice, illegal aliens, dishonorably discharged soldiers, and Americans who have renounced their citizenship—are also ones viewed, rightly or wrongly, with suspicion by society.[105]

Persons detained as prisoners of war, on the other hand, generally are not subject to such stigmas. The long list of persons who are prohibited from possessing firearms does not include persons who served as prisoners of war in American POW camps, even though such persons may have raised, if not fired, weapons against Americans. Likewise, prisoners of war are not subject to impeachment at trial. Finally, there is scant evidence of any reluctance to hire former enemy POWs.[106]

5. Summary

Although criminal law and the law of armed conflict serve the same general purpose of protecting society from those who would do it harm, they accomplish this goal in distinctly different ways. Criminal law stigmatizes offenders as wrongdoers who deserve the imposition of punishment, whereas the law of

---

102. *See, e.g.*, Walter Matthews Grant et al., *Special Project—The Collateral Consequences of a Criminal Conviction*, 23 VAND. L. REV. 929, 1229–30 (1970) (quoting JOINT COMM'N ON CORR. MANPOWER AND TRAINING, THE PUBLIC LOOKS AT CRIME AND CORRECTIONS 15 (1968)); *see also* Kevin Osborne, *Second Chance Summit Goal*, CINCINNATI POST, Mar. 11, 2004, at A13.

103. 18 U.S.C. § 922(g) (2000).

104. A felon is prohibited from "possess[ing] in or affecting commerce, any firearm" so long as the crime of conviction is punishable by more than a year of imprisonment. *Id.*

105. *See id.* §§ 922(g)(2), (5)–(7). Sections 922(g)(3)–(4) also prohibit anyone "who has been adjudicated as a mental defective or who has been committed to a mental institution" or who is a drug addict from possessing a firearm. Those are categories of persons who might not be viewed in the same light as felons, dishonorably discharged soldiers, and the like.

106. Indeed, considering that such persons may have been inducted involuntarily into fighting, one could argue that discrimination against them would be a form of discrimination based on national origin in violation of Title VII. 42 U.S.C. § 2000e-2(a) (2000).

armed conflict justifies the detention of enemy soldiers (who may indeed be viewed as honorable) for the sole purpose of incapacitating them. With regard to detainees, those on whom we seek to inflict punishment should, of course, be dealt with through substantive criminal law (whether in civilian or military court). However, just as not all enemy soldiers in a nation-state war are "punished," it also may be the case that not all Guantanamo detainees are to be punished. Criminal law may not be the appropriate set of rules to apply to them if the goal is preventative incapacitation based on their status as members of an organization declared by Congress to be belligerent.

At the same time, there is a major problem with the application of the law of armed conflict to the Guantanamo detainees, for the law of armed conflict presumes that enemy soldiers are inherently dangerous because of loyalty to or coercion by their national government. That is, the only event (apart from incapacitating injury) that ends their inherent dangerousness is the end of military conflict between the detaining nation-state and the enemy nation-state; until that happens, there is no need even to inquire into the individual soldier's dangerousness. With the Guantanamo detainees, however, application of that same presumption of inherent dangerousness leads to the troubling possibility of endless detention, for there may never be a clear end to the military conflict between the United States and al Qaeda.

B. Criminal Law and the War on Terrorism

While criminal law and the law of armed conflict can thus be said to serve different purposes, there is an undeniable area of overlap with regard to certain al Qaeda members. Already, the government has prosecuted a number of defendants for violating federal laws.[107] Other high-level al Qaeda suspects currently in custody, such as Khalid Sheikh Mohammed and Ramzi Binalshibh, could presumably face criminal charges in the near future. To the extent that our goal is to inflict *punishment* on such persons, criminal law, with its procedural stan-

107. *See, e.g.*, United States v. Lindh, 212 F. Supp. 2d 541 (E.D. Va. 2002) (refusing to dismiss the indictment of American citizen who joined and provided material support to the Taliban); United States v. Reid, 214 F. Supp. 2d 84 (D. Mass. 2002) (upholding pretrial detention of defendant charged with attempting to detonate a bomb on a transatlantic flight).

dards, is the appropriate vehicle. Yet, constitutional and statutory requirements may prevent us from prosecuting certain detainees.

1. Murder, Conspiracy, and Other Common Law-Based Crimes

Almost 3,000 people were killed in the 9-11 attacks in New York, Washington, D.C., and Pennsylvania.[108] The al Qaeda members responsible for planning and executing the attacks could theoretically be prosecuted for numerous counts of murder, conspiracy, and other such crimes, though, of course, the direct perpetrators of the 9-11 attacks, the nineteen hijackers, all died. But persons who aided and abetted the attacks can be punished as if they were principals,[109] so long as the prosecution proves that the aider and abetter intended to facilitate the crime and took some affirmative action toward doing so.[110]

Binalshibh, for example, was a key member of 9-11 leader Mohammed Atta's "cell" in Hamburg, Germany, and, in 1999, the entire cell (consisting of Binalshibh, Atta, Ziad Jarrah, and one other) ended up in Afghanistan after being recruited by an al Qaeda member.[111] There, Atta, Binalshibh, and Jarrah allegedly pledged loyalty to bin Laden, after which they were assigned by bin Laden's deputy, Mohammed Atef, to a "highly secret mission" that involved flight training.[112] Not only did Binalshibh allegedly receive the same instructions that 9-11 hijackers Atta and Jarrah did on blending in with Americans, he also researched flight schools in Europe.[113] According to the government, Binalshibh probably would have been one of the hijackers but was unable to obtain a visa to enter the United States; he stayed behind and "provide[d] critical assistance from abroad to his co-conspirators."[114] Among other things, Binalshibh supposedly wired money to Atta, passed along a status report from Atta to al Qaeda, and conveyed instructions

108. 9-11 REPORT, *supra* note 24, at 311.

109. *See, e.g.*, 18 U.S.C. § 2 (2000) ("Whoever commits an offense against the United States or aids, abets, counsels, commands, induces or procures its commission, is punishable as a principal.").

110. *See, e.g.*, United States v. Dolt, 27 F.3d 235, 238 (6th Cir. 1994).

111. 9-11 REPORT, *supra* note 24, at 164–66.

112. *Id.* at 166.

113. *Id.* at 167–68.

114. *Id.* at 168; *see also id.* at 225 (noting that Binalshibh had even sent a deposit to a flight school in the United States).

from bin Laden to Atta to carry out the attacks as soon as possible, as well as bin Laden's preference for the White House as an attack target.[115] It is hardly surprising that the 9-11 Commission described Binalshibh as a "key player" in the 9-11 attacks.[116] If the government is correct, he helped the 9-11 hijackers by making funds available so that they could stay in the United States, obtain flight training, and test airline security systems; he also served as a conduit of information between Atta and bin Laden.

Khalid Sheikh Mohammed allegedly conceived of using hijacked airplanes as weapons back in 1995, at a time when bin Laden and his deputy Atef were still limited to thinking about blowing planes up.[117] In 1999, bin Laden approved the "planes operation" and created a list of targets with Atef and Mohammed; bin Laden insisted on the White House and the Pentagon, while Mohammed pushed for the World Trade Center.[118] Mohammed taught some of the 9-11 hijackers English, familiarized them with various aspects of Western culture so that they could better blend in,[119] and gave them advice on obtaining "clean" passports to enter the United States.[120] The 9-11 attacks were a scaled-down version of Mohammed's brainchild. Despite bin Laden's overall leadership role, Mohammed was essentially the master criminal who directed the underlings. The federal prisons are full of criminal masterminds serving time for playing similar roles in their criminal enterprises.[121]

But even when we narrow our focus to Mohammed, Binalshibh, and Abu Zubaydah (another high-level al Qaeda member in custody who was involved in the 9-11 attacks), there are still barriers ranging from minor to significant in the use of criminal law as the basis for their continued detention. First, as a matter of jurisdiction, federal murder prosecutions would be limited to a subset of all murders, such as murders of federal

---

115. *Id.* at 225, 227, 243–44.

116. *Id.* at 160.

117. *Id.* at 153–54. Mohammed's original plan involved hijacking as many as ten planes, striking the 9-11 targets as well as six others in New York, D.C., California, and Washington, and then landing the tenth plane to deliver a speech denouncing the United States for supporting Israel. *Id.* at 154.

118. *Id.* at 155.

119. *Id.* at 157.

120. *Id.* at 235.

121. *See, e.g.*, United States v. Gigante, 166 F.3d 75, 78 (2d Cir. 1999).

officers and employees;[122] foreign officials, official guests, and "internationally protected persons";[123] or murders within the special maritime and territorial jurisdiction.[124] Although most of the 9-11 victims killed when American Airlines Flight 77 crashed in the Pentagon fit within at least one of those categories (federal employees), most of the other victims do not. Thus, while a federal prosecution is feasible, the United States would be unable to hold Mohammed, Binalshibh, and Zubaydah accountable for all 2,973 murders.[125]

A more serious obstacle to successful prosecution lies in the manner in which the government learned of the roles allegedly played by Binalshibh, Mohammed, Zubaydah and others. Information provided by Mohammed about his role in the 9-11 attacks has likely been procured through un-Mirandized and coercive interrogation. One unnamed U.S. intelligence officer stated shortly after Mohammed's capture that "[s]ome al Qaeda just need some extra encouragement [to talk]," which the officer defined as "a little bit of smacky-face."[126] There was also speculation that the United States could use two of Mohammed's children who were captured along with him as leverage against him.[127] As for Zubaydah, who was reportedly shot three times during the firefight in which he was captured, there are unconfirmed reports that U.S. personnel withheld painkillers from him in an effort to coerce him into giving in-

---

122. 18 U.S.C. § 1114 (2000).

123. *Id.* § 1116.

124. *Id.* § 1111(b).

125. Of course, it seems quite unlikely that it would make a difference in terms of sentencing whether they were convicted of murdering 29 persons or 2,973. But we should not be so quick to discount the symbolic value of criminal prosecutions and convictions. After the 1995 Oklahoma City bombing, the United States prosecuted Timothy McVeigh and Terry Nichols. McVeigh was sentenced to death and Nichols was sentenced to life imprisonment. *See* United States v. Nichols, 169 F.3d 1255, 1260 (10th Cir. 1999); United States v. McVeigh, 153 F.3d 1166, 1179 (10th Cir. 1998). Subsequently, an Oklahoma prosecutor indicted Nichols on 160 counts of first-degree murder. Although one survey showed that a majority of Oklahomans thought that a state trial would be a waste of money, a family member of a victim who favored a trial explained that "[o]ne hundred sixty people never had their loved one's name on an indictment." Ralph Blumenthal, *Another Oklahoma City Bomb Trial, and Still Questions Remain*, N.Y. TIMES, Mar. 16, 2004, at A18.

126. Jess Bravin & Gary Fields, *How Do U.S. Interrogators Make a Captured Terrorist Talk?*, WALL ST. J., Mar. 4, 2003, at B1; *see also* Mark Bowden, *The Dark Art of Interrogation*, ATLANTIC MONTHLY, Oct. 2003, at 51, 53–56.

127. Bravin & Fields, *supra* note 126, at B1.

formation.[128] Further inculpatory evidence against other al Qaeda members might likewise be inadmissible under the rules of evidence. For example, as Ruth Wedgwood noted in arguing in favor of military tribunals, "bin Laden's telephone call to his mother, telling her that 'something big' was imminent, could not be entered into evidence if the source of information was his mother's best friend."[129] Thus, the very acquisition of much of the evidence against detainees disqualifies it from use in traditional criminal prosecution.

The difference between the post-capture treatment of Mohammed and John Walker Lindh is illustrative of the different treatment afforded in a criminal context. Lindh, the so-called "American Taliban," converted to Islam in his teens and later traveled to Afghanistan to fight in support of the Taliban. He was captured by the Northern Alliance in November 2001 and turned over to the United States in December 2001.[130] FBI agents interrogated Lindh, and while there is a question of whether federal agents violated the no-contact ethics rule given that Lindh's father had retained a lawyer to represent Lindh,[131] the agents apparently did advise Lindh of his *Miranda* rights and obtain a waiver of those rights.[132] Even the potential ethical violation was of concern to Department of Justice lawyers, who fretted about whether Lindh's confession to an FBI agent would have to be sealed for use only for national security purposes.[133]

The very paradigm of criminal investigation is largely inapposite in the context of the war on terror. Just as confessions or admissions by detainees obtained through torture or coercion may be inadmissible in criminal trials,[134] physical evidence gathered during the war on terrorism may also be inadmissi-

---

128. *See* Michelle Shephard, *Informant Likely Tortured, Harkat Defence Tells Hearing*, TORONTO STAR, Oct. 27, 2004, at A9 (noting allegation by lawyer for an Algerian man facing deportation by Canada).

129. Ruth Wedgwood, Op-Ed., *The Case for Military Tribunals*, WALL ST. J., Dec. 3, 2001, at A18.

130. United States v. Lindh, 212 F. Supp. 2d 541, 545–47 (E.D. Va. 2002).

131. Michael Isikoff, *The Lindh Case E-Mails*, NEWSWEEK, June 24, 2002, at 8.

132. *See* Richard Willing & Toni Locy, *Case Built on Walker's Own Words*, USA TODAY, Jan. 16, 2002, at A3.

133. Isikoff, *supra* note 131. Lindh's waiver of his *Miranda* rights reduced the issue to whether he was coerced into making the waiver.

134. *See* Rogers v. Richmond, 365 U.S. 534, 540–41 (1961).

ble. Fundamentally, a criminal investigation is usually conducted with an eye toward prosecution:[135] police officers can seek warrants to conduct lawful searches,[136] and they are trained to maintain a chain of custody over evidence that they discover. The chain of custody is procedurally significant because it allows the prosecution to authenticate physical evidence presented at trial.[137]

When soldiers gather physical evidence, however, criminal procedure is not their chief concern. First and foremost, any evidence-gathering by soldiers is incidental to their main purpose, which is to capture or kill enemies of the United States. There may not be time to process the physical evidence.[138] Second, during operations outside the United States, it may not be practicable to seek a search warrant.[139] Third, law enforcement personnel can reasonably expect to be called into court to testify in connection with their investigations, but military soldiers may well be unavailable to testify in criminal trials.[140]

Thus, while some of the likely barriers to murder prosecutions of suspected al Qaeda members are self-imposed, others result from the different operational goals of criminal investigation versus warfare. Regardless of the source of these problems, it seems likely that, absent some new exception to the current constitutional criminal procedure doctrines, the federal government may well be unable to prosecute Mohammed, Binalshibh, and Zubaydah. One could plausibly argue that the government's inability to prosecute these architects of the 9-11

---

135. *But see* Dombrowski v. Pfister, 380 U.S. 479 (1965) (allowing federal court injunction to issue against state prosecution where plaintiffs showed that prosecution was aimed not at securing conviction but rather at harassing them and preventing them from helping vindicate the constitutional rights of African-Americans).

136. One of the advantages to seeking a search warrant is that even if the warrant turns out to be defective, any evidence obtained as a result of the search will not be excluded unless the warrant was facially defective. *See* United States v. Leon, 468 U.S. 897, 922–23 (1984).

137. *See* FED. R. EVID. 901(a).

138. Weapons, for example, might be destroyed outright rather than collected and retained.

139. Admittedly, however, the Court has recognized this problem and held that an alien who has no voluntary connection with the United States has no Fourth Amendment rights relating to searches that take place outside the United States. *See* United States v. Verdugo-Urquidez, 494 U.S. 259, 274–75 (1990).

140. Indeed, U.S. soldiers on active duty may still be posted overseas, or they may have been killed in subsequent military action.

attacks is a simple consequence of constitutional limitations on the government, and that circumventing those limitations in the interest of expediency is a cure worse than the disease. One could thereby conclude that we cannot *punish* these men, regardless of the heinousness of their actions, but it seems to go too far to suggest that we must therefore either release them outright or repatriate them to their home countries and hope that they can be detained there.

The discussion has thus far focused on high-level detainees not held at Guantanamo Bay, whose involvement in the 9-11 attacks is relatively well documented. With other detainees, we may know little more than that they were captured in Afghanistan fighting against the United States or the Northern Alliance and that they continue to see themselves as at war with us. Even if we have more information, such as whether they participated in terrorism training at the al Qaeda camps in Afghanistan, we may have no evidence of their participation in the 9-11 attacks. Indeed, they may have had no role in the 9-11 attacks. As with high-level detainees, we may lack the legal basis to punish them under the criminal law, but that does not necessitate releasing them and thereby permitting them to continue their war against us.

2. Recent Antiterrorism Statutes: 18 U.S.C. §§ 2339A, 2339B

After the 1993 bombing of the World Trade Center, Congress recognized new terrorism-related crimes such as 18 U.S.C. § 2339A. Section 2339A makes it a federal crime to provide "material support or resources" knowing that such material support or resources will be used to carry out terrorist acts.[141] After the 1995 bombing in Oklahoma City, Congress enacted 18 U.S.C. § 2339B. A person who "knowingly provides material support or resources to a foreign terrorist organization" violates section 2339B and faces up to ten years in prison.[142]

As drafted, section 2339A suffers from a severe limitation that cripples its use as an antiterrorism tool. To convict a defendant under this provision, the government must prove that the defendant provided material support or resources that

141. 18 U.S.C. § 2339A(a) (2000).
142. *Id.* § 2339B(a)(1).

were used for a *particular* act of terrorism. The 9-11 Commission described this requirement as "a practical impossibility."[143]

Section 2339B, on the other hand, has been a mainstay of post-9-11 terrorism prosecutions. Notable defendants prosecuted after 9-11 under this provision include John Walker Lindh,[144] the "Lackawanna Six,"[145] and Earnest James Ujaama.[146] Section 2339B has helped federal prosecutors, because its prohibitions are susceptible to broad interpretation; persons who traveled to Afghanistan and trained with al Qaeda can be convicted of providing personnel (themselves) to a designated terrorist group.[147] Yet its potential for broad interpretation has also resulted in a number of successful constitutional vagueness challenges to the statute. For example, in *United States v. Sattar*,[148] the government attempted to prosecute criminal defense attorney Lynne Stewart, among others, for violating section 2339B. Part of the prosecution's theory was that Stewart and her co-defendants provided "material support" in the form of "personnel" when they made themselves available to their client, convicted terrorist Omar Ahmad Ali Abdel Rahman, to communicate his messages to the Islamic Group.[149] Stewart allegedly delivered a statement by her client to the press; the government contended that the statement was a message from Rahman to his terrorist group to commit additional violent acts. The district court dismissed this count of the indictment on vagueness grounds: "[T]hese terms and concepts applied to the prohibited provision of personnel provide no

---

143. *See* John Roth et al., Staff Report to the National Commission on Terrorist Attacks upon the United States, Monograph on Terrorist Financing, at 31–32 (2004) http://www.9-11commission.gov/staff_statements/911_TerrFin_Monograph.pdf. Notwithstanding the 9-11 Commission's interpretation of section 2339A, the United States recently convicted criminal defense attorney Lynne Stewart of violating the provision without any allegation in the indictment of a particular act of terrorism to which the material support (defined as providing Stewart's client access to his followers in Egypt by disseminating a statement) was tied. *See* United States v. Sattar, 272 F. Supp. 2d 348 (S.D.N.Y. 2003).

144. *See* United States v. Lindh, 212 F. Supp. 2d 541 (E.D. Va. 2002).

145. *See* United States v. Goba, 220 F. Supp. 2d 182 (W.D.N.Y. 2002).

146. *See* Mike Carter, *"Unusually Light" Prison Sentence Given to Ujaama*, SEATTLE TIMES, Feb. 14, 2004, at B1.

147. *See, e.g.*, *Goba*, 220 F. Supp. 2d at 194.

148. 272 F. Supp. 2d 348.

149. *Id.* at 354–57. Rahman was convicted in 1995 of a number of crimes, including the 1993 World Trade Center bombing. *Id.* at 354.

notice to persons of ordinary intelligence . . . ."[150] Similarly, in *United States v. Khan*,[151] a district court held that a defendant did not violate section 2339B despite an intent to fight for the Taliban; while the defendant's action would have benefited al Qaeda indirectly, it did not constitute "material support or resources" to al Qaeda.[152]

Recent amendments to section 2339B make clear that Congress intended for the statute to cover situations where a person knowingly "provides" himself to a designated foreign terrorist organization.[153] Even so, section 2339B's usefulness in the war on terrorism is subject to an important jurisdictional limitation built into the statute: The statute's prohibition applies only to conduct "within the United States or subject to United States jurisdiction." Thus, while section 2339B remains a potent tool for federal prosecutors to use against U.S. residents, it does not cover conduct by the Guantanamo detainees, who are being held for acts they allegedly committed in Afghanistan or Pakistan.

3. War Crimes

Strictly speaking, war crimes fall within the ambit of the law of armed conflict, not civilian criminal law. Nonetheless, it is logical to discuss war crimes prosecutions here, because the goal of war crimes prosecutions is to punish those who violate the law of armed conflict. Thus, under U.S. law, enemy persons convicted of war crimes may be imprisoned past the end of hostilities,[154] whereas mere prisoners of war must be repatriated.[155]

On November 13, 2001, President Bush issued an executive order that authorized military tribunals to prosecute appropriate al Qaeda and Taliban suspects "for violations of the laws of war and other applicable laws."[156] In July 2004, the United States had designated fifteen detainees for military prosecu-

150. *Id.* at 359.

151. 309 F. Supp. 2d 789 (E.D. Va. 2004).

152. *Id.* at 821.

153. *See* Pub. L. No. 108-458, §§ 6602–03 (2004); 118 Stat. 3638, 3761–64.

154. *See, e.g.*, Johnson v. Eisentrager, 339 U.S. 763 (1950).

155. *See supra* Part II.A.3.

156. Detention, Treatment, and Trial of Certain Non-Citizens in the War Against Terrorism, 66 Fed. Reg. 57,833 (Nov. 13, 2001) [hereinafter Military Tribunal Executive Order].

tion,[157] the first of which involved the alleged personal driver to Osama bin Laden. A district court blocked the military tribunals from proceeding on the ground that their composition was unconstitutional,[158] but the D.C. Circuit reversed,[159] prompting the detainee to seek certiorari from the Supreme Court. In November 2005, with 450 law professors having signed a letter urging the Court to accept review,[160] the Supreme Court granted certiorari.[161]

In any event, war crimes trials suffer from the same general lack of reach present in criminal trials. For some of the detainees, there may be insufficient evidence to tie them to any past crimes, even though there is evidence that they are members of al Qaeda or the Taliban and that they present a threat of dangerousness. Furthermore, war crimes trials are reserved for those enemy combatants who violate the laws of war. The United States held millions of POWs of the Axis Powers at the end of World War II, but only a few thousand were prosecuted for war crimes. Lack of evidence of past crimes may well preclude prosecution, and hence punishment, but the inability to prosecute does not mean that protective detention is similarly unavailable.

III. NONCRIMINAL DETENTION

The notion that the Guantanamo detainees must either be charged with crimes or released is simply inconsistent with U.S. law. It is true that, as a general rule, a "finding of dangerousness, standing alone, is ordinarily not a sufficient ground on which to justify indefinite involuntary confinement."[162] Yet

157. *See Bush Selects 9 More to Face Tribunals*, KAN. CITY STAR, July 8, 2004, at A8.

158. *See* Hamdan v. Rumsfeld, 344 F. Supp. 2d 152 (D.D.C. 2004). The court did not hold that military tribunals were per se unconstitutional; thus, if the district court's ruling is ultimately upheld, the government would still be able to prosecute detainees in properly constituted military tribunals.

159. *See* Hamdan v. Rumsfeld, 415 F.3d 33, 43 (D.C. Cir. 2005).

160. Statement of 450 Professors of Law on Hamdan v. Rumsfeld (Oct. 26, 2005), http://www.law.yale.edu/outside/html/Public_Affairs/666/professorstatement.pdf; Signatures of October 26, 2005 Letter, http://www.law.yale.edu/outside/html/Public_Affairs/666/stmtsignatures.pdf.

161. 74 U.S.L.W. 3287 (U.S. Nov. 7, 2005) (No. 05-184). *See also* Linda Greenhouse, *Justices to Rule on a Challenge to U.S. Tribunals*, N.Y. TIMES, Nov. 8, 2005, at A1.

162. *See* Kansas v. Hendricks, 521 U.S. 346, 358 (1997).

we confine persons against their will for reasons other than punishment in a variety of circumstances: criminal defendants yet to stand trial are sometimes held without bail, sex offenders and mentally ill persons who pose a danger to society can be civilly committed in appropriate circumstances, and those with virulent and contagious diseases have been forcibly quarantined.[163]

A. Pretrial Detention

The first category of noncriminal detention is pretrial detention of persons awaiting trial. Under the Bail Reform Act of 1984,[164] a federal district court judge is empowered to deny the option of bail and order the detainment of persons facing indictment (but not yet convicted) when, upon a hearing, the judge "finds that no condition or combination of conditions will reasonably assure the appearance of the person as required and the safety of any other person and the community."[165]

In *United States v. Salerno*, the Court held that the Constitution does not categorically prohibit the detention of criminal defendants who are awaiting trial, even though the defendants have not yet been convicted of any crime.[166] The justifications for such detention are that the defendant, having been "indicted or held to answer for a serious crime, presents a demonstrable danger to society,"[167] and that the defendant's liberty interest is outweighed by society's "greater needs."[168] Key to the Court's opinion is the distinction between regulatory and penal objectives. If Congress does not indicate explicitly that it intends for detainment to be punitive, then such detainment may be seen as regulatory if there is an "alternative purpose to which [it] may rationally be connected" and if it is not "excessive in relation to the alternative purpose."[169]

---

163. The purpose here is not to engage in a normative analysis of such detentions, but rather to note first, the circumstances under which detention is constitutionally permissible and, second, the procedural and substantive requirements to effect detention under those specified circumstances.

164. Pub. L. No. 98-473, 98 Stat. 1976 (codified at 18 U.S.C. §§ 3141–3156 (2000)).

165. *Id.* § 3142(e).

166. 481 U.S. 739 (1987).

167. *Id.* at 750.

168. *Id.* at 751.

169. *Id.* at 747 (quoting Kennedy v. Mendoza-Martinez, 372 U.S. 144, 168–69 (1963)).

The legislative history of the Bail Reform Act demonstrates that Congress did not consider pretrial detention to be punitive. The maximum length of pretrial detention is limited by speedy trial requirements,[170] and therefore cannot be indefinite.[171] In addition, pretrial detention contains procedural requirements aimed at protecting the detainee's liberty interest. For example, the defendant facing potential pretrial detention is entitled to the presence of counsel, may testify and present other evidence, and is entitled to cross-examine adverse witnesses.[172]

B. Civil Commitment of the Mentally Ill

The second form of involuntary noncriminal detention is civil commitment of the mentally ill. Constitutionally, a state may use its police powers to protect its citizens "from the dangerous tendencies of some who are mentally ill,"[173] but only when the factual predicates of mental illness and dangerousness are established by at least clear and convincing evidence.[174] Similarly, when a criminal defendant is acquitted by reason of insanity, with mental illness and dangerousness having been determined in the trial, the state is entitled to detain that person in a mental health hospital until the person regains his sanity.[175]

A number of courts have held that indigent persons facing involuntary civil commitment must be appointed counsel to represent them.[176] The detainee's lawyer plays an important

---

170. *See id.* at 747; 18 U.S.C. § 3161 (2000).

171. Moreover, *Salerno* emphasized the milder conditions of detention, including the requirement that pretrial detainees be housed separately from convicted criminals. 481 U.S. 739. In practice, however, that requirement has not been enforced, and therefore it would be difficult to justify pretrial detention on the basis of a difference in the conditions of confinement.

172. 18 U.S.C. § 3142(f) (2000).

173. Addington v. Texas, 441 U.S. 418, 426 (1979); *see also* Donald H.J. Hermann, *Barriers to Providing Effective Treatment: A Critique of Revisions in Procedural, Substantive, and Dispositional Criteria in Involuntary Civil Commitment*, 39 VAND. L. REV. 83, 92 (1986) (noting that most states have adopted rules requiring findings of mental illness and dangerousness).

174. *Addington*, 441 U.S. at 433.

175. *See, e.g.*, Jones v. United States, 463 U.S. 354, 366 (1983).

176. *See e.g.*, Heryford v. Parker, 396 F.2d 393, 396 (10th Cir. 1968); Dorsey v. Solomon, 435 F. Supp. 725, 733 (D. Md. 1977); Lessard v. Schmidt, 349 F. Supp. 1078, 1098 (E.D. Wis. 1972); *see also* Specht v. Patterson, 386 U.S. 605, 610 (1967).

and often complicated role in the commitment proceedings: "The attorney's role is to challenge the basis of the application for involuntary admission and elicit facts and opinions challenging the need for hospitalization. In addition, the attorney should advocate that a less restrictive form of intervention is consistent with the welfare and safety of the individual."[177]

C. Civil Commitment of Sex Offenders

A number of states currently have laws that permit the government to seek civil commitment of sex offenders who suffer from mental abnormalities or personality disorders that render them likely to engage in future violent sex crimes.[178] These laws permitting civil commitment of violent sex offenders arose largely during the 1990s and differ from earlier such laws in that the more recent crop is justified in terms of preventative detention, rather than rehabilitation.[179]

Kansas's sexual predator law, at issue in *Kansas v. Hendricks*,[180] provides an example of this sort of law. The state must satisfy two substantive requirements in order to commit a person under this law. First, the person facing commitment must have been convicted of a violent sexual offense, or been found not guilty of a violent sexual offense by reason of insanity or mental defect or disease, or been charged with a violent sexual offense but found incompetent to stand trial.[181] Second, the person must have a mental abnormality or personality disorder, which is defined as a "congenital or acquired condition affecting the emotional or volitional capacity which predisposes the person to commit sexually violent offenses in a degree" that makes the person a "menace to the health and safety

177. Donald H. Stone, *Giving a Voice to the Silent Mentally Ill Client: An Empirical Study of the Role of Counsel in the Civil Commitment Hearing*, 70 U. MO. KAN. CITY L. REV. 603, 614 (2002).

178. *See, e.g.*, ARIZ. REV. STAT. ANN. §§ 36-3701, -3702, -3704 to -3709, -3712 (2003); CAL. WELF. & INST. CODE §§ 6600–6608 (2001); KAN. STAT. ANN. §§ 59-29a01 to -29a15 (2004); MINN. STAT. ANN. § 253B.185 (2003); N.J. STAT. ANN. § 30:4-82.4 (1997); WASH. REV. CODE ANN. §§ 71.09.010–.092 (2002); WIS. STAT. ANN. §§ 980.01–.13 (1998).

179. *See, e.g.*, Eric S. Janus, *Sex Offender Commitments: Debunking the Official Narrative and Revealing the Rules-in-Use*, 8 STAN. L. & POL'Y REV. 71, 72 (1997); Peter C. Pfaffenroth, *The Need for Coherence: States' Civil Commitment of Sex Offenders in the Wake of* Kansas v. Crane, 55 STAN. L. REV. 2229, 2233–42 (2003).

180. 521 U.S. 346 (1997).

181. KAN. STAT. ANN. §§ 59-29a03(a), 22-3221 (2004).

of others."[182] If those conditions are proven to the court, the person can be ordered into confinement subject to further periodic review.

A person faced with civil commitment as a violent sexual predator, like one who faces civil commitment or pretrial detention, has the right to counsel, as well as the rights to call on witnesses, to cross-examine the government's witnesses, and to review documents.[183] In addition, an indigent detainee has the right to an examination by mental health professionals and the right to the assistance of counsel at public expense.[184] Despite the procedural protections, the net effect of these statutes is to provide a mechanism for what is "essentially indefinite" detention during the time that the detainee continues to suffer from a mental defect, which may be congenital, that renders him or her dangerous.[185]

D. Medical Quarantines

Although quarantines have been long accepted as a public health measure to control the spread of infectious diseases, they have not been used on a widespread basis in the United States for at least eighty years.[186] In an old case that has not been revisited in the last hundred years, the Court upheld a Louisiana board of health decision that forbade the *S.S. Britannia*, carrying 408 immigrants from Italy, to dock at major ports in the state.[187] The ship owners argued that the decision vio-

---

182. *Id.* § 59-29a02(b).

183. *Hendricks*, 521 U.S. at 353 (citing KAN. STAT. ANN. § 59-29a06–07 (1994)).

184. *Id.* (citing KAN. STAT. ANN. § 59-29a06 (1994)).

185. Eric S. Janus, *The Use of Social Science and Medicine in Sex Offender Commitment*, 23 NEW ENG. J. ON CRIM. & CIV. CONFINEMENT 347, 349 (1997).

186. *See* Joseph Barbera et al., *Large-Scale Quarantine Following Biological Terrorism in the United States: Scientific Examination, Logistic and Legal Limits, and Possible Consequences*, 286 J. AM. MED. ASS'N 2711, 2712 (2001). The possibility of an outbreak of pandemic influenza, especially in light of the current "avian flu" crisis, has the White House at least talking about quarantines as among potentially necessary steps. *See* HOMELAND SECURITY COUNCIL, NAT'L STRATEGY FOR PANDEMIC INFLUENZA 8 (Nov. 2005) http://www.whitehouse.gov/homeland/nspi.pdf ("In support of our containment strategy, we will . . . [w]here appropriate, use governmental authorities to limit non-essential movement of people, goods and services into and out of areas where an outbreak occurs.").

187. Compagnie Francaise de Navigation a Vapeur v. State Bd. of Health, 186 U.S. 380, 382 (1902). The state board also ruled that if the *Britannia* were to try to dock in a place adjacent to the listed ports, that area would also be declared off-limits. *Id.* at 382.

lated the Fourteenth Amendment by depriving them of their property without due process of law, a contention that the Court rejected on the ground that it, if accepted, would "strip the government . . . of all power to enact regulations protecting the health and safety of the people."[188] In a subsequent case upholding a Massachusetts mandatory vaccination law, the Court reasoned that "in every well-ordered society . . . the rights of the individual in respect of his liberty may at times, under the pressure of great dangers, be subjected to such restraint . . . as the safety of the general public may demand."[189] The quarantine cases that have arisen since these early cases have generally been restricted to special circumstances, such as the quarantining in prisons of prisoners who suffer from AIDS[190] or of persons who suffer from tuberculosis.[191]

Under federal law, the Surgeon General may enforce quarantines to prevent the "spread of communicable diseases from foreign countries into the States . . . or from one State . . . into any other State."[192] Detention of any persons under this statute must, however, be limited to "such time and in such manner as may be reasonably necessary"[193] for "the purpose of preventing the introduction, transmission, or spread of such communicable diseases" specified by presidential executive orders.[194]

As it is believed that al Qaeda has sought biological, chemical, radiological, and nuclear weapons for use against the United States,[195] the possibility of a breakout of a lethal infectious disease such as smallpox cannot be discounted. In the event of an outbreak of smallpox, the Centers for Disease Con-

---

188. *Id.* at 393.

189. Jacobson v. Massachusetts, 197 U.S. 11, 29 (1905).

190. *See e.g.*, Harris v. Thigpen, 941 F.2d 1495, 1512 (11th Cir. 1991); *see also* Jicarilla Apache Tribe v. United States, 601 F.2d 1116, 1135 (10th Cir. 1979) (upholding the power of states to enforce quarantine laws on Indian reservations).

191. *See e.g.*, Moore v. Draper, 57 So. 2d 648 (Fla. 1952).

192. 42 U.S.C. § 264(a) (2000).

193. *Id.* § 264(d)(1).

194. *Id.* § 264(b). The current list of such communicable diseases includes cholera, diphtheria, tuberculosis, plague, smallpox, yellow fever, viral hemorraghic fevers (such as Ebola and Marburg), and SARS. Exec. Order No. 13,295, 68 Fed. Reg. 17,255 (Apr. 4, 2003).

195. *See, e.g.*, Rahimullah Yusufzai, *Conversation with Terror*, TIME, Jan. 11, 1999, at 39 (statement of Osama bin Laden) ("Acquiring weapons for the defense of Muslims is a religious duty. If I have indeed acquired these [chemical and nuclear] weapons, then I thank God for enabling me to do so.").

trol (CDC) call for "ring vaccinations" of the exposed population, coupled with "concentric levels of quarantine to restrict movement of individuals and conveyances between communities."[196] The goal is to vaccinate those persons who may have been exposed to smallpox to protect them from contracting it, while isolating those who have begun to exhibit the symptoms of the disease until they are no longer contagious.[197] Generally, a person who contracts smallpox is contagious from the time that he develops a rash (twelve to fourteen days after exposure) to the time that the scabs resulting from lesions separate from the skin (five to six weeks after exposure), although the degree of infectiousness decreases once the lesions scab.[198] Thus, in the event of an outbreak of smallpox, an infected person could face quarantine of up to a month.

Such quarantine is, of course, different from imprisonment. Quarantined persons would still be allowed caregivers so long as the caregivers could be vaccinated against smallpox.[199] Moreover, in the event that an entire community was quarantined, there would be no need to isolate persons in their own homes. Still, despite these relatively humane touches, quarantine is a liberty-restricting tool in which the crucial decisions about whom and how long to isolate are made by medical professionals, not judicial personnel.

IV. NONCRIMINAL DETENTION IN THE WAR ON TERRORISM

As the previous Part demonstrated, noncriminal detention arises in a number of different situations where the detainee is not criminally culpable but does pose some potential danger to society. The danger is tied to some specific trigger, such as a mental defect, indictment, or infectious disease, as opposed to a

196. CENTERS FOR DISEASE CONTROL SMALLPOX RESPONSE PLAN AND GUIDELINES, GUIDE C, PART 2: QUARANTINE GUIDELINES 7 (2003), http://www.bt.cdc.gov/agent/smallpox/response-plan/files/guide-c-part-2.pdf.

197. *See* CENTERS FOR DISEASE CONTROL SMALLPOX RESPONSE PLAN AND GUIDELINES, GUIDE B: VACCINATION GUIDELINES FOR STATE AND LOCAL AGENCIES 5–6 (2003), http://www.bt.cdc.gov/agent/smallpox/response-plan/files/guide-b-part1of3.pdf ("Vaccination within 3 days of exposure will prevent or significantly lessen the severity of smallpox symptoms in the vast majority of people.").

198. *See* CENTERS FOR DISEASE CONTROL SMALLPOX RESPONSE PLAN AND GUIDELINES, EXECUTIVE SUMMARY 1–2 (2003), http://www.bt.cdc.gov/agent/smallpox/response-plan/files/exec-sections-i-vi.pdf.

199. *See id.* at 4.

general propensity for criminality. Without such limiting factors, the utilitarian calculus of protection of society would overwhelm individual liberties in virtually every instance where the government could prove that an individual poses a danger to society; thus, the police are not empowered to round up and detain gang members indefinitely, even though they may well present a danger to their local community. The limiting factor also provides a basis from which to determine the end of the justification for detention, which occurs when the person no longer suffers from that limiting factor (even if the person remains dangerous). Thus, the Constitution does not forbid noncriminal detention when there is some appropriate limiting factor (or trigger) and the detainee is shown to be a danger to society. These two requirements—limiting factor and dangerousness—help to articulate a model for detention of nonstate actors captured during a military conflict.

A. War or Military Force as the Limiting Factor

1. Defining the Class of Persons to be Detained

The limiting factor with respect to the Guantanamo detainees is the status of the military conflict itself. First, it is important to distinguish the rhetoric of the "war on terrorism" from the congressional authorization for military force (and the attendant dimensions of that military conflict). A war (or military conflict) with the objective of defeating "terrorism" is, as many others have noted, of potentially never-ending duration, and it is against a relatively large class of undefined persons. On the other hand, the Authorization for Use of Military Force that Congress passed on September 14, 2001, specifies that the President may "use all necessary and appropriate force against those nations, organizations, or persons he determines planned, authorized, committed, or aided the terrorist attacks that occurred on September 11, 2001, or harbored such organizations or persons . . . ."[200] Whatever other targets may be encompassed when the President refers to the global war on terrorism, the entities that Congress has designated to be the subject of military force are limited to those that either played a

200. Authorization for Use of Military Force, Pub. L. No. 107-40, § 2(a), 115 Stat. 224, 224 (2001).

role in planning or carrying out the 9-11 attacks or sheltered those responsible.[201] We may also be in the position of not knowing when *this* war is over, for it is possible for us to capture or kill every person who fits within the statutory definition of the "enemy," but that is a different problem from not knowing when we have defeated a *concept* such as terrorism.

The current war on terrorism, as defined by the Military Force Authorization, thus differs from other rhetorical wars, such as the "war on poverty" and the "war on drugs." The war on poverty involved no use of military force or assets, so we need not spend any more time analyzing it.[202] The war on drugs, on the other hand, presents a more challenging question, since U.S. military forces have been involved in interdicting the illegal drug trade. As early as 1981, Congress relaxed the Posse Comitatus Act,[203] which forbids the military from engaging in domestic law enforcement, to allow the Department of Defense to provide indirect assistance to domestic law enforcement agencies in the form of information, training, equipment, and advice.[204] In 1989, Congress directed the Department of Defense to further assist local authorities in stopping the flow of illegal drugs.[205] That same year, President Bush "authorized sending a small number of U.S. Special Forces troops to Colombia to train its police and military in rapid-strike tac-

---

201. *See* Curtis A. Bradley & Jack L. Goldsmith, *Congressional Authorization and the War on Terrorism*, 118 HARV. L. REV. 2047, 2107–09 (2005). *But see* Ryan Goodman & Derek Jinks, *Replies to Congressional Authorization: International Law, U.S. War Powers, and the Global War on Terrorism*, 118 HARV. L. REV. 2653, 2654–58 (2005) (arguing that Bradley and Goldsmith overstate the category of persons subject to the congressional military force authorization).

202. John Hart Ely has commented on whether Congress would do something as outrageous as authorize the use of military force against an abstract concept such as poverty: "It is an entirely legitimate response . . . to note that it couldn't pass and refuse to play any further. In fact it can only deform our constitutional jurisprudence to tailor it to laws that couldn't be enacted . . . ." JOHN HART ELY, DEMOCRACY AND DISTRUST: A THEORY OF JUDICIAL REVIEW 183 (1980). *But see* Luban, *supra* note 5, at 222 (discussing a hypothetical in which members of the Mafia are declared enemy combatants).

203. Ch. 1041, § 18(a), 70A Stat. 626 (1956) (codified as amended at 18 U.S.C. § 1385 (2000)).

204. *See* 10 U.S.C. §§ 371–75 (2000); *see also* Peter M. Sanchez, *The "Drug War": The U.S. Military and National Security*, 34 A.F. L. REV. 109, 122 (1991) (suggesting that Congress merely "legalize[d] an existing practice").

205. *See* National Defense Authorization Act, Fiscal Year 1989, Pub. L. No. 100-456, § 1104, 102 Stat. 1918, 2042–47 (1988) (codified at 10 U.S.C. §§ 371–80).

tics."[206] While there has yet to be an instance of the U.S. military capturing and detaining persons as enemy combatants in the war on drugs, the United States did use military force in Operation Just Cause to capture General Manuel Noriega of Panama and bring him to stand trial on his indictment for drug trafficking and money laundering. The assault, which started on December 20, 1989, involved 24,000 American soldiers tasked with capturing Noriega, overthrowing Noriega's party (the Panamanian Defense Force), and protecting American lives and property in Panama;[207] by January 3, 1990, it was over, with Noriega in military custody on his way to be handed over to the Drug Enforcement Agency.[208] A federal jury subsequently convicted Noriega, and he received a sentence of forty years.[209] Still, the Noriega example is not one where the law of armed conflict was used to justify *military* detention of a person captured in the "war on drugs"; rather, it was a situation of using military force to apprehend a person, who was then brought to stand trial in federal court.

The Military Force Authorization is, to be sure, something of an imperfect trigger: It broadly defines its key term—"enemy"—as "those . . . organizations [that] planned, . . . committed, or aided" the 9-11 attacks, as well as "nations, organizations, or persons who harbored" them.[210] An executive branch, bent on expanding the scope of the war on terrorism, could exploit that broadness to include persons who do not reasonably fit within the statutory definition. Consider, for example, that President Bush's order establishing military tribunals to prosecute al Qaeda suspects for war crimes also allows for their use against

---

206. MARK BOWDEN, KILLING PABLO: THE HUNT FOR THE WORLD'S GREATEST OUTLAW 64–65 (2001).

207. *See* PETER HUCHTHAUSEN, AMERICA'S SPLENDID LITTLE WARS: A SHORT HISTORY OF U.S. MILITARY ENGAGEMENTS: 1975–2000, at 117, 121 (2003); *see also* Sanchez, *supra* note 204, at 134 & n.235.

208. HUCHTHAUSEN, *supra* note 207, at 126. In August of that same year, U.S. Special Forces soldiers were prepared to storm a Panamanian home where Colombia drug lord Pablo Escobar was rumored to be hiding; like Noriega, Escobar was under federal indictment. The plan called for Delta Force soldiers to capture Escobar and turn him over to DEA agents. However, the assault was called off when the military determined that Escobar was still in Colombia. *See* BOWDEN, *supra* note 206, at 65.

209. *See* United States v. Noriega, 117 F.3d 1206, 1210 (11th Cir. 1997).

210. Authorization for Use of Military Force, Pub. L. No. 107-40, § 2(a), 115 Stat. 224, 224 (2001).

persons who have "engaged in, aided or abetted, or conspired to commit, acts of international terrorism, or acts in preparation therefor, that have caused, threaten to cause, or have as their aim to cause, injury to or adverse effects on the United States, its citizens, national security, foreign policy, or economy."[211] The class of persons who face potential military tribunals under President Bush's executive order is thus significantly broader than the class of persons fitting within the terms of Congress's Military Force Authorization.

On the other hand, the definition might also be too restrictive, in that al Qaeda may, at this point, be little more than a symbolic threat to us; the real danger may come from various Islamic terrorist groups that are not technically part of al Qaeda but that share the same goals.[212] A strict reading of the Military Force Authorization might well preclude the President from using military force against such groups, even though the members of such groups may have similar training and intentions as members of al Qaeda. This latter problem is troubling, but at least there is a clear solution: Unless an attack by such groups is imminent, in which case the President could respond by invoking the "repel doctrine," the President should go to Congress and seek a broader authorization for the use of military force that would encompass such groups.

Unlike civil commitment of the mentally ill or of violent sexual predators, where the detainee is unable to control his dangerous tendencies, an enemy combatant does not necessarily lack volitional control over his actions. That is, the mentally ill person facing commitment has been selected for the commitment proceeding because of his mental illness, which in turn is what causes him or her to be dangerous in a way that he cannot control. The military conflict, on the other hand, need not

---

211. Military Tribunal Executive Order, *supra* note 156, at 57834.

212. *See* Syed Saleem Shahzad, *The Remaking of Al-Qaeda*, ASIA TIMES, Feb. 25, 2005, *available at* http://www.atimes.com/atimes/South-Asia/GB25Df04.html. This may be a more serious problem than was initially apparent, given the mutation of al Qaeda after the destruction of its training bases in Afghanistan. Al Qaeda may exist more as an inspirational entity than an operational one; yet, as other violent terrorist groups seek to replicate al Qaeda's goals, might the President charge that such other groups are "part" of al Qaeda? Consider, for example, that Abu Musab al-Zarqawi, the leader of the insurgency in Iraq, declared his loyalty to Osama bin Laden and al Qaeda in late 2004. *Al-Zarqawai Group Threatens to Kill Japanese Hostage*, CNN.com, Oct. 27, 2004, *available at* http://www.cnn.com/2004/WORLD/meast/10/16/Iraq.main.

cause the detainee to be dangerous in a way that he cannot control; indeed, the Guantanamo detainees, if they are members of al Qaeda, voluntarily joined that organization. Thus, war or authorization of military force might not seem to be a trigger analogous to other triggers of noncriminal detention, which incapacitate those who cannot control their own dangerousness (as opposed to those who choose not to do so).

There are two responses to that difference. First, not all types of noncriminal detention require that the limiting factor actually cause the detainee to lose volitional control; pretrial detention, for example, is predicated upon the detainee's indictment for a serious crime, and yet the indictment itself is not even admissible in trial as evidence of wrongdoing and certainly does not cause the defendant to become dangerous. Second, under the law of armed conflict as applied to traditional nation-state conflicts, captured enemy soldiers can be detained because of their status as members of enemy forces; the implicit assumption is that their status as the enemy makes them dangerous, either because of their national loyalty to or coercion by their home government, or both.[213]

Another issue to consider is that this trigger would not foreclose the executive branch from being able to detain persons captured on U.S. soil, so long as it could plausibly classify them as persons who "planned, authorized, committed, or aided the terrorist attacks that occurred on September 11, 2001, or harbored such organizations or persons." The persons who "committed" the attacks are, of course, dead, and those who "planned" the attacks are either dead (the suicide hijackers) or detained (Khalid Sheikh Mohammed, Ramzi Binalshibh) or obvious high-level al Qaeda members still at large (Osama bin Laden). However, there may well be persons who "aided" the 9-11 attacks in this country.[214]

Like the 9-11 hijackers, the German saboteurs in *Ex parte Quirin*[215] were helped by at least fourteen Americans—friends

213. Indeed, that presumption explains why we do not need to hold hearings on whether enemy POWs are dangerous.

214. 9-11 REPORT, *supra* note 24, at 215 ("We believe it unlikely that [Nawaf al] Hazmi and [Khalid al] Mihdhar—neither of whom . . . had any prior exposure to life in the West—would have come to the United States without arranging to receive assistance from one or more individuals informed in advance of their arrival.").

215. 317 U.S. 1 (1942).

or family members who served as contacts—offered the use of their homes as hideouts, and concealed illicit cash brought by the saboteurs for use in their mission.[216] The federal government convicted six of the fourteen, obtaining death sentences against two men and twenty-five-year prison sentences against four women.[217] But the fact that the civilian criminal justice system was capable of handling these aiders and abettors need not have foreclosed the possibility of using the law of armed conflict to handle them.

Finally, it is not obvious that requiring that detention occur pursuant to authorized use of military force actually imposes a significant obstacle to the President's ability to detain persons indefinitely. Recent history, which one might think of as coincident with the rise of congressional "authorization to use military force" in place of formal declarations of war, suggests that the President is able to persuade, cajole, or otherwise coerce Congress into giving him the authority to use military force.[218] What if the President were to seek to militarize further the "war on drugs" such that U.S. military forces would be involved in fighting, killing, or capturing persons involved in producing cocaine to be smuggled into the United States? Could the President order that Colombian drug lords be detained as "enemy combatants"?

If Congress were to authorize the President to use military force against a defined set of drug lords, then the military would presumably be entitled to detain persons captured during ensuing military conflicts. To be sure, the potential for such an outcome might give one pause about the thesis concerning military force authorizations as the trigger for detentions. However, the discomfort lies in the political determination that such nonstate actors—drug lords in this hypothetical—must be dealt with through military force, rather than extradition and prosecution. That is, one may question seriously whether drug

---

216. *See* ALEX ABELLA & SCOTT GORDON, SHADOW ENEMIES: HITLER'S SECRET TERRORIST PLOT AGAINST THE UNITED STATES 222–28 (2002).

217. *Id.* at 229. The convictions were reversed on appeal on the ground that the defendants' confessions were obtained involuntarily. *See* United States v. Haupt, 136 F.2d 661, 671 (7th Cir. 1943). The government successfully retried one of the six (the father of one of the eight saboteurs); two others pleaded guilty to lesser charges. ABELLA & GORDON, *supra* note 216, at 231.

218. Use of Force in Iraq (2003): House, 296-133; Senate, 77-23; Use of Force in Afghanistan (2001): House, 420-1; Senate, 98-0.

lords are appropriate targets for military action.[219] After all, if they are, then presumably the military would be able to capture and detain domestic drug dealers (who would be the equivalent of foot soldiers).[220] The prevention of such a spectacle lies in Congress's political accountability, rather than any hard and fast rule restricting the scope of military force to nation-states.[221] The Constitution commits the power to declare war to Congress, and the determination of whether particular nonstate actors (whether al Qaeda or drug lords) should be dealt with through military force is therefore left to the political

---

219. *But see* DEFENSE POLICY PANEL & INVESTIGATIONS SUBCOMM. OF THE H.R. COMM. ON ARMED SERVS., 100TH CONG., REPORT ON NARCOTICS INTERDICTION AND THE USE OF THE MILITARY: ISSUES FOR CONGRESS 8 (Comm. Print 1988) (quoting Congressmen Les Aspin and Bill Nichols as saying that "international drug trafficking is a national security matter"); BOWDEN, *supra* note 204, at 65 (quoting President George H.W. Bush as saying that "[t]he rules have changed," and that "[w]hen requested, we will for the first time make available the appropriate resources of America's armed forces"); Michael Isikoff, *Drug Plan Allows Use of Military*, WASH. POST, Sept. 10, 1989, at A20 (noting Presidential National Security Decision Directive stating that illegal drugs are a "national security" threat).

220. Nor would the Posse Comitatus Act be a barrier to such a result, for that law only forbids the military from being used for civil law enforcement purposes. *See* 18 U.S.C. § 1385 (2000). In this hypothetical, the military would be fighting the "enemy," who just happened to be on U.S. soil. Certainly, if the United States were invaded by a foreign military force, the Posse Comitatus Act would not bar the Marines from acting to capture or kill the invaders.

221. Under international law (as opposed to domestic law), there may not be a barrier to the invocation of the law of armed conflict against nonstate actors. *See, e.g.*, Michael Novak, *Just Peace and the Asymmetric Threat: National Self-Defense in Uncharted Waters*, 27 HARV. J. L. & PUB. POL'Y 817, 828–29 (2004) ("[T]here have occasionally been precedents for making war on nonstate actors, such as organizations of pirates preying upon international shipping lanes."); Yoo & Ho, *supra* note 14, at 207–15 (arguing that al Qaeda's actions on September 11, 2001 were acts of war); John Yoo, *Transferring Terrorists*, 79 NOTRE DAME L. REV. 1183, 1197 (2004) (arguing that the four Geneva Conventions "make it plain that the laws of armed conflict may apply to intense levels of hostilities conducted by a nonstate actor," and that "Al Qaeda members do not escape the laws of war because they are nonstate actors"). *But see* AMNESTY INT'L, YEMEN: THE RULE OF LAW SIDELINED IN THE NAME OF SECURITY 23 (2003) ("Under existing international humanitarian law, it is not possible to have an international armed conflict between a state on the one hand and a nonstate actor on the other, should the armed group not form part of the armed forces of a Party to the Geneva Conventions."), *available at* http://web.amnesty.org/library/pdf/MDE310062003ENGLISH/$File/MDE3100603.pdf; Jordan J. Paust, *Post-9/11 Overreaction and Fallacies Regarding War and Defense, Guantanamo, the Status of Persons, Treatment, Judicial Review of Detention, and Due Process in Military Commissions*, 79 NOTRE DAME L. REV. 1335, 1342 (2004) ("[A]ny conflict between the United States and al Qaeda as such cannot amount to war or trigger application of the laws of war.").

branches. Indeed, any bright-line rule restricting the use of military force to nation-states, as opposed to nonstate actors in appropriate circumstances, would be not only unworkable but probably also unconstitutional. Consider the often-cited *Prize Cases*,[222] in which the Court held that when the President recognizes a threat grave enough to the nation to require immediate military action to repel, the President is entitled—indeed, obligated—to act even in the absence of congressional authorization. It hardly makes sense to require the President to ascertain that an imminent attack is indeed brought by a nation-state, as opposed to nonstate actors, before repelling it. One might argue that repelling an armed attack by using military force is not the same as authorizing the use of military force, and that military force should be authorized against only nation-states, even if there would be an exception to allow such force to be used against nonstate actors to repel an imminent attack. But that argument is ultimately incoherent: If a group of nonstate actors has mounted an armed attack sufficient to warrant repulsion through military force, why should Congress be unable to authorize the President to pursue the group and destroy it before it mounts another attack?

2. *Application to the War on Terrorism*

In the domestic examples of noncriminal detention, one way in which a person may escape detention is to show that the limiting factor is no longer applicable to him or her. Thus, a person who recovers from a mental illness may seek release from a mental hospital, and a criminal defendant who is no longer under indictment (through either acquittal at trial or dismissal of the indictment) must be released from pretrial detention. In the latter example, detention is not permissible even if the government could show that the defendant remains dangerous.[223]

Because the limiting factor in the terrorism circumstance is (generally) membership in an organization deemed by Congress to be belligerent to the United States, a detainee should theoretically be able to show that the limiting factor no longer applies to him because he has renounced membership in that

222. 67 U.S. 635, 668 (1862).

223. Imagine, for example, that the judge dismisses the indictment due to expiration of the statute of limitations or suppression of key evidence due to a constitutional violation.

organization. Because of the nature of this particular issue, the burden should be placed on the detainee to make the showing. Unfortunately, it is nearly impossible to detail how the detainee would make such a showing beyond noting that the adjudicative factfinder would have to made a judgment on the believability of the detainee's renunciation of the organization. One might expect willingness to take a sworn oath and explanation of the detainee's charge of heart to be among the factors evaluated by the factfinder.[224] The appropriate forum in which to allow the detainee the opportunity to make this showing would be the Combatant Status Review Hearings, since those hearings are aimed at determining the detainee's status, in which the detainee would be alleging a change.

A more troubling situation arises where the detainee was incorrectly classified as a combatant at the outset and that mistake is not rectified until after potentially brutal conditions have enraged or radicalized the detainee. While one might be uncomfortable releasing such a person, he simply did not—and does not—fit within the scope of the persons, organizations, or nations deemed by Congress to be appropriate targets of military force. The closest analogy in support of continued detention of the person would consider what would be done in the ordinary criminal justice system with such a person. If this wrongfully detained individual out of anger and frustration from continued mistaken confinement commits a crime while in prison, the subsequent crime would be sufficient to justify continued detention. Absent such a derivative offense, however, the detainee, while angry and dangerous, cannot be detained further on that ground alone, as dangerousness alone is not enough to warrant involuntary detention.

B. Proving Dangerousness

The examples of noncriminal detention all require the government to prove dangerousness. That dangerousness can be proven in different ways depending on the context and justification for detention. According to the Court, conviction of a

224. While it might be difficult for a detainee to make this showing, a detainee is much more likely to be released under the second prong (dangerousness). The only instance in which this problem is likely to arise as a practical matter would be the unlikely event in which a detainee still wanted to inflict harm on the United States but no longer wanted to be a member of al Qaeda.

felony, including nonviolent felonies, is one way of proving dangerousness.[225] In the situation of pretrial detention, Congress specified that dangerousness would be determined based on factors such as "the nature and seriousness of the charges, the substantiality of the Government's evidence against the arrestee, the arrestee's background and characteristics, and the nature and seriousness of the danger posed by the suspect's release."[226] There is, of course, no guarantee of proving dangerousness with absolute certainty, but one endures that degree of uncertainty about the provability of future dangerousness in other circumstances. As Justice Stevens noted in 1976, "[t]he fact that such determination [of dangerousness] is difficult, however, does not mean that it cannot be made."[227] Not only do federal judges have to make findings on future dangerousness in deciding whether to grant bail, so also do state judges presiding over civil commitment proceedings.[228] In addition, in some death penalty jurisdictions, juries at the penalty phase are charged with determining, as a potential aggravating factor, "whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society."[229]

This is not to say that there are no problems with these other instances of factfinders who must determine future dangerousness. There is a reason to believe, for example, that psychiatric experts cannot correctly predict the future dangerousness of death penalty defendants better than one time in three,[230] which is good for baseball hitters but not for the criminal justice system. To minimize the negative consequences of the uncertainty, the system often relies upon the burden of proof and the standard of proof; thus, for both pretrial detention and civil commitment, the court must find dangerousness by clear and

225. Lynch v. Overholser, 369 U.S. 705, 714 (1962).

226. United States v. Salerno, 481 U.S. 739, 742–43 (1987) (citing 18 U.S.C. § 3142(g) (1982)).

227. Jurek v. Texas, 428 U.S. 262, 274–75 (1976).

228. *See supra* Part III.B–C.

229. *See Jurek*, 428 U.S. at 269 (quoting TEX. CODE. CRIM. PROC. ANN. art. 37.031 (Supp. 1975–1976)).

230. *See* Barefoot v. Estelle, 463 U.S. 880, 900 & n.7 (1983) (citing sources).

convincing evidence.[231] Nevertheless, it cannot be denied that the system is imperfect.

Still, our nation's historical experience with repatriating German soldiers after the end of World War II suggests that it is not impossible to distinguish enemy soldiers who are hardened enemies from ones who are not. In 1945, with the end of the war in sight, the U.S. War Department determined that "German noncommissioned officers and incorrigible Nazis will be the last to be released from a prisoner of war status."[232] In an eerie parallel to current circumstances, activists in 1945 demanded that the United States immediately repatriate German POWs, a position rejected by the American military.[233]

Some of the relevant factors in determining dangerousness in this context of detention of suspected al Qaeda and Taliban fighters might include, nonexclusively, the amount of terrorism training that the detainee has received, the detainee's expressed willingness to engage in terrorism and the type of terrorist attack contemplated, and the prior violent conduct of the detainee. The 9-11 attacks demonstrated how the level of terrorist training can greatly increase a terrorist's dangerousness; for example, the leaders of the 9-11 hijackings had received training on blending into American society to better evade detection, which enabled them to engage in pre-9-11 plotting, such as testing airport and airline security. Similarly, the more devastating a terrorist attack in which a detainee is willing to take part (regardless of whether such an attack has reached the stage of a criminal attempt), the more likely it should be to conclude that the detainee is dangerous and will remain dangerous for a longer period of time. Prior violent conduct, even if not a war crime, would be further evidence of dangerousness, though overreliance on prior violent conduct would probably

---

231. 18 U.S.C. § 3142(f) (2000) (stating clear and convincing evidence as standard for pretrial detention); Addington v. Texas, 441 U.S. 418, 433 (1979) (holding clear and convincing evidence as standard for civil commitment proceedings).

232. *See* KRAMMER, *supra* note 96, at 231 n.11 (quoting Memorandum for Planning Comm., War Dep't., Status and Employment of Italian and German PWs on Collapse of Germany, (Feb. 19, 1945)).

233. *See id.* at 230–32.

result in over- and underinclusiveness.[234] A final relevant factor might be the age of the detainee.[235]

Additionally, we might ask detainees to take an oath not to attack the United States or American interests abroad. Some detainees may agree to do so and others may refuse. Unwillingness to take the oath would be almost (if not absolutely) dispositive evidence of continued dangerousness.[236] On the other hand, a detainee's willingness to take the oath would be evidence, though not conclusive, of nondangerousness, as a detainee might take the oath with no intention of honoring it. Presumably, the government would also be able to gather helpful evidence from monitoring conversations between detainees at Camp Delta.

Finally, experts may have a role in the dangerousness hearings. One terrorism expert sees commonalities in the way that terrorists view the world: "They are extremely sensitive to slights and humiliations inflicted on themselves or on members of social groups to which they belong or with which they identify themselves."[237] Another terrorism researcher believes that terrorists tend to share seven characteristics: (1) "oversimplification of issues"; (2) "frustration about an inability to change society"; (3) "a sense of self-righteousness"; (4) "a utopian belief in the world"; (5) "a feeling of social isolation"; (6) "a need to assert his own existence"; and (7) "a cold-blooded willingness to kill."[238] A third commentator observes that "Islamic extremists mount their jihad, construed as self-defense against tyrannical, decadent infidels who seek to enslave the Muslim

---

234. *See* Paul H. Robinson, *Punishing Dangerousness: Cloaking Preventative Detention as Criminal Justice*, 114 HARV. L. REV. 1429, 1450 (2001). Robinson was concerned with prior record, which is not exactly the same as prior violent conduct; prior criminal conduct has further predictive value because criminals have "'thumb[ed their] nose[s]' at the justice system." *Id.* at 1436.

235. *Id.* at 1451 ("Evidence suggests that criminality is highly age-related.").

236. In rare circumstances, a detainee might distrust the U.S. government and fear that taking the oath would be tantamount to confessing to some misdeed, but every effort should be made to persuade the detainee that the oath is only to agree to refrain from attacking the United States.

237. Paul B. Davis, *The Terrorist Mentality*, in POWERWEB: VIOLENCE AND TERRORISM, 36, 36 (Thomas J. Badey ed., 6th ed. 2003) (citing Frederick Hacker, *Dialectical Interrelationships of Personal and Political Factors in Terrorism*, *in* PERSPECTIVE ON TERRORISM 24–25 (Lawrence Freedom & Yonah Alexander eds. 1983)).

238. *Id.* at 37 (citing Rushworth M. Kidder).

world," a view that allows them to "redefin[e] the morality of killing, so that it can be done free from self-censuring restraints."[239]

Psychiatric expertise, however, may be of limited use in this situation. As noted earlier, the ability of psychiatrists to opine accurately as to future dangerousness in death penalty cases is disputed, perhaps no better than one time in three chances,[240] though the predictive ability increases when there is a consistent history of violent behavior. Similar doubts have arisen as to the ability of psychiatric experts to opine accurately in civil commitment proceedings. Compounding the matter is the fact that "[p]sychological pathology does not seem to be present in higher rates among terrorism perpetrators than it is among members of the general public."[241] If true, the lack of pathology would render psychiatric expertise of even less use here than in civil commitment proceedings. We may not even need psychiatric experts to opine about a particular detainee's dangerousness. Trained military interrogators could interview each detainee with an aim toward eliciting the detainee's attitudes toward the West, toward violence against civilians, and other such matters outlined above.[242] Because the close-knit nature of terrorist cells tends to reinforce the self-defensive, oppressed worldview, detention in a facility such as Camp Delta may, over a period of time, mute such attitudes and reduce the detainee's dangerousness.

Thus, drawing upon the previous examples of pretrial detention, civil commitment, and others, one can build a model for detaining persons captured in military conflicts against nonstate actors. Assuming that such persons have been correctly

---

239. Albert Bandura, *The Role of Selective Moral Disengagement in Terrorism and Counterterrorism*, *in* UNDERSTANDING TERRORISM: PSYCHOSOCIAL ROOTS, CONSEQUENCES, AND INTERVENTIONS 121, 124–25 (Fathali M. Moghaddam & Anthony J. Marsella eds., 2004).

240. *See supra* note 230 and accompanying text.

241. *See* Audrey Kurth Cronin, *Sources of Contemporary Terrorism*, *in* ATTACKING TERRORISM: ELEMENTS OF A GRAND STRATEGY 19, 24 (2004); *see also* PHILIP JENKINS, IMAGES OF TERROR: WHAT WE CAN AND CAN'T KNOW ABOUT TERRORISM 68 (2003).

242. An interesting question arises if a detainee refuses to be interviewed, thus denying the United States the opportunity to gather evidence of continued dangerousness. Such refusal to be interviewed could be viewed as evidence of non-cooperation and thus potential dangerousness; it should not, however, be conclusive evidence of dangerousness.

classified as within the ambit of the military force authorization,[243] continued detention would depend on the government's ability to demonstrate the detainees' dangerousness, subject to the following procedural requirements.

1. Counsel

The first feature of the model permits detainees to be represented by counsel. The majority of the examples of noncriminal detention discussed earlier—pretrial detention, civil commitment of the mentally ill, and civil commitment of violent sexual predators—allow detainees to seek counsel, and, in the case of indigent detainees, provide counsel at the state's expense.[244] The reasoning underlying the provision of counsel in those instances varies depending on the nature of the detention. At minimum, it is based on the importance of the individual's stake in the matter: his liberty. The classic example of the right to counsel arises in criminal prosecution, where the Court has interpreted the Sixth Amendment to require that indigent defendants who cannot afford to hire attorneys be provided them at public expense.[245] As other situations demonstrate, however, the right of indigents to be provided counsel is not limited to criminal defendants.

To be sure, not all instances of potential loss of liberty require the allowance or provision of counsel. In *Wolff v. McDonnell*,[246] the Court held that prison inmates facing disciplinary proceedings were not entitled to legal counsel at their hearings.[247] *Wolff* noted that allowing counsel into such proceedings "would inevitably give the proceedings a more adversary cast and tend to reduce their utility as a means to further correctional goals."[248] If the purpose of noncriminal detention of al Qaeda and Taliban fighters is to rehabilitate them so that they no longer seek to attack the United States, *Wolff* might seem applicable. Nevertheless, *Wolff* is ultimately less persuasive as an

---

243. *See supra* Part IV.A.

244. *See supra* Part III.

245. Gideon v. Wainwright, 372 U.S. 335 (1963); *see also* Johnson v. Zerbst, 304 U.S. 458 (1938) (requiring that counsel be provided to indigent defendants in federal criminal trial).

246. 418 U.S. 539 (1974).

247. *Id.* at 570.

248. *Id.*

analogy than *Salerno* and the Bail Reform Act, in part because *Wolff* dealt with a disciplinary process to punish criminals for whom there was no dispute that imprisonment was proper. Here, the hearings inquire whether the detainee remains dangerous and subject to continued detention, not whether an incremental restriction on already limited liberty is justified.

Clearly, lawyers would be useful in this model. As trained advocates, they can help present the affirmative case of nondangerousness (should the detainee testify or present witnesses and evidence) as well as test the government's showing of dangerousness. As the Court explained in *Powell v. Alabama*, even an "intelligent and educated layman . . . lacks both the skill and knowledge adequately to prepare his defense, even though he have a perfect one," and without counsel, "he faces the danger of conviction because he does not know how to establish his innocence."[249] Here, of course, it is not innocence that is at issue, but dangerousness; nevertheless, the point remains that the nonlawyer may not be trained to *argue* for a particular outcome.

Elsewhere, this author has considered the argument that the Combatant Status Review Hearings, used to determine if each detainee was properly classified as a combatant, could prohibit detainees from obtaining counsel on the ground that the factual question to be determined was one where detainees could be as effectively assisted by nonlawyer military officers:

> [T]his is not to deny that lawyers would be able to help the detainees present more effective cases. However, it does suggest that the marginal gain in accuracy from having lawyers present is measured from a different baseline than in the criminal cases: there are not complex rules of litigation to contend with, the issue to be resolved is relatively simple and factual, and to the extent there are complexities in the definition of "combatant," the personal representative should have the training to assist the detainee.[250]

Although the government interest in excluding counsel in the Combatant Status Review Hearings was outweighed by the marginal benefits that counsel would offer, given that there were lawyers who were willing to represent individual detain-

249. 287 U.S. 45, 68–69 (1932).

250. Yin, *supra* note 22, at 1086.

ees,[251] counsel might yet be excludable upon a sufficient showing of national security needs. That possibility, however, rests in large part on the fact that the personal representative assigned to the detainee—a military officer—would necessarily be trained to make the determination of "combatant" status, because the laws of war require that soldiers be able to distinguish combatants, who are legitimate targets, from noncombatants, who are not legitimate targets.[252] Whether one's views on the necessity of counsel in the Combatant Status Review Hearings, the argument against counsel seems much weaker in this context, where military officers cannot be assumed to have the training to assist a detainee in arguing nondangerousness. To a military officer, all members of enemy forces are presumed dangerous, and hence are legitimate targets for killing or capturing. That presumption is entirely reasonable on the battlefield, but it begs the question that the dangerousness hearings are meant to resolve.

One important qualification remains. In order to provide counsel to represent the detainees in dangerousness hearings that would take place on Guantanamo Bay, the government might be able to draw only from a limited pool of lawyers who possess the appropriate security clearances, both to set foot on the base and to see classified information.[253] If the number of detainees were to increase dramatically, the government might be temporarily unable to provide enough lawyers to represent detainees. In such an event, the frequency of the hearings might have to be scaled back.[254]

2. Periodic Evaluations

Another feature of the model is periodic reevaluation of the dangerousness of each detainee. This reevaluation system is also a feature of the civil commitment of the mentally ill and of violent sexual predators, where the state must, on at least an annual basis, prove that the confined person remains danger-

---

251. *Id.* at 1087–88.

252. *Id.* at 1084–85.

253. *See id.* at 1087–88.

254. *See, e.g.*, Mathews v. Eldridge, 424 U.S. 319 (1976). Perhaps *Eldridge* could be read as requiring the government to provide more lawyers, rather than cut the number of hearings. At the same time, it seems like that the courts would defer to the executive branch's military judgment about the feasibility of allocating more military lawyers and officer to staff the additional hearings.

ous.[255] In upholding the constitutionality of such confinement, the Court relied on the fact that the detention was only "potentially" indefinite: "If Kansas seeks to continue the detention beyond that year, a court must once again determine beyond a reasonable doubt that the detainee satisfies the same standards as required for the initial confinement."[256] The detainees, on the other hand, face potentially endless detention. True, a few of the Guantanamo detainees are slated for prosecution in military tribunals,[257] but they may remain in detention even if they are acquitted. Thus, the military tribunals will not, in themselves, end detention so much as provide an opportunity for the government to inflict punishment in addition to preventative detention.

In a traditional nation-state conflict, enemy POWs may also face potentially endless detention; certainly, during World War II, German and Japanese POWs died in captivity before the end of the war. For enemy POWs in traditional nation-state wars, the event that terminates their captivity—the end of the conflict—is one that is out of their hands. As long as a state of military conflict exists between the U.S. and the nation state of the enemy POWs, those POWs are considered belligerents who would be ordered to fight us against if released. Nevertheless, in traditional conflicts, there is an entity (the enemy nation) that continues to exist and that can negotiate an end to the conflict. It may not be the same entity, but it is the same sovereign. Some military conflicts, such as the 1991 Gulf War, have limited objectives that do not involve changing the enemy national government; others, such as the recent conflicts in Iraq and Afghanistan, as well as World War II, have a goal of replacing a sovereign's government, but not the sovereign itself. That is, before World War II had ended, we ousted the Nazi government of Germany, but Germany as a sovereign remained intact, with a new government installed by the Allied Powers.

When military force is applied against nonstate actors such as al Qaeda, the continuing existence of such an entity is much less clear, as is the ability of such entity to negotiate an end to the conflict. Al Qaeda is not, and never was, a sovereign. The goal of the military conflict against al Qaeda is not to replace its

---

255. *See supra* Part III.B–C.

256. Kansas v. Hendricks, 521 U.S. 346, 364 (1997).

257. *See supra* note 156.

"government" with another, but rather to destroy al Qaeda. In a very real sense, the war has become individual to each detainee. Accordingly, the possibility exists that a given individual will decide that he no longer wishes to fight against the United States.[258] If the detainee reaches such a decision, he is by definition no longer dangerous, and further detention is not necessary. Requiring the government to demonstrate continued dangerousness on an annual basis would presumably result in the release of detainees who have decided that they no longer want to fight the United States. Necessarily, the frequency of hearings would also depend on other factors, including the number of detainees. Annual hearings that would be feasible for five hundred detainees might well be impossible for five thousand detainees, given the demand for ten times as many lawyers, decisionmakers, and expert evaluators (not to mention physical space for the hearings).

3. Choice of Forum

Finally, it is necessary to consider the type of forum in which these hearings should take place. Should it be a federal judge who determines whether the government has shown that a particular detainee remains dangerous? Should it be a so-called Article I judge? Or should it be a panel of neutral military officers?

Certainly, there are non-frivolous arguments that, as a matter of policy, such hearings should take place before Article III judges. The attributes of Article III status, life tenure and protection from salary diminution,[259] are intended to provide federal judges with independence from the political branches and from the public.[260] In this context, a federal judge charged with determining whether a particular Guantanamo detainee is dangerous can theoretically do so strictly based on the merits of the matter, without fear of the consequences of second-guessing by superiors. Decisionmakers who are military officers, on the other hand, may feel, rightly or wrongly, that they need to rule

---

258. The reasons for reaching such a decision could vary from fearing the military power of the United States to rethinking the *jihad* to wishing to return home to one's family.

259. U.S. CONST. art. III, § 1; N. Pipeline Constr. Co. v. Marathon Pipe Line Co., 458 U.S. 50, 59–60 (1982).

260. *See* THE FEDERALIST NO. 78 (Alexander Hamilton).

against the detainees in close cases or suffer negative career consequences.[261]

Thus, the advantages of Article III judges over other decisionmakers who lack life tenure are undoubtedly real. However, there are many adjudicative decisionmakers, ranging from state court judges to administrative law judges to military judges who lack life tenure but who make judgments on questions of federal law. Potentially indefinite detention raises serious liberty concerns. Death penalty cases, however, raise even more serious liberty concerns, and no constitutional impediment exists to having non-life-tenured judges presiding over capital cases, even though a state judge who must stand for retention or reelection may feel pressure (again, rightly or wrongly) to steer the jury toward conviction of a notorious defendant. The mere superiority of life-tenured decisionmakers, however, does not translate into a requirement of life-tenured judges. Due process merely requires that the decisionmaker be free from actual bias.[262]

Moreover, use of non-Article III decisionmakers may provide the government with a degree of flexibility and the detainees with a degree of process that might not be available with the limited resource of federal judges. Neutral or unbiased military officers would be the most convenient decisionmakers, given that the detainees are being held on a military base outside the United States; having an Article III judge conduct the dangerousness hearings would require that either the judge be transported to the military base or the detainee be transported to the United States. The former imposes not insignificant burdens on federal judges, particularly if the detainees are held far from the United States. The latter imposes security demands on the military and on federal courts.

C. A Note on the Conditions of Detention

Conditions of confinement were important factors in other confinement situations detailed above. In *Salerno*, for example,

---

261. *Cf.* Yin, *supra* note 22, at 1075–78 (discussing unlawful command influence and other due process issues relating to the staffing of the Combatant Status Review Hearing panels by military officers).

262. *Cf.* Ann Althouse, *Tapping the State Court Resource*, 44 VAND. L. REV. 953, 960–961 (1991) (arguing that the relevant question is whether "the state courts are good enough, even if they fall short of parity").

the Court, in upholding the constitutionality of pretrial detention, placed some weight on its belief that pretrial detainees would be confined in separate facilities from convicted criminals.[263] Similarly, in 1945, the American Psychiatric Association issued a committee report outlining minimum standards for mental health hospitals, including the conditions of confinement, which included housing "either in single rooms or in small dormitories."[264] Furthermore, the conditions of prisoner-of-war camps as mandated by the Geneva Convention are consistent with a nonpunitive purpose.[265] A tremendous chasm exists, however, between theory and practice. For example, *Salerno*'s observation that pretrial detainees should be housed separately from convicted criminals has been routinely ignored on the grounds of expedience and necessity. Some mental health hospitals are criticized for substandard or even inhumane living conditions. This gap between theory and practice in these other examples of noncriminal detention demonstrates the futility in setting forth rules about the conditions of confinement.

D. Comparison to the Bush Administration's Administrative Review Hearings

On September 14, 2004, Navy Secretary Gordon England issued a directive establishing "administrative review procedures for enemy combatants detained" at Guantanamo Bay.[266] Even before this date, the government had periodically re-

263. United States v. Salerno, 481 U.S. 739, 747–48 (1987).

264. *See* ALBERT DEUTSCH, THE MENTALLY ILL IN AMERICA 452–53 (1949); *see also* Youngberg v. Romeo, 457 U.S. 307, 315, 324 (1982) (holding that a civil commitment detainee must be provided with safe conditions and "reasonably nonrestrictive confinement conditions").

265. *See* Geneva Convention (III), *supra* note 70, art. 25 ("Prisoners of war shall be quartered under conditions as favourable as those for the forces of the Detaining Power who are billeted in the same area. . . . The premises provided for the use of prisoners of war individually or collectively, shall be entirely protected from dampness and adequately heated and lighted, in particular between dusk and lights out.").

266. Memorandum from Gordon England, Sec'y of the Navy, to Sec'y of State (Sept. 14, 2004), http://www.defenselink.mil/news/Sep2004/d20040914adminreview.pdf [hereinafter England Memorandum]; News Release No. 905-04, Office of the Asst. Sec'y of Def. (Pub. Aff.), Administrative Review Implementation Directive (Sept. 15, 2004), *available at* http://www.defenselink.mil/releases/2004/nr20040915-1253.html.

leased detainees from Camp Delta,[267] through apparently less formal procedures. The Administrative Review Hearings are held before a panel of three military officers and will take place annually for each detainee.[268] The panel will determine "whether [the detainee] represents a continuing threat to the U.S. or its allies" by "consider[ing] all relevant and reasonably available information."[269] The burden of proof is not stated explicitly, in part because the Navy Secretary has described the proceedings as "non-adversarial";[270] however, the implementing document could be read as presuming either the need for continued detention or the need for release.[271]

Although the detainee is entitled to testify at this hearing, the detainee has no right to call other witnesses.[272] If the detainee wants family members or officials from his national government to help establish his lack of dangerousness, he can do so only by presenting written statements from those persons.[273] The detainee is not allowed counsel but is provided an "Assisting Military Officer" to help prove his non-dangerousness.[274]

The establishment of formal hearings to review annually the justification for continued detention is undoubtedly a welcome step, and the focus on long-term dangerousness is appropriate. The actual procedures used in the hearings, however, fall well short of those suggested in this Article. The combination of lack of counsel and the unclear burden of proof is particularly troublesome. While either of those provisions alone might be reasonable, it would be the presence of the other provision that remedies the absence of the first. As William Buss has noted in the context of disciplinary measures in public schools, where the right to cross-examination is denied, "participation by a lawyer becomes even more imperative to reveal weaknesses in

---

267. *See supra* notes 64–67 and accompanying text.

268. England Memorandum, *supra* note 266, enclosure (3) para. 2.

269. *See id.* at 2 para. 1(c).

270. *Id.*

271. *Compare id.* enclosure (3) para. 1(a) ("This process will provide an annual review to determine the need to continue to detain enemy combatants . . . . [T]he ARBs shall conduct such proceedings as necessary to make a written assessment of . . . whether there are other factors bearing upon the need for continued detention.") *with id.* ("This process will . . . explain why the enemy combatant's release would otherwise be appropriate.").

272. *Id.* enclosure (3) para. 3(d).

273. *Id.*

274. *Id.* enclosure (3) para. 2(c).

hostile testimony and to identify weaknesses in the testimony that cross-examination might have exposed."[275] Here, a lay person who must prove his lack of dangerousness may be at a loss to do anything other than to proclaim his own peacefulness.

Placing the burden of proof on the detainee, as the implementing memorandum may do, presents other problems when combined with the prohibition on presenting live testimony (other than the detainee's). Written statements from family members or officials of the detainee's national government will most likely be able to opine only on matters relevant to the detainee's personal history, as opposed to the detainee's present and future intentions. Yet, it is the latter (combined with the detainee's status as an enemy combatant within the terms of the Congressional authorization for the use of military force) that is indicative of whether the detainee still poses a threat to the United States.

Thus, the Administrative Review Hearings should be reconstituted so that detainees are provided counsel to represent them at the hearings. Furthermore, the burden of proof should be clarified as being on the government to show continuing dangerousness.

V. CONCLUSION

The perhaps inaptly named war on terrorism has exposed a gap between criminal law and the law of armed conflict. During the middle of the Twentieth Century, when the 1949 Geneva Convention was adopted, only nation-states were capable of inflicting the level of damage that would plausibly be seen as a warlike act; individuals committed crimes and were dealt with through the civilian criminal justice system. Even before September 11, 2001, however, the crime-war distinction had begun to disintegrate. Had the 1993 World Trade Center bombing been successful from the standpoint of its intended outcome, it would be remembered as the first instance of megaterrorism in the United States.

Yet simply because military force might be reasonably directed against nonstate actors under domestic and international

275. *Cf.* William G. Buss, *Procedural Due Process for School Discipline: Probing the Constitutional Outline*, 119 U. PA. L. REV. 545, 610 (1971).

law,[276] it does not follow that the law of armed conflict can be transferred seamlessly into this new context. Detention of enemy soldiers for the duration of a traditional nation-state conflict is premised upon the assumption that such soldiers, if released prior to the end of the conflict, would be ordered back into the conflict by their home nation. With nonstate actors, however, the same assumption might not be valid. Therefore a new method must be developed to determine how long nation-states may detain persons captured by them when military force is applied against such nonstate actors.

By synthesizing common elements from domestic forms of noncriminal detention, the proposed model creates a process whereby nonstate actors captured during authorized military campaigns can be detained for as long as the United States can demonstrate their continued dangerousness. The burden of proving such dangerousness rests with the United States, and the detainee is to be provided counsel to assist him in defending himself at the hearing. This approach is consistent with the nature of the detention pursuant to military force, and it is more flexible than the "prosecute or release" mindset. It also provides a measure of process above the government's current Administrative Review Hearings that presumably will lead to more confidence in the hearing determinations.

This Article focused on domestic law in creating a model for noncriminal detention of nonstate actors captured during authorized uses of military force. In this day and age, however, international law cannot be wholly ignored. Members of the international community, notably Western European nations, may object that certain aspects of the noncriminal detention model violate international law, such as the Universal Declaration of Human Rights,[277] the International Covenant on Political and Civil Rights,[278] and the Body of Principles for the Protec-

---

276. *See supra* Part IV.A.

277. Universal Declaration of Human Rights, G.A. Res. 217A art. 9, U.N. Doc. A/810 (Dec. 10, 1948) ("No one shall be subject to arbitrary arrest, detention, or exile."); *id.* at art. 10 ("Everyone is entitled in full equality to a fair and public hearing by an independent and impartial tribunal, in the determination of his rights and obligations and of any criminal charge against him.").

278. International Covenant on Civil and Political Rights, G.A. Res. 2200A (XXI) art. 9(4), 21 U.N. GAOR, Supp. No. 16, U.N. Doc A/6316 (Dec. 16, 1966) ("Anyone who is deprived of his liberty by arrest or detention shall be entitled to take

tion of All Persons under Any Form of Detention or Imprisonment,[279] by, among other things, not providing for direct judicial determination of the detainee's dangerousness.

The United States has an opportunity here to engage in a dialogue with the international community about how international law should treat detention of nonstate actors against whom a nation-state has authorized the use of military force. If this Article's analysis is persuasive—that is, if conventional criminal law is insufficient to address the conduct of such nonstate actors but the traditional law of armed conflict does not translate perfectly either—then the model proposed herein may be a start toward reconciling military force and criminal prosecution into a coherent international-law framework for the problems presented by nonstate belligerents.

---

proceedings before a court, in order that court may decide without delay on the lawfulness of his detention and order his release if the detention is not lawful.").

279. Body of Principles for the Protection of All Persons under Any Form of Detention or Imprisonment, G.A. Res. 43/173 Principle 11(1), 43 U.N. GAOR, Supp. No. 49, U.N. Doc. A/43/49 (Dec. 9, 1988) ("A person shall not be kept in detention without being given an effective opportunity to be heard promptly by a judicial or other authority. A detained person shall have the right to defend himself or to be assisted by counsel as prescribed by law."); *id.* at Principle 11(3) ("A judicial or other authority shall be empowered to review as appropriate the continuance of detention.").

USING INFORMATION MARKETS TO IMPROVE PUBLIC DECISION MAKING

ROBERT W. HAHN* AND PAUL C. TETLOCK**

I. INTRODUCTION....214
II. AN INTRODUCTION TO INFORMATION MARKETS....219
III. GETTING BETTER INFORMATION FOR MAKING POLICY CHOICES....224
A. A New Approach....224
B. Extending the Framework....234
C. The Potential for Improving Fairness....238
D. Potential Problems with This Approach....241
1. Versatility of the Approach....241
2. Project Governance....243
3. Measurement Issues....246
4. Market Design Issues....251
IV. A BENEFIT-COST ANALYSIS OF INFORMATION MARKETS....261
A. A More Informed Assessment of Policy Proposals....262
B. Greater Transparency and Accountability in Decision Making....263
C. Greater Availability of Assets for Financing Projects and Spreading Risks....265
D. Cost of Information Markets....266

---

* Co-Founder and Executive Director, American Enterprise Institute-Brookings Joint Center for Regulatory Studies; Resident Scholar, American Enterprise Institute.

** Assistant Professor of Finance, University of Texas at Austin, McCombs School of Business.

The authors would like to thank Michael Abramowicz, Matthew Adler, Jason Burnett, Peter Cramton, Christopher DeMuth, Robin Hanson, Randall Kroszner, Lester Lave, Robert Litan, Paul Milgrom, Terry Murray, Sam Peltzman, Eric Posner, Richard Schmalensee, Cass Sunstein, Philip Tetlock, and Scott Wallsten for helpful comments and Laura Goodman and Rohit Malik for research assistance. This article represents the views of the authors and does not necessarily represent the views of the institutions with which they are affiliated.

E. Comparison of Direct and Indirect Approaches for Information Markets.....268
V. SIGNIFICANCE FOR POLICY DESIGN AND EVALUATION 270
A. Information Markets in the Policy Process.....270
B. Solving Some Difficult Government Oversight Problems.....273
C. The Potential Role for Government and Researchers.....277
VI. CONCLUSION 283
VII. APPENDIX.....287

I. INTRODUCTION

Many legal scholars have studied how to improve public decision making. Justice Breyer, for example, argues that technical problems could benefit from greater scientific expertise, and suggests using such analysis to help prioritize among competing social needs.[1] Cass Sunstein argues for the judicious use of cost-benefit analysis in a variety of areas, but also points out its limitations.[2] Sunstein's proposal, and the proposals of other scholars, would rely heavily, albeit not exclusively, on cost-benefit analysis to evaluate public policy decisions.[3]

Cost-benefit analysis is a tool used by decision makers to help inform the policy process. Cost-benefit analysis examines how different policies affect the overall level of net benefits to society, or benefits minus costs. A cost-benefit analysis may also be used to explore equity issues, examining how the distribution of net benefits varies across key groups, such as minorities or small businesses.[4]

---

1. *See, e.g.*, STEPHEN BREYER, BREAKING THE VICIOUS CIRCLE: TOWARD EFFECTIVE RISK REGULATION (1993); Stephen Breyer, Assoc. Justice of the U.S. S. Ct., Economic Reasoning and Judicial Review, AEI-Brookings Joint Center 2003 Distinguished Lecture (Dec. 4, 2003).

2. *See, e.g.*, CASS R. SUNSTEIN, RISK AND REASON: SAFETY, LAW, AND THE ENVIRONMENT (2002).

3. *See, e.g.*, *id.*; KENNETH J. ARROW ET AL., BENEFIT-COST ANALYSIS IN ENVIRONMENTAL, HEALTH, AND SAFETY REGULATION: A STATEMENT OF PRINCIPLES (1986), *available at* http://aei-brookings.org/publications/books/benefit_cost_analysis.pdf.

4. A project can be said to yield net benefits if the economic benefits less the economic costs are positive. These calculations frequently ignore distributional implications; that is, a policy will typically have winners and losers. *See generally*

A fundamental problem with cost-benefit analysis of new policies is that the analysis is conducted before such policies are implemented. When conducting ex ante analyses, it is difficult to predict the future values of key variables that could be affected by a policy.[5] For example, an analyst might predict that a worldwide carbon tax of $100 per ton would reduce world GDP by 1% in 2010.[6] How confident should we be in such a prediction?

In this paper, we present a new framework for addressing such uncertainty; this framework has the potential to substantially improve public decision making. We argue that decision makers can be more confident in analytical results if these results are based more directly on market data. Our framework introduces "information markets" that allow people to profit from superior knowledge about the future.[7] For example, if an information market suggested that expected GDP would fall by 1% with a carbon tax,[8] this estimate would theoretically incorporate all publicly available information about that policy's effects. We also argue that if these information markets are designed well, information from the prices in these markets is likely to be much more accurate than other forecasts.

An information market allows individuals to purchase contracts, using real money, that yield returns to their owners contingent upon the uncertain outcome of a future event.[9] With the

---

E.J. MISHAN, COST-BENEFIT ANALYSIS (2d ed. 1976); D.W. PEARCE, COST-BENEFIT ANALYSIS (2d ed. 1983); EDITH STOKEY & RICHARD ZECKHAUSER, A PRIMER FOR POLICY ANALYSIS (1978). For an analysis of the appropriate use of cost-benefit analysis in federal regulations, see ARROW ET AL., *supra* note 3.

5. We use the terms "project" and "policy" interchangeably. From the standpoint of a firm, it is undertaking a project. From the standpoint of the government, it is implementing a policy. Policies include regulations, research, standards, laws, programs, and public works projects.

6. For a similar prediction, see, William R. Cline, *Meeting the Challenge of Global Warming*, *in* GLOBAL CRISES, GLOBAL SOLUTIONS (Bjorn Lomborg ed., forthcoming), *available at* http://www.copenhagenconsensus.com/Default.asp?ID=165.

7. *See* Robin Hanson, Shall We Vote on Values, but Bet on Beliefs? (September 2003) (unpublished manuscript, *available at* http://hanson.gmu.edu/futarchy.pdf); *see also* Michael Abramowicz, *Information Markets, Administrative Decisionmaking, and Predictive Cost-Benefit Analysis*, 71 U. CHI. L. REV. 933 (2004).

8. When we say the market may "know," "believe," or "suggest," we are referring to the knowledge and beliefs of speculators in the market, which will be reflected in the market price. In this article, when we ascribe a view to the market, such as "the market expects," we use such phrases as shorthand.

9. For a useful definition of information markets, see Justin Wolfers & Eric Zitzewitz, *Prediction Markets*, J. ECON. PERSP., Spring 2004, at 107, 108 ("Anal-

advent of the Internet, information markets are becoming more common. They are used in a number of contexts, ranging from assessing the likelihood that the Federal Reserve will raise interest rates to assessing the odds that a particular presidential candidate will be elected.

As an example, consider the online exchange at TradeSports.com. This exchange allowed its members to trade contracts that yielded $10 to their owners if President Bush was reelected in November 2004. The contracts yielded $0 if Bush was not reelected. The prevailing price of these contracts on July 6, 2004 ($5.42)[10] revealed the price at which the supply for Bush contracts equaled the demand for these contracts. Assuming that the market was efficient, this price suggested the probability of Bush's reelection was 54.2% on that date.[11]

The idea of making greater use of information markets to promote social objectives is not new. Indeed, several scholars have suggested using information markets in a number of different contexts. Robin Hanson suggests that governments use information markets to identify whether particular policies will improve national welfare; he proposes relying on such markets exclusively, for example, when information markets predict a

---

ytically, these are markets where participants trade in contracts whose payoff depends on unknown future events."). The literature also refers to these markets as "speculative markets" and "betting markets." For a discussion of speculative markets, see Hanson, *supra* note 7, at 6 ("Most markets for stocks, bonds, currency, and commodities futures are called *speculative markets* because they allow people to bet on future prices by buying or selling today in the hope of later reversing such trades for a profit."). For a discussion of betting markets, see Paul W. Rhode & Koleman S. Strumpf, *Historical Presidential Betting Markets*, J. ECON. PERSP., Spring 2004, at 127; Hanson, *supra* note 7, at 6 ("*Betting markets* are speculative markets that trade assets that are specifically designed to allow people to bet on particular matters of fact, such as which horse will win a race.").

10. This price is as of 3:18 p.m. EDT, July 6, 2004.

11. Efficient prices equal "risk-neutral" probabilities, which are equal to actual probabilities if either the marginal trader is risk-neutral or the contract in question is not exposed to *systematic* (or aggregate) risk. It is possible, for example, that traders believe that future aggregate labor income will be lower if Bush is reelected. If they hold this belief, risk-averse traders will be willing to pay a premium for contracts that insure against the risk that aggregate income will fall in the future. The equilibrium price of the Bush contracts will therefore exceed the probability that he is reelected. However, through the clever use of other information markets, one can still infer the probability that Bush will be reelected, even in this more complicated example. The price of a contract yielding payments dependent upon future labor income conditional on Bush being reelected would identify the necessary adjustment between Bush's risk-neutral reelection probability and his true reelection probability.

policy will increase GDP.[12] Michael Abramowicz also notes the potential of such markets and advocates using them to predict the results of a cost-benefit analysis that would be done in the future by a designated expert.[13] Ronnie Horesh offers a novel proposal for creating incentives for private investors to aid the government in achieving social policy objectives: government-issued "Social Policy Bonds" that pay a fixed amount after a certain performance objective is met, such as achieving a particular unemployment rate.[14] His proposal has some similarities to ours, but does not address how to obtain information on the costs and benefits of a policy prior to implementation.[15]

To this point, no one has explicitly linked the improved information obtained through these markets to the subsequent implementation of policies that can maximize net benefits. Our central contribution is to propose an efficient way to implement well informed policy decisions.[16] Our hope is that this general approach will induce the government to make more efficient decisions—and, in particular, curb the appetites of

12. *See* Hanson, *supra* note 7, at 10 ("Therefore, if one is willing to recommend policies that statistical studies suggest will increase (a time average of future) GDP, one should be willing recommend [sic] policies that speculative markets estimate will increase GDP, and so one should be willing to consider a form of government which relies more on such market estimates in choosing policies.").

13. *See* Abramowicz, *supra* note 7, at 997 ("The information markets described so far can provide inputs into governmental decisions, but they cannot provide comprehensive assessments of the decisions themselves. Predictive cost-benefit analysis provides at least a solution to these dilemmas, by creating an information market to predict the outcome of a cost-benefit analysis."); *id.* at 940 ("This Article thus imagines predictive cost-benefit analysis, an information market used to predict a cost-benefit analysis that would be performed some time after a decision whether to enact a policy.").

14. *See* Ronnie Horesh, Social Policy Bonds: Explained in 4200 Words, http://www.geocities.com/socialpbonds/aeu.html (last visited Nov. 20, 2005).

15. Horesh's proposal differs from ours in two other ways. First, it does not pay for incremental progress toward a specific goal. That is, the bond would only pay the fixed amount if a specific objective were achieved. Second, in the absence of costless bargaining and coordination among firms holding social policy bonds, individual firms will not have adequate incentives to help meet the performance objective under his proposal because of problems with free riding. By contrast, our approach explicitly deals with the free-rider problem.

16. By the word "efficient," we mean relatively efficient but not necessarily perfectly efficient. Our proposal enhances the efficiency of existing policies, but may only be optimal in special circumstances.

politicians for introducing policies that have costs far exceeding their benefits.[17]

Our new approach builds upon the literature on information markets as well as the literature on the design of efficient policy mechanisms. Our claim is that the prices in information markets can inform the mechanism design process, thereby making previously infeasible mechanisms feasible for the policy maker. Specifically, information markets substantially broaden the scope for offering pay-for-performance contracts that yield useful information on the net benefits of a policy.

First, we show how it is generally possible to design markets whose prices will convey useful information on the costs and benefits of a number of policy choices, ranging from regulation to public works projects. Second, we describe one way of providing incentives for self-interested agents to implement policies that maximize net social benefits. Third, we show how information markets can be used to provide a stronger foundation for implementing a variety of government oversight mechanisms that, up to this point, have been stymied because of difficulties in estimating costs and benefits. We also show how legislators can use traditional budgetary controls in conjunction with information markets to exercise more effective oversight.

Finally, we identify and analyze the strengths and limitations of using information markets to help improve policy. While we focus on public-sector decision making, our analysis also holds for cost-benefit analyses used to inform private-sector and not-for-profit sector decision making.

This Article proceeds as follows: Part II provides a brief introduction to information markets. Part III provides a template for designing contracts that can supply better information on costs and benefits. Part IV discusses key costs and benefits of using information markets. Part V addresses the significance of

---

17. We are not so naïve as to think that a particular mechanism advanced by economists will be sufficient to stem the tide of inefficient government programs. On the other hand, the introduction of policy tools that increase transparency and provide better information on efficiency could help hold elected officials and bureaucrats more accountable. We do not, however, wish to imply that all regulation or taxation policies are inappropriate. For example, some environmental regulations can be expected to have benefits in excess of their costs if designed well. The key is to assess such regulations carefully with the best policy tools available. For a general discussion of the strengths and limits of economic analysis in designing more effective policy, see George J. Stigler, *Economists and Public Policy*, REGULATION, May–June 1982, at 13.

these contracts for policy design and evaluation, and suggests a new approach to legislative and regulatory oversight. Part VI presents our conclusions.

II. AN INTRODUCTION TO INFORMATION MARKETS

All markets can be thought of as providing some kind of information. We use the term "information market" to denote a market for contracts that yield payments based on the outcome of an uncertain future event.[18] In this paper, we specifically consider contracts that yield financial payoffs to their owners contingent on the status of key policy variables. The relevant policy decisions could be private or public, but this paper focuses on public decisions. Information markets could provide information related to costs, benefits, net benefits, or the likelihood that a certain event will occur, such as the probability that a President is reelected.[19] They can also be used to address potential market failures in the provision of information on public policy matters.[20]

Information markets differ from traditional equity markets in that they are not typically tied to a claim of an ownership stake in a firm. Instead, the assets are claims that will pay off an amount that depends upon the state of the world, such as the monetary value of actual policy benefits.[21] Although information markets for claims on benefits do not correspond directly to traditional equity markets for claims on corporate profits, there is a clear analogy. By monetizing policy benefits, information contracts allow organizations implementing policies to

---

18. Information market contracts are a subset of futures market contracts. *See* WILLIAM F. SHARPE, INVESTMENTS 521 (3d ed. 1985) ("Whenever something is ordered instead of purchased on the spot, a *forward* or *future* contract is involved. The price is decided at the time the order is placed, but cash is exchanged for merchandise later. . . . *Futures contracts* (*futures* for short), provide a standardized means of engaging in such transactions for agricultural and other commodities and for financial instruments and stock indices."); *see also* MCGRAW HILL ENCYCLOPEDIA OF ECONOMICS 449 (Douglas Greenwald ed., 2d ed. 1994) (providing a similar definition of a futures market contract).

19. *See generally* Robert Forsythe et al., *Anatomy of an Experimental Political Stock Market*, 82 AM. ECON. REV. 1142 (1992).

20. *See* discussion *infra* Part V.

21. *See* Abramowicz, *supra* note 7, at 934 ("The securities in such a market do not serve as claims to corporate ownership, but rather offer payoffs contingent on the occurrence of some future event specified by the market's sponsor.").

transfer risks from these policies, just as traditional equity contracts allow corporations to transfer risks from their projects.

Information markets have already been used in a variety of contexts. The most well known information markets are for small-stakes political contracts; researchers at the University of Iowa, for example, conduct an electronic market for political futures contracts. In the corporate world, Hewlett Packard has experimented with information markets to forecast sales, while Eli Lilly has used these markets to help predict the success of pharmaceuticals. TradeSports.com offers information contracts in a number of areas including sports, politics, finance, law, entertainment, and even the weather. Goldman Sachs supports an exchange called economicderivatives.com, which hosts call auctions for securities based on economic indices.

The use of information markets is becoming more widespread.[22] Hedgestreet.com recently received regulatory approval from the Commodity Futures Trading Commission to list several contracts based on economic indices; Macro Securities Research has developed another contract design that enables claims on economic indices to be bought and sold; and Case Shiller Weiss has constructed real estate indices that could provide the basis for information market contracts whose prices would convey the public's knowledge of housing prices.[23]

The prices in information markets not only provide an additional source of information for use in cost-benefit analyses, they also dominate existing sources of information in many applications. This price discovery process seems to work well in theory and in practice. Las Vegas odds and point spreads predict the outcomes of sporting events better than sports ex-

22. In 2003, there was a controversial proposal to use information markets to predict terrorist events. A discussion of this proposal is beyond the scope of this paper. The issue is complicated because both terrorists and the government can affect the outcome. In addition, the government may not wish to disclose the kind of sensitive intelligence that an information market would reveal. For a discussion of the issues surrounding this proposal, see David M. Pennock, The Good Side of the "Terror Futures" Idea (Yes, There is One) (2004), http://dpennock.com/pam.html; Scott Wallsten, Congress Shorts Future Terror-Fighting Innovation, AEI-Brookings Joint Center (July 2003), http://www.aei.brookings.org/policy/page.php?id=147.

23. For example, the owner of such a contract might receive a payment proportional to the real estate index if a certain policy were implemented.

perts.[24] The prices in Iowa political markets are more accurate than concurrent opinion polls in forecasting elections 451 times out of 596.[25] Hewlett-Packard information markets beat official forecasts in predicting printer sales in two separate trials: fifteen times out of sixteen in one trial and six out of eight in the other.[26] Even play-money markets are a dominant source of information, outperforming four out of five columnists at forecasting Oscar winners in 2000.[27]

Markets are generally the best available mechanism for gathering and aggregating dispersed information from private, self-interested economic agents.[28] People have something to lose if they are wrong, unlike in a standard cost-benefit analysis. That is, if the prices in information markets are poor predictors of the future, speculators have a direct economic incentive to get better information and trade on their superior information, moving prices toward the expected value of the contract payments.

The prices in properly designed information markets provide information on key cost-benefit analysis parameters in different states of the world. The economic theory and evidence cited above suggests the error of the market's implicit net benefits forecast will be no more than the error from a traditional cost-benefit projection.[29] Information markets not only provide in-

---

24. *See* PerformanZ Football Ratings, http://tbeck.freeshell.org/ (last visited Nov. 20, 2005) (comparing experts, polls and the updated Las Vegas betting line in college and professional football and basketball); *see also* Raymond D. Sauer, *The Economics of Wagering Markets*, 36 J. ECON. LIT. 2021 (1998).

25. *See* Joyce Berg et al., Results from a Dozen Years of Election Futures Markets Research (Mar. 2003) (unpublished manuscript, http://www.biz.uiowa.edu/faculty/trietz/papers/iemresults.pdf).

26. *See* Kay-Yut Chen & Charles R. Plott, *Information Aggregation Mechanisms: Concept, Design and Implementation for a Sales Forecasting Problem* (Cal. Inst. of Tech., Social Science Working Paper No. 1131, Mar. 2002), *available at* http://www.hpl.hp.com/personal/Kay-Yut_Chen/paper/ms020408.pdf; *see also* Ajit Kambil, *You Can Bet on Idea Markets*, HARV. BUS. SCH. WORKING KNOWLEDGE, Dec. 1, 2003, *at* http://hbswk.hbs.edu/pubitem.jhtml?id=3808&t= innovation.

27. *See* David M. Pennock et al., *Extracting Collective Probabilistic Forecasts from Web Games*, PROC. 7th ACM SIGKDD INT'L CONF. KNOWLEDGE DISCOVERY & DATA MINING 2001.

28. *See* Chen & Plott, *supra* note 26; Hanson, *supra* note 7; *see also* Berg et al., *supra* note 25. *See generally* THE HANDBOOK OF EXPERIMENTAL ECONOMICS (John H. Kagel & Alvin E. Roth eds., 1995); F.A. Hayek, *The Use of Knowledge in Society*, 35 AM. ECON. REV. 519 (1945).

29. This particular error measure is the mean squared error. The mean squared error from a traditional cost-benefit projection is the average squared difference

formation about the true mean of important quantities, but also have the potential to shed light on their entire distribution. So, for example, one could assess the probability that a policy would have a small, medium, or large impact on GDP, where the size of a small impact is defined as less than a 1% increase, and so forth. Abstracting from hedging considerations, the price of a contract that yields $100 in the event that GDP growth falls below 1% will indicate the probability that the policy will cause GDP growth to fall below 1%.[30]

Evidence from the psychology literature buttresses the case for using information markets instead of experts. In many instances, decision makers cannot outperform even simple statistical rules.[31] In fact, many experts fail to outperform statistical rules based only on the experts themselves; the experts merely add noise to their own forecasting "rules."[32] In contrast, financial markets outperform simple statistical rules; failure to do so would undermine the reason for their existence.[33] Other markets, such as sports wagering markets, also appear quite efficient at aggregating information. As previously mentioned, Las Vegas odds forecast sports game outcomes better than any expert, outperform any ranking or poll system, and perform as well as any betting rule.[34]

This Article considers only markets for contracts that use real currency as a medium of exchange. Information markets for contracts that use alternatives, such as points, may also prove

---

between actual net benefits and the government's net benefit estimate. For the market forecast, the mean squared error is the average squared difference between actual net benefits and the implied market expectation of net benefits estimated from the current price.

30. With an estimate of the underlying distribution of net benefits, regulators would be better able to avoid regulatory projects that could be major failures, if that were viewed as desirable. This opens up the question of whether the government should be risk neutral in its choice of policy instruments, which is beyond the scope of this paper. *See* Kenneth J. Arrow & Robert C. Lind, *Uncertainty and the Evaluation of Public Investment Decisions*, 60 AMER. ECON. REV. 364 (1970); *see also* DISCOUNTING AND INTERGENERATIONAL EQUITY (Paul R. Portney & John P. Weyant eds., 1999).

31. *See* Robyn M. Dawes et al., *Clinical Versus Actuarial Judgment*, 243 SCIENCE 1668 (1989) (highlighting the superiority of statistical judgments over clinical judgments).

32. See id.

33. *See* BURTON G. MALKIEL, A RANDOM WALK DOWN WALL STREET (8th ed. 2003); Eugene F. Fama, *Efficient Capital Markets: A Review of Theory and Empirical Work*, 25 J. FIN. 383 (1970); *see also* Michael C. Jensen & George A. Bennington, *Random Walks and Technical Theories: Some Additional Evidence*, 25 J. FIN. 467 (1970).

34. *See* PerformanZ Football Ratings, *supra* note 24; *see also* Sauer, *supra* note 24.

helpful. The scope for the manipulation of prices is probably heightened, however, in markets in which contract sales and purchases amount to "cheap talk." Participants in markets that have alternative mediums of exchange may not have a strong incentive to acquire costly information about fundamental values. Research by Robin Hanson and Ryan Oprea shows that, in contrast, attempts at price manipulation can actually enhance the accuracy of real-money information markets by giving incentives to other market participants to acquire costly information.[35]

A number of scholars have addressed the issue of how information markets could contribute to policy design. Hanson, for example, suggests that information markets can be used to assess whether certain government policies should be implemented. Specifically, he suggests using the information market's estimate of the likely impact that a proposed policy has on GDP as the core determinant of whether it should be implemented.[36]

Abramowicz amplifies Hanson's insight in the context of cost-benefit analysis. He suggests a novel approach to estimating the net benefits of a policy. In essence, Abramowicz suggests using an information market to forecast how a cost-benefit analyst, specified by the government in the future, will, for example, estimate the net benefits of a regulation to reduce arsenic after she has a chance to observe these net benefits.[37]

Instead of estimating key parameters that could be used in a cost-benefit analysis now, Abramowicz suggests that his approach may actually be better. Abramowicz would go even further in terms of the flexibility given to the analyst by allowing the analyst to define costs and benefits in a way that seems appropriate to her at that time. Allowing the analyst complete

---

35. *See* Robin Hanson & Ryan Oprea, Manipulators Increase Information Market Accuracy (July 2004), http://hanson.gmu.edu/biashelp.pdf (suggesting that manipulation indirectly increases the accuracy of prices in information markets by increasing incentives for informed trading).

36. *See* Hanson, *supra* note 7, at 14 ("The basic rule of government would be: When an approved betting market clearly estimates that a proposed policy would increase expected GDP+, that proposal becomes law.") Hanson defines GDP+ as a measure of national welfare that incorporates national income and values such as "lifespan, leisure, environmental assets, cultural prowess, and happiness." *See id.* Hanson recognizes that GDP may be an imperfect measure, and suggests using other measures as they are developed. *See id.* at 9–10.

37. *See* Abramowicz, *supra* note 7, at 939–40.

flexibility has the advantage that the analyst can respond to changes in our understanding of cost-benefit analysis.

Although the approach Abramowicz suggests has merit, there is a great deal to be gained from using information markets to estimate parameters that would be *directly* useful for doing a cost-benefit analysis.[38] We will show how this can be done for a wide range of problems. In addition, we suggest an efficient mechanism for implementing such projects.

III. GETTING BETTER INFORMATION FOR MAKING POLICY CHOICES

This section presents a new approach to assessing and implementing policies that builds on two ideas. The first is to use information markets to help inform policy makers, the public, and interested parties about the likely benefits of policies. The second is to implement policies using a design that pays for actual results, so that the party implementing the policy has the appropriate incentive to produce the desired results. By connecting these two ideas, the decision maker can obtain valuable information on the costs, benefits, and net benefits of a particular policy prior to actually implementing that policy.

A. A New Approach

We begin by showing that it is possible to define a set of market mechanisms that yields useful information on net benefits, and that this information can be used to implement desired policies in an efficient manner. The essence of our approach is illustrated in Figure 1. The basic idea is for the government to gather information to decide if a policy proposal is worthwhile, then if it is worthwhile, to implement that proposal in a way that is efficient—that is, that maximizes net social benefits.[39]

The approach consists of four steps. First, the government must specify and monetize all verifiable benefits accruing from

---

38. *See* discussion *infra* Parts IV, V.

39. We do not claim that our framework is optimal. For example, the performance-based policy design advocated in this Article has advantages and disadvantages relative to "book building" methods for selling project equity contracts. Book building, the informal process used by most U.S. investment banks to assess interest, allocate shares, and determine prices in initial public offerings, may not be feasible for the government because of the potential for unequal treatment of investors. A comprehensive comparison of our framework with all possible mechanisms is beyond the scope of this paper.

a possible project.[40] The specification of verifiable measures is a prerequisite for designing appropriate information market contracts.[41] It is also a key to designing policies that reward those implementing a project based on actual performance.

Second, a contract that provides information on expected benefits with no change in policy would need to be issued. The expected benefits with no change in policy would provide the benchmark against which performance is measured if the policy is actually implemented. This benchmark is critical for getting a reasonable assessment of net benefits.

As part of this second step, the government would have a trading exchange list the information market contract for trading. The exchange would serve as a meeting place for buyers and sellers making contract orders on the not-yet-issued contracts.[42] Generally, potential trades would occur when and if a particular policy is implemented. Unlike normal securities, however, government information market benefit contracts would be traded even if there is no change in policy because a project is not implemented.[43]

Third, the government would conduct an auction for the right to implement and receive the monetized incremental benefits from that proposal.[44] The government would have the option of setting a minimum acceptable price for the auction.[45] This price would reflect unquantifiable benefits and costs as assessed by the government and would be set in advance.

---

40. *See* discussion *infra* Part III for a more formal model. A project could be thought of as any action that yields quantifiable and verifiable benefits. Examples include enforcing a policy or regulation, researching a new technology, or manufacturing a product.

41. We relax the requirement that the government must know the monetary value of benefits later in Part III. We also consider the case in which not all benefits or costs can be quantified.

42. The exchange could be subsidized or run by the government.

43. Trades in securities generally take place only when and if the security is issued and are referred to as "when-issued" contracts. Examples include treasury notes in the U.S. and initial public offerings in other countries, such as Germany.

44. The baseline for the incremental benefits would be determined by the estimate of benefits without the policy, which would be obtained from the information market. Without some estimate of the benefits that would result with no change in policy, it is not possible to measure the net benefits that are associated with a policy. It is, of course, possible to sell off a pay-for-performance contract without such a benchmark, but the resulting sale would not provide useful information on the incremental net benefits of the project.

45. In the economics literature, this is often referred to as a "reservation" price.

Fourth, the performance contract would be sold to the highest bidder if the auction revenues exceed the minimum acceptable price. That bidder would win the right to implement the policy and receive the monetized incremental benefits.[46] That right would be freely transferable, in whole or in part.

In a competitive bidding market with risk-neutral bidders, the revenue raised by this auction would approximate the project's expected net social benefits.[47] This is because different agents are competing for the right to get the social benefits from the project, as defined by the government. The winning bidder's profits would equal the project's social benefits less its private costs.[48]

If the project does not go ahead because the bids are less than the minimum acceptable price, then trades in the information market benefit contracts would occur at the prices already agreed upon by market participants. The final value of the contracts would depend on actual measured benefits at some prespecified point in time. If the project does go ahead, then these

---

46. Many types of firms could be winners in the auction. We hypothesize that certain firms with good access to information, funding, and an ability to resell securities would be prime candidates; examples include investment banks, venture capital firms, hedge funds, and institutional investors. Involvement of such firms would increase the likelihood that this process will work because these firms could serve as underwriters for the monetized benefits. Underwriters help solve the information asymmetry problem and coordinate buyers' and sellers' interests in the security in a manner analogous to their function in private securities markets.

47. The highest bid will not necessarily equal the highest bidder's valuation, but could be a close approximation, depending on how others value the contract. *See* Peter Cramton, *Ascending Auctions*, 42 EUR. ECON. REV. 745 (1998). This auction is efficient (that is, it allocates the good to the agent with the highest value) and optimal (that is, raises the most revenues) when there is perfect resale. It allows for the presence of private and public information on the value of a good. Furthermore, the auction is flexible enough to allow the sale of multiple benefits payments with different reserve prices. *See* Lawrence M. Ausubel & Peter Cramton, *Auctioning Many Divisible Goods*, 2 J. EUR. ECON. ASS'N 480 (2004). Depending on the structure of the social problem, it might be preferable for the government to auction benefits payments to multiple firms.

48. Perfect competition is not required. For example, if two risk-neutral firms value the project identically and there is Bertrand price competition, this result will hold. Without risk-neutrality, the winning bid will be less than expected net benefits by an amount reflecting the firm's cost of capital (that is, the project's systematic risk). Generally, the difference between the auction revenue and net benefits will be equal to the economic rent obtained by the winning bidder (that is, the difference between the project's net benefits for the high bidder and the second highest bidder). *See* William Vickrey, *Counterspeculation, Auctions, and Competitive Sealed Tenders*, 16 J. FIN. 8 (1961).

trades would never materialize because they were conditioned on the assumption that there would be no change in policy.

A stylized example shows how the framework could work in practice. Suppose policy makers are interested in improving average standardized test scores of public high school students in a major metropolitan area. In step one, the government decides that it is willing to pay a firm $1 million for each point that average test scores improve in a year as a result of the firm's educational reform project.

In step two, to estimate the benefits, the government (or an exchange) can list a contract for trading that is issued only if there is no change in policy. If the contract is issued, it will pay $\$x$ when the average test score is x. Suppose the price of this contract is $75 just before the government auctions off the project. This implies that the average test score is expected to be 75 with no change in policy.[49]

In step three, the government auctions off the right to receive compensation for the education reform project. It offers to pay $1 million for each one-point increase in test scores above 75 points, the market's estimate of points scored in the absence of the reform project. Assume the winning bid reflects the expected net benefits from the policy, and that this bid is $7 million.

Finally, in step four, the government decides whether to proceed with the project. The government will fund the project if the winning bid and therefore the expected net benefits exceed some minimum amount. Suppose the government's decision rule is that net benefits must be greater than zero. In this case, the government will proceed with the project because the net benefits are estimated to be $7 million.[50] The firm winning this auction would work with the schools to increase performance measures, with the firm paid on the basis of the performance increase it delivers.

The framework in Figure 1 allows the decision maker to implement only those projects with high expected net benefits. For these projects, the auction mechanism selects the firm that

---

49. The possibility that the market price is a biased measure of expected test scores with no change in policy is discussed below.

50. In the case of multiple proposals, the government only needs to choose whether or not to implement the proposal yielding the highest auction bid (less the reserve price). All other proposals will yield lower expected net benefits.

is likely to deliver the highest net benefits. In some cases, the decision maker may also want to know information on the expected benefits prior to making a decision, as well as the likely direct costs to the decision maker of undertaking the project.

By introducing one additional information market in benefits, the decision maker can get information on both the expected benefits and the expected costs of the policy. This idea is illustrated in Figure 2. There are two key differences between Figure 2 and Figure 1. First, in step 2 of Figure 2, we add an information market on the benefits that would result if the project were implemented. This allows an estimate of the expected benefits of the project based on the difference between the prices in the two information markets. If the government does decide to go ahead with the auction, it obtains an estimate of expected net benefits using the auction price. To obtain an estimate of expected costs, it can subtract the expected net benefits from the expected benefits.[51]

The second change in the figure is that the government has a choice whether to auction the right to receive benefits payments in step 3. We add this choice here because the government may not want to go ahead with the auction if the expected benefits from the project are small.[52]

To illustrate how Figure 2 can be applied by extending the stylized example on test scores, suppose the government or an exchange issues a second type of contract that pays $\$y$ if the government chooses the pay-for-performance model and the average test score is y. Suppose the price of that contract is $85 just before the government decides whether to implement the policy. Recall that the price of the contract associated with no change in policy was $75. These prices would suggest that the market predicts a ten-point increase in average test scores (85 minus 75). Policy makers value this expected increase at $1 million per point or $10 million in total. The prices give the government advance knowledge of the likely benefits of a particular policy.

---

51. When costs are verifiable the government can measure costs directly rather than via benefits minus net benefits by creating contracts with payoffs dependent on the future costs with and without the proposed policy. This is not feasible, however, when costs are not verifiable.

52. This choice was omitted from Figure 1 to keep the framework simple. But, as discussed in Part III.B, there are many ways of extending the basic framework.

As before, assume the auction occurs and the winning bid is $7 million. The government then has information on benefits and net benefits. By subtracting benefits from net benefits, it gets an estimate of the costs of the project. In this case, costs are estimated to be $3 million.[53]

Before dealing with some limitations of the framework, it is important to note a number of points about its advantages and general applicability. First, note that the government is only paying for the cost of the project. Paying for costs is better than paying for benefits because this requires less revenue to be raised by the government. Most politically acceptable means of raising revenues distort after-tax prices in the economy,[54] leading to inefficiencies in production or consumption, or both.[55] In the preceding example, the government only has to raise $3 million to cover the expected costs of the project, as opposed to $10 million to cover the expected benefits.

Second, the government does not need to monitor the costs of the project because the firm has every incentive to minimize costs. Although the government pays the firm on the basis of its actual performance—in this example, the improvement of average test scores—the firm also bears all costs incurred in the process of improving scores. The firm thus has an incentive to efficiently manage its productive inputs for improving test scores, which means minimizing its costs.[56]

---

53. $10 million in benefits minus $7 million in net benefits equals $3 million in costs.

54. Only lump sum taxes (sometimes called poll or head taxes) do not distort decision making.

55. *See* Martin Feldstein, *Tax Avoidance and the Deadweight Loss of the Income Tax*, 81 REV. ECON. & STAT. 674 (1999). Feldstein estimates that as much as $2.06 per dollar of taxes is lost through deadweight efficiency losses. *See id.* at 678.

56. By design, this approach characterizes the problem of a social planner trying to maximize expected social surplus. *See* Mark A. Cohen & Paul H. Rubin, *Private Enforcement of Public Policy*, 3 YALE J. ON REG. 167 (1985) (applying this idea to private enforcement); Marc J. Roberts & Michael Spence, *Effluent Charges and Licenses Under Uncertainty*, 5 J. PUB. ECON. 193 (1976) (using a similar idea in the context of maximizing the net benefits of pollution control). That is, the firm maximizes social benefits less private costs, which are assumed to equal social costs. If there were externalities associated with the firms' investments, these could be dealt with in standard ways to internalize those externalities. *See, e.g.*, WILLIAM J. BAUMOL & WALLACE E. OATES, THE THEORY OF ENVIRONMENTAL POLICY (2d ed. 1988). Externalities could also be dealt with by including them in the information markets as part of a social benefit function or social cost function. For example, suppose introducing traffic lights in a town reduced fatalities but added delays. Both of these factors could be considered in an expanded benefit

Third, the transferability of the project (or a component thereof) permits additional efficiency gains. This transferability allows the project manager to sell shares in the project conferring the right to some fraction of the benefits payments from a policy. Any party could purchase these shares, whether it was interested in managing the project differently, managing the project more efficiently, or perhaps terminating the project. Transferability allows for the transfer of project control between managers in response to changing market conditions.

These project shares are very similar to the corresponding information market contracts. Both project shares and information market contracts entitle their owners to the same cash flows—that is, a fraction of the project benefits.[57] The only material difference is that a project share confers project control rights upon its owner as well.[58] This difference is akin to the difference between voting shares and non-voting shares in public equity markets. Under certain conditions, the price between the project shares and information market contracts may differ.[59] This price difference poses no obvious theoretical or practical problems for the proper functioning of information markets.

The market for the transfer of projects would function analogously to the market for corporate control. Through efficient bargaining, the most capable firm will end up doing the project. The price of the security in the shares market and the par-

---

function. This equivalence implies that the profit-maximizing firm will solve the problem optimally from a social standpoint. In the interest of simplicity, we abstract from explicitly including the deadweight losses to the government associated with taxation. We will discuss this issue in more detail in a related paper. By purchasing the monetized benefits from the government, the firm is able to internalize the social benefits and reap the surplus. Of course, in a competitive bidding market, the firm must pay this surplus to the government in order to win the right to implement the project.

57. The precise relationship between shares and contracts will depend on the number of shares and number of contracts issued.

58. There is another, immaterial difference between the two claims. In the event of project liquidation, the owner of a claim on project benefits receives a fraction of the actual proceeds from the liquidation, whereas the owner of the information market contract receives the equivalent amount in cash from the counterparty to his transaction. Counterparty margins secure the payment in one case, whereas the value of the asset secures the payment in the other case.

59. *See* Luigi Zingales, *What Determines the Value of Corporate Votes?* 110 Q.J. ECON. 1047, 1052–55 (1995) (requiring that (1) there must be a contest over corporate control; (2) the participants in this contest must receive private benefits of control; (3) there must be some information asymmetry).

allel market for contracts would reflect the market's belief that the most efficient firm will undertake the project.

The market for transfer of projects differs from the market for corporate control in one crucial respect: The government could repurchase the rights to the project benefits. Government repurchase allows abandonment of a project or regulation, or assumption of control over project implementation. Thus, the government can revise its policy over time in light of new information about the social benefit function. Foreseeable changes in the social benefit function need not lead to government repurchase because these changes can be incorporated in the original government benefits payments at the outset. To return to the test score example, if the government expects the value of a one-point rise in test scores to increase from $1 million to $2 million, it can promise $1 million per point initially and $2 million per point after the date when it anticipates a change in the value of a one-point rise in scores. The market price for project shares will reflect the possibility of project takeover or termination, ensuring that all voluntary buyers and sellers in such a transaction will be adequately compensated for the transfer of control rights.

Government project takeover will not be necessary unless the social benefit function changes unexpectedly.[60] Even if new information about the likely success of the policy is revealed after the benefits payments have already been sold, the project manager will still optimally implement the project. For example, if it becomes clear that the project will not improve average test scores, the project manager will cease spending money without any prompting by the government.

In general, a project could give rise to other changes in government behavior that would not have occurred without the existence of the project. Thus, the market price of the education performance contract is a gross measure of benefits that includes the benefits from the policy changes that are expected to arise from the initial project. The measure of net benefits (auction proceeds) will also include the present value of all expected net benefits from anticipated policy changes.

---

60. The government may also repurchase the project when the government is the most efficient owner.

To see this point, reconsider the amount of money a project manager would be willing to bid in the step-three auction for the right to the benefits payments. Although the successful firm will not directly incur the costs of any other, unimplemented projects that could improve test scores, it will have to pay for these expected benefits in the auction for the right to the benefits payments. Upon winning the auction, the project manager becomes the claim holder of the benefits payments, which depend on the success or failure of this project and future projects. After the project manager commits to spending the funds on education reform, she can resell the rights to the benefits in the form of project shares to whoever can efficiently implement future education reform projects. In many instances, this repurchasing party is likely to be the government. Thus, the government will buy back the rights to the project shares when it appears that the manager has less use for these rights than the government. If no repurchasing party is likely to be interested in the rights, then the initial auction price and the subsequent project share and information contract prices will reflect only the benefits from the manager's project. In either case, the measure of project net benefits will be accurate.

The key point is that the benefits and net benefits measured in both the project share prices and the information market contract prices include all expected policy changes that result from the current decision. Information market prices signal the full marginal value of a social project by accounting for both its direct and indirect effects on welfare.[61]

Extending this argument, the anticipated benefits from one policy depend upon market participants' expectations about the government's future behavior *contingent on the policy being implemented*. If traders believe the government will choose more socially valuable projects following a particular initial policy, market participants will expect greater future benefits from that initial policy. Accordingly, the price of a claim on the policy's benefits, either a project share or an information market contract, will reflect these expectations.

Another advantage of transferability of shares is that the winning firm can hedge its investment in the project by selling

61. The prices can measure only the effects of policies that are not deterministically linked. That is, if one policy implies with certainty that another policy will be implemented, the government cannot construct an information market to measure the impact of each policy separately.

a fraction of the project payoffs. In the preceding education reform example, suppose that the project is implemented and awarded to the winning firm. The information contract yields one dollar for every one-point increase in test scores; so too, would rights to a 1-in-1,000,000 fractional share of the project benefits.

Suppose an educational testing company, say Princeton Review, is the winner of the right to collect $1 million for each point that test scores increase above 75 points. Assume also that Princeton Review defines a share in the project as a 1 in 1,000,000 stake in the project's monetized benefits. For simplicity, suppose that the price of one share is equal to $10, which is exactly the difference in price between the corresponding information contracts—that is, there is no premium for the control rights. Princeton Review could sell z shares to raise $\$10z$ (current price times z) and use these proceeds to finance the education reform project. If Princeton Review sells at least 300,000 shares, then it can fully finance the project (300,000 times $10 = $3 million).[62]

Note that Princeton Review will earn only ($1 million minus $\$z$) per one-point increase in test scores, implying that increases in the number of shares it sells reduces its incentive to invest in increasing scores. There is a trade-off between allowing the firm to sell more project shares for hedging purposes[63] and wanting the firm to hold shares to maintain its incentive to invest the socially optimal quantity in education funds. For example, if Princeton Review sells 300,000 shares, it would then have expected profits of $7 million (that is, the 10 point increase times $1 million per point, less $3 million). The net effect of agreeing to fund the project and selling these shares would be to pay nothing today in exchange for the right to receive

62. In general, the project manager can undertake myriad financial transactions to divide the rights to project control and cash flows. An important lesson from corporate finance theory is that these transactions will not affect efficiency unless they create or destroy value (for example, by affecting the implementation of the project). *See* Franco Modigliani & Merton H. Miller, *The Cost of Capital, Corporation Finance and the Theory of Investment*, 53 AM. ECON. REV. 261 (1958) (proving the sufficient conditions for this result to hold).

63. Compared to its profits under an unhedged position, Princeton Review's profits would be higher than they would have been if the actual point increase is less than ten, and lower than they would have been if the point increases exceed ten. By selling short, Princeton Review reduces the variation in expected profits associated with an unhedged position.

$700,000 per one-point increase. In this case, Princeton Review would still have a strong incentive to make the project succeed. If Princeton Review performed research that led to further improvement of test scores, it would receive even more money from the government.[64]

B. Extending the Framework

We consider three extensions of the framework illustrated in Figure 2: allowing for benefits to accrue over time; using the minimum acceptable price for different purposes; and adding information markets to supply more detailed information and improve hedging opportunities.

First, consider the problem of benefits accruing over time. If benefits accrue over time, the government can introduce a suitable discount factor into the performance measure. Suppose test scores improved over time. The contract could offer annual payoffs based on the average improvement in test scores each year. The price of this contract would now reflect the present discounted value of improvements in test scores, sending the appropriate signals to firms trying to value the project.[65] In principle, the timing of benefits is not critical, so long as the benefits are measurable.

Second, the government can use the minimum acceptable price to incorporate both information on unquantifiable benefits and costs and a budget constraint. If all costs and benefits are included in the framework, then an optimal decision rule would be to undertake the project if net benefits exceed zero. On the other hand, if there are some unquantifiable benefits and costs, then the government could add an amount to the minimum acceptable price that represents the government's best estimate of these unquantifiable costs and benefits.[66] The government could then impose the requirement that it will

64. *See* discussion *infra* Part III.C (examining the more general case of the incentive versus hedging trade-off).

65. If, for some reason, the social discount rate differs from the project discount rate, then the government should adjust future payments by the difference between the discount rates. This change induces the private sector to use the social discount rate.

66. If so-called unquantifiable benefits or costs were thought to be correlated with quantifiable benefits or costs, then the government could take account of this relationship directly in the benefit function rather than by adding an amount to the reserve price. In addition, if unquantifiable benefits or costs were expected to vary with the scale of the project and the expected benefits the project, then it would be more efficient to take this into account directly in the benefit function.

only sell the benefits payments if the winning bid in the auction lies above some pre-specified cutoff value. While the accuracy of these government estimates of unquantifiable costs and benefits is not guaranteed, there is some reason to think that a budget-constrained agency would behave reasonably in assessing these amounts.[67] Specifically, the agency would have some incentive, as under a regulatory budget, to maximize net benefits as it defined them.

The government can also use the minimum acceptable price to limit its expected payouts to the firm winning the contract.[68] Limiting payouts could be important if the government faces budget constraints. Suppose, as before, that the government estimates benefits of $10 million from the increase in test scores, but it is not willing to pay out more than $4 million. The government can then set the minimum acceptable price in the auction at $6 million. Of course, even if the minimum acceptable price were $6 million, this does not assure that the government would pay no more than $4 million. To eliminate the risk entirely, the government could insure against the possibility that its payouts exceeded $4 million.[69] If the government faces a budget constraint on several projects, the same logic applies.[70] It may, however, need to hold a series of auctions simultaneously to satisfy the overall budget constraint.[71]

---

67. A complete treatment of the unquantifiability issue is beyond the scope of this paper.

68. If the government wants to satisfy both a budget constraint and account for unquantifiable costs, then it should set its reserve price equal to the maximum of the reserve prices implied by the two cutoffs.

69. This transaction could entail the government swapping its obligation to pay the project benefits in exchange for the revenues from the net benefits auction and a bond with a present value of $4 million. Corporations and financial institutions, such as investment banks and insurance companies, routinely conduct similar transactions in "over-the-counter" markets. There is little reason to expect the government would have difficulty executing such a transaction.

70. The same logic on budget constraints also applies to firms. A possible application for firms would be capital budgeting in a world in which they face capital constraints.

71. A good example is the case where the government considers two projects, both of which have expected benefits of $10 million. Suppose the government cannot spend more than $8 million total on the two projects. Then the government could impose a rule that both benefits payments are sold only if the combined revenues from the two auctions are greater than $12 million (and each auction raises positive revenues). If the combined revenues fell short of $12 million, then the government could implement whichever project raised greater revenues,

At first glance, it may seem inconsistent to allow the decision maker discretion in choosing a minimum cutoff value for auction proceeds while at the same time forcing her to use a quantitative rule to implement the project. These seemingly inconsistent directives merely require the decision maker to impose a binding minimum price prior to the auction of benefits. The good is not sold unless some bidder meets the policy maker's pre-specified price.[72] This price must be determined in advance of the information market for net benefits and the auction to ensure there is no ex post manipulation of policy decisions. Although the minimum price could be private or public during the information market for net benefits and the auction, announcing that price has the advantage of transparency.[73]

Third, we consider the role that additional information markets play in providing information for the decision maker and parties participating in the markets. Our framework can be modified to allow policy makers to assess the sensitivity of estimated net benefits to changes in assumptions about the value of policy benefits, which would be particularly important in cases where the benefit function is uncertain, or key parameters in the benefit function are uncertain.[74] The sensitivity analysis could also be useful when the government's budget is limited.

Consider, for example, the case where a government facing budget constraints is unsure whether to value a statistical life at $1, $5, or $10 million.[75] In step one, the government could specify all three possible benefits values. In step two, the govern-

---

assuming these revenues exceeded $2 million. If neither project raised $2 million in revenues, then no project would be implemented.

72. There could be other restrictions on auction implementation, such as invalidation of the auction in the event of corruption or collusion in bidding.

73. On the other hand, a private reserve price will induce uncertainty whether bidders in the auction will meet the reserve price. This uncertainty may enhance the liquidity of the information market for benefits contracts by increasing the likelihood that trades will take place. Thus, uncertainty may be desirable in this instance.

74. More generally, information markets can help with sensitivity analysis on key parameters, such as costs, benefits, and net benefits to the extent that useful proxies for these measures are available. *See generally* M. GRANGER MORGAN & MAX HENRION, UNCERTAINTY: A GUIDE TO DEALING WITH UNCERTAINTY IN QUANTITATIVE RISK AND POLICY ANALYSIS (1990) (discussing different approaches to dealing with uncertainty).

75. *See generally* W. KIP VISCUSI, FATAL TRADEOFFS: PUBLIC AND PRIVATE RESPONSIBILITIES FOR RISK 17–98 (1992) (discussing the value of a statistical life); W. Kip Viscusi & Joseph E. Aldy, *The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World*, 27 J. RISK & UNCERTAINTY 5 (2003) (reviewing different estimates of the value of a statistical life).

ment would establish liquid information markets for contracts based on the benefits from each possible benefit value. In step three, the government would conduct multiple simultaneous auctions—one for each benefit value—in order to estimate the net benefits under the three alternative assumptions about the value of a statistical life.[76] The government could choose its benefit payment based on how the amount of the winning bid changes as a function of the value assigned to a life. Only the benefits payments corresponding to the value chosen by the government would be implemented in step four, and only the trades in the market corresponding to these benefits payments would take place.

More generally, this modified framework allows the government to assess how changes in its declared monetized benefit function would affect the estimated social surplus from a given policy. In contrast to sensitivity analyses performed in traditional cost-benefit analyses, this sensitivity analysis adjusts net benefits to account for general adjustments by firms to changes in the assigned monetary value of benefits.

The ability of information markets to accommodate a wide range of assumptions regarding the monetary value of benefits is particularly important for certain applications where this value is uncertain and likely to vary across the voting public. For example, the implied value of a statistical life depends strongly on the context in which the life is saved. Important factors include whether the program saving the life is voluntary or involuntary, whether the risk is "dreaded" or not, and whether the risk is familiar or not.[77] Also, different people value statistical lives differently, suggesting that the optimal policy depends upon the person selecting that policy. Sensitivity analysis using information markets allows policy makers to observe the dependence of optimal policies on different preferences.

The government may want to consider introducing more information markets on both benefits and net benefits if that information were useful. For example, if a proposed regulation

---

76. In fact, the government could explore an entire range of possible values of a statistical life by asking firms to submit demand functions for the project.

77. *See generally* Paul Slovic et al., *Regulation of Risk: A Psychological Perspective*, *in* REGULATORY POLICY AND THE SOCIAL SCIENCES 241 (Roger G. Noll ed., 1985) (discussing more thoroughly how and why the implied value of a life can vary).

were expected to reduce mortality and morbidity, the government could issue contracts that yield payments to their owners based on the level of these two quantities with and without the policy. Similarly, the government could use multiple information markets if it wished to compare more than two policy alternatives, such as a comparison of the status quo with a regulation that provides varying degrees of flexibility in implementation.

Finally, the government could also introduce an information market contract to estimate net benefits, as measured by auction revenues, prior to the auction. Using the education example, an information contract could pay $1 for each $1 million of revenues that would be received, contingent on proceeding with the auction. There are a number of benefits to having such a market, including providing early information to decision makers on the expected benefits and costs of a policy, allowing hedging opportunities for firms bidding in the auction, and reducing the likelihood of underbidding in the auction due to the winner's curse.[78]

C. The Potential for Improving Fairness

It is often difficult to estimate the distributional impact of a proposed policy and, indeed, many cost-benefit analyses fail to do so.[79] Fortunately, information markets can be used to help assess the impact of proposed policies on equity. Furthermore, under certain circumstances, information market contracts can assist policy makers in implementing equitable policies without harming efficiency. In economic terms, it may be possible for the government to approximate *any* efficient policy using transfer payments determined by information market prices. Although there are practical difficulties with obtaining reliable information on the equity implications of individual projects, it may be possible to obtain better information on general sets of policies and use this information in deciding on appropriate redistributive policies. We use two tax policy examples to illus-

78. The winner's curse occurs in auctions where the bidders have private information about the value of the good. The winning bidder is likely to be the one with the most optimistic private information. Adjusting for this fact, all bidders shade their bids downward to prevent over-bidding. *See generally* RICHARD H. THALER, THE WINNER'S CURSE: PARADOXES AND ANOMALIES OF ECONOMIC LIFE (1992).

79. *See, e.g.*, ARROW ET AL., *supra* note 3 (explaining the desirability of addressing such issues when data are available).

trate the potential of information markets to estimate the distributional impact of policy.[80]

The first example is a case in which the government wants to know the impact of a certain tax proposal on a specific income group, say those at or below the tenth percentile in income. The government could design two contracts that yield $1 for every $1,000 in that group's average after-tax income: One would be conditional on no change in policy and the other would be conditional on a change in policy. From the difference in the prices of these contracts, one could infer the estimated impact of the proposal on those with the lowest incomes. Perhaps more important, if the poor have access to such markets, they could hedge the risk that their income would fall by assuming a short position on this contract.[81] In addition, the contract price could inform policy makers seeking to offset the distributional impact of a project.[82]

A second example reveals how government can meet redistributive objectives without sacrificing efficiency by using information markets to uncover valuable information about the effect of a policy in advance of its implementation.[83] Consider a proposed income tax cut in a simple hypothetical society with just two types of workers. The current tax rate is 50% for the more productive type of worker, who earns $1 million per year before taxes ($500,000 after taxes); the tax rate is 0% for the less productive type of worker, who earns $10,000 per year (before and after taxes). Policy makers would like to reduce all mar-

---

80. We choose tax policy because distributional issues are salient in this area. Assuming availability of data, these ideas are applicable to other areas of policy, such as the impact of a policy on educational test scores for different income groups.

81. We recognize that the poor often have limited access to asset markets, limited means to invest in these markets, and limited awareness of opportunities in these markets. Notwithstanding these considerations, this particular strategy requires very little capital or specialized knowledge. Policy makers could alert the poor to this opportunity to purchase inexpensive income insurance, or could even purchase it on behalf of the poor. *Cf.* Cass R. Sunstein & Richard H. Thaler, *Libertarian Paternalism Is Not an Oxymoron*, 70 U. CHI. L. REV. 1159 (2003).

82. *See* ROBERT J. SHILLER, THE NEW FINANCIAL ORDER: RISK IN THE 21ST CENTURY 237–41 (2003) (emphasizing the value of hedging markets for addressing distributional concerns, but not suggesting hedging markets conditional on policy implementation); Hanson, *supra* note 7 (proposing conditional information markets, but not for hedging purposes).

83. Our example relies on lump sum transfers that may not be politically feasible.

ginal tax rates to zero in order to reduce distortionary impacts on work effort, but they are concerned that this will increase inequality.

Suppose policy makers would also like to retain the current absolute difference in after-tax income between the two types of workers ($490,000). To redistribute income efficiently, policy makers want to utilize lump sum taxes. But policy makers do not know the amount to tax the productive type after the tax cut, because this requires advance knowledge of the productive type's behavioral response to the tax cut, that is, his increase in work effort. To solve this problem, policy makers could create an information market contract that yields a payment equal to the average pre-tax income of the productive type after the tax.[84]

Suppose the price of the contract is $2 million, implying the productive workers will double their pre-tax incomes in response to the tax cut. Then the government could levy a per-capita lump-sum tax of $750,000 on the productive workers and give the proceeds to the unproductive workers (assumed to be equal in number to the productive workers and assumed not to alter their work effort choice). This transfer would leave the productive workers with $1.25 million in after-tax income and the unproductive workers with $760,000 in after-tax income. The two types can share equally in the efficiency gains only because the government is able to uncover valuable information about the effect of its policy in advance of the policy's implementation.[85]

These examples reveal that information markets can provide new insights into the equity impact of proposals. There is a broader question, however, of how equity concerns should enter into decision making. We do not have an answer to that question, but offer some suggestions for accommodating equity concerns in different ways.

The government could, for example, add some kind of equity constraint to its decision-making algorithm. If a policy has rela-

---

84. Assuming there are many productive workers, the contract's payment is only barely influenced by each individual worker's effort choice. Thus, the contract poses negligible moral hazard problems.

85. Although the example appears to rely heavily on the absence of uncertainty, this procedure can be adjusted to accommodate random shocks to each income group. *See* SHILLER, *supra* note 82, at 237–41 (noting that the contracts described above are ideal hedging vehicles for the productive workers, enabling exact redistributive transfers even in the presence of uncertainty).

tively modest expected net benefits, then the government could add a requirement that it not be harmful to groups at the lower end of the income range. This constraint need not apply to policies with higher levels of net benefits.[86] Government could also incorporate the equity constraint by altering the benefits payment formula to account for the distribution of the benefits.[87] Still another approach is for the government to measure the equity impact of particular policies, but not to address them on a case-by-case basis. Instead, government could address questions of income redistribution more broadly through fiscal policy.[88]

D. Potential Problems with This Approach

This section examines four groups of potential problems: versatility of the approach, project governance, measurement issues, and market design issues.

1. Versatility of the Approach

We begin by examining the versatility of the approach from the government's perspective. The approach is *designed* to be performance based, and in that sense, would represent a paradigm shift in many areas, including regulation. This performance-based paradigm has been endorsed by a number of leading regulatory scholars, including Justice Breyer.[89] Our approach differs from that of Justice Breyer in that it relies less on the expertise of government regulators and more on actual markets.

Although the benefit function explicitly pays out on the basis of performance, it is still possible for the government to pursue traditional command-and-control regulation if it so desires.

---

86. The rationale for exempting policies with high net benefits is that they yield a "bigger pie," some of which could be used for redistribution if desired. The definition of "high" is somewhat arbitrary. From an economic standpoint, the most logical measure would be return on investment.

87. The downside to this approach is that the government could not separately identify the expected impacts of the policy on efficiency and equity, which is crucial for certain applications.

88. *See, e.g.*, ARROW ET AL., *supra* note 3, at 8 (suggesting that regulation is a blunt instrument for redistribution); *see also* SHILLER, *supra* note 82, at 149–64 (suggesting ways of reducing inequality).

89. *See* BREYER, *supra* note 1, at 61–65 (describing and advocating a partially performance-based system).

This goal, if desired, could be met by modifying the framework slightly. Suppose, for example, the government wished to introduce an environmental standard with limited flexibility.[90] The general procedure would remain as described in Figure 1. The government could, however, auction the benefits to the project with a restriction on its implementation—for example, by specifically requiring a technology-based approach.[91]

Furthermore, the winning bidder would be allowed to contract with the government to actually implement the regulation if this contract were the most cost-effective way to do so. In this scenario, a private firm could negotiate with the Environmental Protection Agency to pay for the EPA's administrative, monitoring, and enforcement costs in exchange for a promise from the EPA to implement the regulation as outlined by the firm. The winning bidder can be thought of as a project manager, consultant, and overseer hired by the government. The performance-based compensation scheme gives the managing firm the appropriate incentives to implement the project efficiently. Thus, there are ways to use the framework to implement traditional and nontraditional regulatory approaches.[92]

Another issue that may be a problem from the government's perspective, or an agency's perspective, is the potential liability incurred. Many government agencies face strict budgetary requirements and may be reluctant to take on projects where they are promising a very large or unbounded amount based on undetermined future benefits. In the discussion of the minimum acceptable price, we suggested one approach for dealing with this problem. There are others that are worth considering. For example, the government could adjust the amount it is willing to pay for results—say for a point increase in average test scores. Alternatively, the government could consider put-

90. A technology-based standard for power plants might be an example. For a discussion of different kinds of standards, see Cary Coglianese & Gary E. Marchant, *Shifting Sands: The Limits of Science in Setting Risk Standards* (Regulatory Policy Program, Working Paper RPP-15, 2003), *available at* http://www.aei-brookings.com/publications/abstract.php?pid=392.

91. Examples of technology-based approaches to regulation include requiring the installation of scrubbers on power plants or catalytic converters on vehicles.

92. For example, one could apply the same idea to market-based approaches to reduce pollution such as the allowance trading used to limit pollution from sulfur dioxide emissions. In this case, a firm such as a trading exchange with expertise in trading systems could be hired to manage and oversee the process. The framework has the advantage that it provides measures of benefits, costs, and net benefits.

ting a cap on payouts. For example, a contract could promise $1 million per point for test score increases over 75 points with a maximum payout of $20 million (achieved when average test scores reach 95 points). Of course, this yearly benefit cap will blunt the incentive for the project manager to increase test scores if scores approach the 95-point cutoff. Finally, some of the problems the government wishes to address have potentially long time horizons or long latency periods. Examples include climate change and reducing exposure to carcinogens. Long latencies present no problem in principle for the information market mechanism described here. There are liquid financial markets for 30-year U.S. Treasury bonds and for 30-year residential mortgage-backed securities. Traditional equities have a potentially infinite maturity.

2. Project Governance

We consider two problems related to the incentives of the project manager: one if the manager does not act as a profit maximizer and a second related to excessive hedging and the incentives of the manager.

A problem could arise if the project manager decides not to behave as a profit maximizer. This could lead to inefficiencies, such as lower average test scores or wasteful expenditures. The potential for such behavior is not new to information markets or corporations. For example, firm management may be risk-averse, have empire-building goals, receive private benefits from control, or be interested in diverting cash flows into their own account.[93] There are already a number of mechanisms, including corporate charters, external financial contracts, internal labor contracts, and the market for corporate control, which address these issues.[94]

---

93. See Philipe Aghion & Patrick Bolton, An Incomplete Contracts Approach to Financial Contracting, 59 REV. ECON. STUD. 257 (1992); Michael C. Jensen, *Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers*, 76 AM. ECON. REV. 323 (1986), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=99580.

94. For a discussion of these and related issues, see Douglas Gale & Martin Hellwig, *Incentive-Compatible Debt Contracts: The One-Period Problem*, 52 REV. ECON. STUD. 647 (1985); S.J. Grossman & O.D. Hart, *Disclosure Laws and Takeover Bids*, 35 J. FIN. 323 (1980); Sanford J. Grossman & Oliver D. Hart, *One Share-One Vote and the Market for Corporate Control*, 20 J. FIN. ECON. 175 (1988); Bengt Holmstrom, *Moral Hazard in Teams*, 13 BELL J. ECON. 324 (1982); Bengt Holmstrom, *Moral Hazard and Observability*, 10 BELL J. ECON. 74 (1979).

Conversely, if the winning bidder tries "too hard" to maximize profits, this could create other problems. For example, because the winning bidder is the only firm with the incentive to solve a particular social problem, it could threaten to fail to perform on the contract with the government after it has won the auction unless the government increases the benefits payments. Even if the government deems this threat to be credible, several mechanisms could curb the firm's incentive to renegotiate the terms of the contract.[95] First, the government could put some requirements in the contract specifying minimum levels of performance, such as requiring a certain level of expenditures over time. Second, the government could refuse to do business with such a firm in the future. Third, consumers could punish the firm by refusing to buy its products once it became clear that the firm was harming the public by not acting in good faith; alternatively, consumer concerns could drive politicians to act. Finally, a regulatory body, such as the Federal Trade Commission, could be enlisted to identify cases in which firms significantly underperform and impose appropriate penalties.[96] All of these actions could negatively impact current and future revenues of the firm, and also affect its share price if it is publicly traded.

Another problem could arise during the implementation of the project if the project manager creates a hedge that reduces the firm's incentive to obtain social benefits. Suppose the project manager in our earlier example would like to completely hedge the project's risk by selling 1 million project shares at $10 each. If the price of shares remained constant as the manager sold its shares, the manager would then receive an up-front cash flow of $10 million in contract sales minus the $7 million in auction payments. If this hedge is possible, it would create a huge problem, because the project manager would have *no* incentive to spend the money on the education reform project. For each point increase in test scores, the manager would receive $1 million from the government, but also have to pay out $1 million to fulfill its obligations to project shareholders ($1 for each share sold). In this scenario, the manager would put forth

95. Because this is a bilateral bargaining situation, the government could try to hold up the firm as well. *See generally* Benjamin Klein et al., *Vertical Integration, Appropriable Rents, and the Competitive Contracting Process*, 21 J.L. & ECON. 297 (1978).

96. The agency would, of course, need to be acting within the law, and this should only be done as a last resort.

no effort, implying that the holders of project shares would expect to receive no payoff.

Fortunately, there is a natural market mechanism that would prevent this outcome. So long as the project manager must register its issuance of securities with the Securities and Exchange Commission (SEC), she would never be able to sell 1 million contracts at $10 each. The reason is that investors would observe that the project manager is selling its entire stake, thus blunting her incentive to undertake the project. A rational investor would be willing to pay nothing for the benefit contracts if the manager were selling its entire stake, because the investor would correctly anticipate that the manager would not bother with the costly project. This implies that the manager would make more money by holding on to the contracts, doing the project, improving the test scores, and receiving the benefits payments from the government.

A far-sighted manager will keep a large stake in the project to inspire confidence in her investors.[97] The project manager does not care about the investors per se, but she does care about the price received for stakes in her project. She will sell the project or hold claims on project benefits, depending upon which option will create greater profits. Recall that profits have been designed to equal social net benefits. So we have really shown that the firm that is the sole claimant to the project will take the necessary actions in securities markets to maximize social net benefits. The manager's project finance decision is analogous to the decision faced by an entrepreneur choosing what fraction of her project to sell to venture capitalists. This analogy suggests the mechanism will work in practice just as well as it does in theory.

There is an additional mechanism that could prevent the project manager from shirking her implementation of the project. The shirking problem exists only to the extent that the manager holds a net short position in the net benefit contract market. This short position implies that other contract holders are effectively financing the project and have strong incentives to monitor its completion. These contract holders could elect a board of directors and use standard corporate governance techniques to

97. In fact, the price of the project stake will already reflect the market's expectation of benefits given the fraction of the project that the auction winner intends to sell.

keep project management from enriching themselves at the expense of the project shareholders. For example, bonuses for management could be based on project performance and management could be replaced if it performed poorly. The governance mechanisms that operate effectively at the corporate finance level could operate equally effectively at the project finance level.

3. Measurement Issues

There are several problems with measurement that the government may want to address. First, it needs to decide on appropriate indicators of performance for particular projects. Second, it may want to assess the amount of noise in key variables and consider the gains from reducing uncertainty in measurement error.

If policy makers have access to multiple unbiased measures of benefits, they will need some guidance on which to use. We think a reasonable approach is to specify performance measures as broadly as possible, subject to the constraint that projects can be implemented on an efficient scale without wasteful coordination costs.[98] Typically, we would advocate specifying benefits on a project-by-project basis, depending on the objectives of the project. Examples might include having performance measures for improving educational outcomes or for reducing fatalities from auto crashes.

Having one firm buy a right to a broad package of benefits, such as GDP, is unlikely to be efficient for several reasons. First, a single firm is not likely to be well informed about the entire set of possible social projects needed to increase GDP. Firms specialize in different areas. Returning to our education example, Princeton Review has expertise and knowledge in improving children's test scores. It presumably lacks expertise in other areas, such as improving the environment.

Second, if a firm wins a performance contract based on increases in GDP, it will have to incur potentially huge costs to coordinate thousands of firms. If there were several performance contracts based on GDP, the coordination problems would be even worse. Narrowing the benefit function reduces these coordination problems. So, for example, Princeton Re-

98. The problem is analogous to determining the optimal size for a firm. *See* R.H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386 (1937).

view may not need to coordinate with any other firms to efficiently increase educational test scores.

Third, benefit functions linked to what a firm or small group of firms can produce allow firms to hedge their risks more easily. Princeton Review may not have much control over GDP, but it may have a lot of control over test scores. If firms are able to hedge their risks, then projects will become more attractive and more firms may participate in the auction. A firm may decide to avoid buying a residual claim on GDP, because it would fear bankruptcy if a general economic downturn occurred. For example, if Princeton Review owned a residual claim on GDP and GDP unexpectedly fell by just 1%, Princeton Review would lose roughly $100 billion, resulting in certain bankruptcy.[99]

To limit the bankruptcy risk, Princeton Review could hedge against risks in GDP beyond its control, but only if these risks are traded. This is really an argument in favor of introducing narrow project-based contracts, which would allow firms to hedge the risks from other projects that are beyond their control. Moreover, if the decision maker funding the performance contract divided it into narrow projects, then firms implementing the projects would not face any of the practical problems associated with hedging.[100]

In addition, there is an important theoretical reason to worry about the efficiency of reselling parts of project benefits to others. The prospective buyers of these hedging contracts could infer that the seller is not motivated by hedging, but by unfavorable private information.[101] This problem is particularly acute when only broad GDP contracts are available because the prices of these contracts do not convey information about the

---

99. Princeton Review could limit its risk by owning only a small fraction (say 0.1%) of the residual claim on GDP, but this would diminish drastically its incentive to reform education, rendering the "performance-based" contract ineffective.

100. These include transactions costs, liquidity risk, and basis risk. *See generally* JOHN C. HULL, OPTIONS, FUTURES AND OTHER DERIVATIVES (5th ed. 2003).

101. *See* George A. Akerlof, *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, 84 Q.J. ECON. 488 (1970) (classic treatment of how private information can lead to the sale of only "lemons" in the used car markets); *see also* Stewart C. Myers & Nicholas S. Majluf, *Corporate Financing and Investment Decisions when Firms Have Information that Investors Do Not Have*, 13 J. FIN. ECON. 187 (1984) (discussing models of adverse selection and its detrimental impact on efficiency).

individual components of GDP being resold. This adverse selection problem could be mitigated by creating information markets for the narrow, project-based contracts. A third party, such as the decision maker, could sell these contracts.

Another problem with issuing only broad contracts is that the information markets may not be very efficient. Distinct pieces of information about U.S. GDP are held by millions of consumers and thousands of firms in the economy. If there are even small fixed costs to participating in the GDP information market, then it is unlikely that everyone with information about GDP will participate. Even though together they may hold substantial information about GDP, most individuals hold an insignificant fraction of the total information. The information available to individuals is likely to address only their specific economic circumstances or expertise; in fact, it could be project-based. Individuals would search for more efficient ways to exploit specific project-based private information by trading in securities other than GDP.[102]

Fourth, there is little cost to the decision maker in specifying a set of narrower benefit functions. Suppose the decision maker is interested in GDP and that she knows that increases in GDP are best accomplished through increasing educational achievement and making better environmental investments.[103] Then, performance contracts could be auctioned off for education and the environment, reducing the systemic risks to which bidders would be exposed if they were rewarded in terms of GDP.

In certain cases, there may be reasons to broaden the scope of the performance contract. First, an excessively narrow performance measure could mean that the project manager ignores important externalities. For example, the education reform project described above could also reduce law enforcement costs in areas where uneducated high school dropouts resort to crime rather than finding a job. In this case, policy makers would need to value the impact of improved test scores

---

102. Traders exploiting the connection between the GDP and project-based markets would ensure that the GDP contract incorporates all project-specific information. Of course, the project-specific securities must exist for this mechanism to work.

103. If the decision maker does not know which categories of GDP are most important, she can design derivative securities to estimate the impact of establishing information market contracts in different areas on GDP. For more details on how this would work in practice, see discussion *infra* on the value of climate change research.

on criminal activity and reward these improvements. This would give the firm implementing education reform an adequate incentive to focus its efforts on areas with the greatest benefits from improved test scores. Second, there may be somewhat greater transaction costs associated with implementing more performance contracts. But these costs should not be significant in comparison to the benefits of writing narrower performance contracts.

In short, there are many factors that should be considered in determining the optimal scope of a performance contract. Our preference would be for the performance measure to include all benefits that can be reasonably quantified and no more. In addition, it should include externalities to the extent reasonable. The same firm should reap the social benefits and bear the full costs of its actions.

Another potential problem with the information markets framework relates to measuring the key quantities. Fortunately, for many policy interventions, there appear to be reasonably accurate proxy measures for benefits to form the basis for information market contracts. Without accurate, verifiable measures of such quantities, it is difficult to construct information markets that provide useful information.

Consider the previous example on education reform. The estimated increase in test scores is 10 points if the reform project is implemented. Because that result may be subject to a great deal of measurement error, it would be useful to know the error bounds. For example, if the variance in realized benefits is quite high, then the policy could just as easily result in a 5-point or 15-point increase in test scores. In this case, the government may not wish to go ahead with the reform project.[104]

The government can establish a market for another contract to estimate the uncertainty in benefits. Consider an option contract that entitles its owner to a payment of $1 for each point increase above 10 points if the policy is implemented. The price of this option will reflect the likelihood that actual benefits will

104. Some have argued that the government should behave as though it is risk-neutral in its policy decisions. Rather than making this normative judgment, we note that real policy makers often behave as though they prefer less risky alternatives. Without endorsing this behavior, we outline a process that a risk-averse policy maker could adopt to accommodate his preferences. For a good introduction to these and related issues, see DISCOUNTING AND INTERGENERATIONAL EQUITY (Paul R. Portney & John P. Weyant eds., 1999).

deviate from expected benefits, allowing the policy maker to infer the degree of uncertainty in realized benefits.[105]

Stock option prices in traditional financial markets already provide analogous measures of implied volatility for the prices of the underlying stocks. The Black-Scholes options pricing formula describes the mathematical relationship between the expected volatility of the underlying asset price and the price of an option on the asset, enabling policymakers to calculate the expected uncertainty in net benefits.[106] Such measures of uncertainty are particularly useful in situations where actual benefits have a large variance, such as climate change policy.

If information markets reveal that measurement error is a problem, they can also point to a solution by helping to assess the value of additional information.[107] If, for example, the relationship between greenhouse gas emissions and rising sea level is too uncertain to estimate accurately using information markets, the government might consider launching a research initiative to improve our understanding of this relationship. To measure the effectiveness of this proposed research, the government could establish markets for claims on the expected sea level with regulation of greenhouse gas emissions and the expected sea level without regulation. Consider the price difference between these two markets: This quantity measures the expected value of monetized sea level reductions from regulating emissions. The government normally would check whether this price difference exceeds cost to determine its optimal policy. In this case, however, the government would be concerned that this price difference and its optimal policy might change after-market participants learn more about the effects of greenhouse gas emissions on climate change.

---

105. Under the assumption that realized benefits are normally distributed, the price of this single option (along with the price of the benefits claims) will allow the government to infer the exact distribution of realized benefits. Even without this parametric assumption, the government can establish markets for multiple options with different strike prices (for example, point increases above and beyond 4 points) to estimate the distribution of realized benefits.

106. Fischer Black & Myron Scholes, *The Pricing of Options and Corporate Liabilities*, 81 J. POL. ECON. 637, 640–45 (1973).

107. Indeed, it is possible that information markets may reveal that more information is needed on many regulatory issues before the government can make an informed judgment on the net benefits of different policies. For a classic treatment of the value-of-information issue, see HOWARD RAIFFA, DECISION ANALYSIS: INTRODUCTORY LECTURES ON CHOICES UNDER UNCERTAINTY 157–61 (1968).

To assess the effect of the research initiative on this price difference and its optimal policy, the government would need to issue two more contracts. Both contracts would offer their owners the option at the projected completion time of the research initiative to receive a payment equal to the change in the price difference.[108] One contract would be issued if the research initiative is implemented, while the other would be issued if the research initiative is not implemented. The prices of these option contracts reflect market participants' beliefs about the new information that would be discovered with and without the research initiative. So, by comparing the prices of the options, the government can evaluate whether doing the research is likely to reduce uncertainty about the benefits of the policy, and by how much. Then it can proceed with steps three and four to choose whether to implement the research.

4. Market Design Issues

There are many important issues in market design.[109] Here we consider three: ensuring the market is liquid, ensuring market prices are not biased measures of policy benefits, and limiting market power in project implementation. This discussion is by no means exhaustive.[110]

Liquidity refers to the ability to execute large volume transactions in a market with minimal impact on the prevailing price. It is important for a number of reasons. First, liquidity is needed to get a reasonable price signal on the "true" underlying value of a quantity, because individual transaction prices in illiquid information markets may not be representative of market participants' beliefs. Second, liquidity allows well informed market participants to conduct profitable trades using their own information, providing incentives for traders to acquire

---

108. Formally, these securities are when-issued at-the-money European call options on the price difference, which expire at the projected completion time of the research initiative.

109. *See, e.g.*, Hanson, *supra* note 7, at 15–25 (discussing thirty engineering design issues for information markets).

110. *See, e.g.*, John O. Ledyard, *Designing Information Markets for Policy Analysis, in* INFORMATION MARKETS: A NEW WAY OF MAKING DECISIONS IN THE PRIVATE AND PUBLIC SECTOR (Robert W. Hahn & Paul C. Tetlock eds., forthcoming 2005) (focusing on several problems, including the potentially large number of markets that may be needed to inform policy decisions and the problems associated with getting an unbiased estimate of a policy's impact).

better information. Third, liquidity is important for information markets to serve their hedging function. Even though hedgers are long-term traders, they will be reluctant to take positions in contracts that they cannot reverse if they need cash at some point in the future.

Liquidity will likely be underprovided in fledgling information markets for several theoretical and practical reasons. First, liquidity has positive networking externalities. Each trader in a given market benefits from the addition of a new trader, because the additional trader offers all existing traders enhanced trading options.[111] Second, market liquidity in information markets allows for improved estimates of prices, implying that liquidity provides improved information to any economic agent that observes the prevailing market price. Because traders who supply liquidity cannot capture the returns from supplying valuable price information to other decision makers, traders will underprovide liquidity.

Liquidity may also be underprovided if traders do not expect the conditional information market contract trades to ever take place. So, for example, it might be hard to get bids and offers on a contract that was contingent on education reform if no trader expected the government to contract for an education reform project. Because traders can only profit when their conditional trades occur, they must believe that the contingency required for valid trades has a reasonable probability of occurring. Otherwise traders will not acquire information or trade.

One consequence of illiquidity is that individual transaction prices may not be representative of the market participants' beliefs. This can happen if the limit order book is so "shallow" that no trader has an incentive to trade on his superior information.[112] In this case, a great deal of information may have been revealed since the last transaction price, suggesting the last price is a poor measure of market participants' current beliefs. Also, when transactions do take place, they often occur at the bid price for sell orders and at the ask price for buy or-

111. *See* Nicholas Economides & Aloysius Siow, *The Division of Markets is Limited by the Extent of Liquidity (Spatial Competition with Externalities)*, 78 AM. ECON. REV. 108 (1988) (modeling market formation with positive liquidity externalities).

112. A shallow order book describes a situation in which there are only bids and offers in the order book for small quantities of a security and at prices that are not close to the quote midpoint.

ders.[113] This leads to observed transaction prices fluctuating predictably back and forth between the two levels as buy and sell orders arrive at the market. In this case, neither the last bid nor the last ask price is necessarily a good measure of current beliefs. This phenomenon, known as bid-ask bounce, is particularly pronounced in situations where the bid and ask prices are far apart, such as in illiquid markets. In these markets, using the last transaction price as a measure of benefits is problematic.

Fortunately, bid-ask bounce can be mitigated through two simple means. First, most theoretical and empirical market microstructure models show that bid-ask spreads are increasing functions of asset prices.[114] By designing contracts with returns that are very sensitive to variations in the measurement of net benefits, it is possible to minimize price noise, though at the expense of increasing measurement noise. Second, using average transaction prices or midpoints of bid and ask prices will mitigate bid-ask price noise. Because price noise is by definition temporary, this technique will reduce price measurement error. Averaging over transactions from multiple traders also blunts the incentive for an individual trader to manipulate prices.

A more general solution for liquidity-induced problems might ensure that information contract prices reflect the best available information on benefits, rather than manipulation or temporary liquidity shocks. Toward this end, the government could subsidize market-making activity either directly or indirectly.[115] A market maker could post continuous bids and offers for the purchase and sale of each contract. The government could act as the market maker or it could solicit bids from pri-

---

113. *See* Richard Roll, *A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market*, 39 J. FIN. 1127, 1128 (1984).

114. *See, e.g.*, Thomas E. Copeland & Dan Galai, *Information Effects on the Bid-Ask Spread*, 38 J. FIN. 1457, 1468; *see also* Lawrence R. Glosten, *Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices*, 42 J. FIN. 1293 (1987); Lawrence R. Glosten & Paul R. Milgrom, *Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders*, 14 J. FIN. ECON. 71 (1985).

115. Government subsidization of these markets could give incentives to lobby for more information markets that are subsidized, but it should not provide an incentive to manipulate the outcomes of particular markets. Only in the case that particular market outcomes were tied to the overall subsidies would this present a potential problem.

vate parties in an effort to minimize perceived conflicts of interest and the cost of providing this service.

The government also could try other methods of inducing price discovery. For example, it could offer payments to traders who post bids and offers at or inside the market quotes based on the duration of time that their orders remain inside the market quotes and the extent to which their orders improve upon the market quotes.[116] In theory, this type of payment should only encourage liquidity provision and should not alter traders' incentives to take positions in the contracts. This mechanism could be used more generally to establish accurate prices for contracts that are not heavily traded.[117]

A second market design issue concerns the possibility that information market prices could be biased measures of policy benefits. Abramowicz specifically addresses biases in conditional information markets for policies, such as those proposed in this article.[118] He considers the situation in which the policy maker has information unavailable to the market. This implies that the policy maker's decision to implement a policy signals to the market that the policy's net benefits are high, thereby inflating the market's estimate of the net benefits conditioned on the policy being implemented.

For the vast majority of policy issues, this potential selection bias can be solved by forcing the decision maker to disclose all private information or allowing the decision maker to trade with advance notices similar to those required by the SEC in U.S. equity markets.[119] In practice, however, it is unlikely that

---

116. A bid at or inside the market refers to a bid at a price greater than or equal to the highest bid price in the order book.

117. There are other possible mechanisms to encourage liquidity. *See* Michael Abramowicz, *The Law-and-Markets Movement*, 49 AMER. U. L. REV. 327 (1999). Specifically there are the "market scoring rules" discussed in Robin Hanson, *Combinatorial Information Market Design*, 5 INFO. SYS. FRONTIERS 107 (2003), and the periodic call market idea outlined in Nicholas Economides & Robert A. Schwartz, *Electronic Call Market Trading*, 21 J. PORTFOLIO MGMT. 10 (1995).

118. Abramowicz even suggests that information markets will not necessarily perform better than traditional cost-benefit analysis. *See* Abramowicz, *supra* note 7, at 951 ("The ultimate question is whether experts or markets are likely to outperform the other on average, assuming that equal resources are provided for each task. . . . Perhaps the most that can be said on the basis of such experimental data is that information markets and well-motivated experts are roughly comparable.").

119. *E.g.*, Form 3: Initial Statement of Beneficial Ownership of Securities, *available at* http://www.sec.gov/about/forms/form3.pdf; Form 4: Statement of Changes in Beneficial Ownership, *available at* http://www.sec.gov/about/forms/form4.pdf;

the decision maker has more information than the market, rendering this a moot point. Thus, properly designed information markets can overcome potential biases in prices caused by policy makers' private information.

Prices could also be biased measures of policy benefits if information markets fail to aggregate information properly. There is a growing literature that describes the situations in which the prices of asset markets do not accurately reflect publicly available information.[120] In general, the presence of irrational and unpredictable traders can cause inefficiencies in prices when there are limits on the extent of trading by rational traders.[121]

Of course, any process designed to aggregate information will perform worse when it incorporates the beliefs of irrational agents. The real question is whether information markets are more robust to the participation of irrational agents than other mechanisms. Judging by the experimental and field evidence cited above, the answer appears to be an overwhelming "yes." Empirically, market prices are better forecasters than not only the average belief in a population, but also the beliefs of only the experts.

There are also strong theoretical reasons to expect markets to perform better than other mechanisms for aggregating beliefs. First, if irrationality is a result of lazy thinking, then monetary rewards will encourage thoughtfulness. Conversely, the threat of monetary punishment could discourage the participation of irrational agents. Second, markets weigh most heavily the beliefs of the rational agents with the most precise information. These agents place the largest trades, so they have the biggest influence on market prices.

Another potential source of bias can arise because of the framework itself. The basic problem is that traders with information about benefits only profit when the trades in the infor-

---

Form 5: Annual Statement of Changes in Beneficial Ownership of Securities, *available at* http://www.sec.gov/about/forms/form5.pdf; Form 144: Notice of Proposed Sale of Securities Pursuant to Rule 144 Under the Securities Act of 1933, *available at* http://www.sec.gov/about/forms/form144.pdf.

120. *See, e.g.*, ANDREI SHLEIFER, INEFFICIENT MARKETS: AN INTRODUCTION TO BEHAVIORAL FINANCE (2000); *see also* ADVANCES IN BEHAVIORAL FINANCE (Richard H. Thaler ed., 1993).

121. *See* J. Bradford DeLong et al., *Noise Trader Risk in Financial Markets*, 98 J. POL. ECON. 703 (1990).

mation market occur. So they have an incentive to trade not only to exploit their information but also to influence the government's decision to auction off the benefits of a policy so that trades take place. This incentive to influence the government's decision is strongest when the price of the benefits contract with no policy change corresponds to a level of net benefits that is close to the government's minimum acceptable price in the auction.

Suppose, for example, that the current price of the benefits contract with no policy change is too high and informed traders know this. The problem is that informed traders will limit their sales of the overpriced benefits contracts to increase the chances that the minimum acceptable price will not be met in the auction. They will pursue this strategy because they prefer that the contract with no policy change remains in force so they can profit from this information. This could result in a possible upward bias in the information market price for that contract.

The general problem is that if traders anticipate the government's use of the market prices, they will recognize that only programs with higher benefits will be implemented. This anticipation will alter their willingness to buy information market contracts based on policy benefits, potentially biasing the resulting market prices and the government's decision based on those prices.

In subsequent work, we propose a mechanism that deals with this concern by separating the information collection and decision tasks. Specifically, if the information market contracts do not depend on the government's decision rule, then their market prices will be unbiased measures of benefits that can be used by the government. The idea is to have the contract settlement depend on the benefits from a random decision, but to commit to using this random-decision rule infrequently. Most of the time, the government can simply implement its preferred policy, using the unbiased market prices as a guide. Some small fraction of the time, however, the government would need to commit to implementing a particular program, even if it turned out not to be in its best interest to do so. In our model, this is essentially the price of getting better information to improve decisions.[122] In general, decision makers will need to

122. To see the problem caused by biased information and how it can be solved, consider the education reform example above. The problem with the mechanism is that the reform program will only be implemented if an organization is willing

weigh the cost of getting better information against the benefits of that information.[123]

There are other ways to address the problem of biased information. The problem is mitigated, for example, when each individual trader has a small amount of wealth invested in the market relative to the size of the overall market. When no individual trader can influence the price, then all traders simply choose their demands without regard for strategic considerations. The problem is also mitigated when information about the policy is widely dispersed across traders. If no individual trader has enough information to be confident of the true pol-

---

to bid on the project. This will not happen if, for example, no organization anticipates being able to implement the project at a cost less than the expected benefits, which were estimated at $10 million by the market. If traders know that no firm will implement the project, they will not bother trading the information market contract because its payoff depends on project benefits that will never be realized.

Traders will only trade the contract linked to project benefits if they expect the project to be implemented. Unfortunately, even if the contract is traded, the equilibrium price of the contract may be biased. For example, if a trader has information that the project is overvalued at the current price, he may hesitate to bet against the project to the extent that he otherwise might, for fear that the price of the contract would decline to the point that the project is not chosen. In that case, his contract would be worthless. Similar concerns about price bias apply to the information market for test scores that would prevail without education reform.

To avoid this bias, a decision maker can commit to implementing the education reform program with at least some positive probability, say 5%, *regardless* of what information is revealed by the market price. Because a trader's payoff no longer depends on the decision maker's choice, the bias described above will be eliminated. Most of the time (95%) the decision maker can still choose whether or not to implement the education reform program. The disadvantage is that the decision maker must sometimes choose the reform program when this choice will not yield the highest expected net benefits. This situation would arise no more than 5% of the time, and it may be a price that the decision maker is willing to pay for the ability to implement an informed and unbiased decision at least 95% of the time.

123. For an alternative approach to this problem, see Justin Wolfers & Eric Zitzewitz, *Five Open Questions About Prediction Markets*, *in* INFORMATION MARKETS: A NEW WAY OF MAKING DECISIONS IN THE PUBLIC AND PRIVATE SECTORS (Robert W. Hahn & Paul C. Tetlock eds., forthcoming 2005) (manuscript at 14–21, *available at* http://faculty-gsb.stanford.edu/zitzewitz/Research/Five%20Questions.pdf). These authors suggest introducing an "instrumental" variable to identify the causal effect of a policy choice on the price of an information market contract. This approach is often used in econometric analysis to separate correlation and causation. A key problem, as the authors note, is that it is difficult to find good instruments that would actually get rid of the bias. Our approach does not require a naturally occurring instrument.

icy benefits, then no trader can know whether the policy's implementation will occur because of favorable information about true benefits or because of noise trading. In a similar vein, to limit further the potential for bias, the government could induce uncertainty in its implementation decision by not revealing its minimum acceptable price in advance.

A more general concern is that information markets are susceptible to price manipulation by those with a vested interest in the policy decision.[124] This need not be the case, however. Hanson and Oprea show that one of the most common forms of price manipulation can actually enhance the accuracy of real-money information markets.[125] Price manipulation is most likely to occur in information markets when a certain interest group prefers that a particular policy is or is not implemented.[126] This opens the possibility that members of this interest group will try to manipulate the prices in information markets as a form of political rent seeking. These members could conduct trades to push prices in a particular direction, regardless of the true underlying value of the asset. When other traders expect prices to deviate from fundamental value because of manipulation, however, they have a strong incentive to acquire costly information.[127] These informed traders will assume large positions in order to profit from the value-independent trades of manipulators.[128]

---

124. *See, e.g.*, Abramowicz, *supra* note 7, at 937 ("Information markets provide *objective* predictions, though only if concerns such as the possibility of market manipulation can be overcome.").

125. *See* Hanson & Oprea, *supra* note 35.

126. Note that manipulation by interest groups is distinct from the above discussion of manipulation by profit-seeking traders who do not inherently prefer a particular policy. Groups with vested interests are willing to sacrifice profits to influence the implementation decision, whereas profit-seeking traders will only attempt to influence the implementation decision to maximize the expected trading profits they obtain from their private information.

127. *See* Hanson & Oprea, *supra* note 35.

128. Although it does capture some important cases, this argument does not apply universally to efforts to manipulate prices. The proposed market for terrorism, *see supra* note 22, is one such exception: Terrorists have the ability to manipulate information market prices in other ways because they have both insider information and an ability to affect the underlying outcome of this type of information market contract. In addition, the government may not wish to disclose the kind of sensitive intelligence that an information market would reveal. Thus, more research on information markets is necessary before establishing such markets in areas of national security and intelligence.

The interaction between informed traders and price manipulators has many consequences.[129] One effect is a large transfer of wealth from those trying to manipulate prices to those gathering information.[130] As a byproduct of this activity, more information will be reflected in the information market price and this result might discourage attempts to use information markets for rent-seeking purposes.

Consider the U.S. government's goal of deregulating local telecommunications services. Policy makers want to provide appropriate incentives to local incumbents to price their local networks efficiently. To realize this goal, it is necessary to measure accurately the riskiness of the telecommunications business. Companies will generally require a higher price to engage in riskier business activities. Not surprisingly, local incumbents tend to represent their business as risky to obtain a higher price for their resources, whereas prospective competitors represent the incumbents' business as riskless to buy the resources cheaply from the incumbents. Both sides of this debate spend substantial sums of money on legal and consulting fees in an effort to convince policy makers of their position.

Suppose, hypothetically, that the local incumbents are able to spend more money convincing policy makers that their business is risky, and therefore obtain a higher price for their networks. How would the existence of a liquid information market for claims on verifiable business risk alter this balance of power?[131] It might seem that local incumbents would also

---

129. Informed traders may also have an interest in a policy decision and may therefore wish to manipulate the information market price when that price affects the decision. For a formal treatment of this issue, see Robert W. Hahn, Donald Lien & Paul C. Tetlock, Designing Information Markets for Decision Making (November 2005) (unpublished manuscript, on file with authors). The paper suggests how to design efficient information markets when the decision maker uses the price in the information market for a specific decision. The paper also shows that the mere act of linking the decision to the market price will typically enhance liquidity in the information market.

130. Price manipulation only refers to cases where manipulators have no control over the underlying contractual outcome. So, for example, this definition does not apply to terrorists participating in terrorist futures markets, but does apply to private citizens predicting the success of education reform.

131. The most important component of the cost of capital is the cost of equity, which is related to the riskiness of equity. The observed "beta" of the incumbent's stock is a reasonable proxy for this risk. Beta is a measure of how much the stock return varies in tandem with the market return; according to one prominent theory of asset pricing, beta is proportional to expected returns. *See* William F.

spend more in the information market, pushing the price of the security in a direction to suggest their business risk is high. However, if the security becomes priced too highly, speculators will sell the security to profit from the mispricing. By pushing the price back to its fundamental value and correcting the measure of business risk, these speculators will effectively act on the behalf of the prospective telecommunications competitors. In this way, information markets can limit the potential for rent seeking.

A third issue relates to specifying the optimal number of firms for implementing a project. It could be undesirable to have only one firm receive project benefits payments, but beyond the hold-up problem addressed in the project governance section, it is not clear what kind of market power this "monopolist" would wield, as the firm is not given an exclusive right to do anything other than receive contingent benefits payments at a price already determined, in advance, at the social value of the good provided.

One possibility is that the owner of the benefit payments could exercise monopsony power in some input markets. For example, if the project manager is the sole buyer of certain pollution control devices, this could lead to a sub-optimal resource allocation. If a firm's monopsony power depressed input market prices below the socially optimal level, then it would induce reductions in the production of inputs and market exit from suppliers of inputs. Monopsony is probably not a viable long-term position, however, because supply resources rarely remain specialized in the long term. Moreover, to maintain a steady supply of inputs, a forward-looking monopsony buyer will not wield its monopsony power even in the short run.[132]

So far we have shown that it is possible to combine information markets with pay-for-performance contracts to improve decisions, at least in theory. At the same time, we have shown that information markets may not be ideal in some settings. These include when disclosure of information is costly, as it

---

Sharpe, *Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*, 19 J. FIN. 425 (1964). All that is needed to calculate beta is time-series data on a firm's stock price and the price of the market index. Both can be readily verified using monthly data from the Center for Research in Securities Prices.

132. For a discussion of monopsony, see DENNIS W. CARLTON & JEFFREY M. PERLOFF, MODERN INDUSTRIAL ORGANIZATION 107–10 (4th ed. 2005).

may be in the case of national defense;[133] when information markets may underperform experts, which appears to be unlikely in most settings;[134] when it is very costly to get an information market to open; and when there is no verifiable measure of performance.

IV. A BENEFIT-COST ANALYSIS OF INFORMATION MARKETS

In this section we provide a general description of the benefits and costs of information markets. We also compare the direct approach and indirect approaches to performing cost-benefit analysis using information markets.

At the outset, we consider the first-order question of whether such markets could, in principle, address a market failure.[135] From a social standpoint, there could be a failure to provide adequate information, because information is a classic example of a public good. The government is frequently trying to develop policies on the basis of limited information that is strategically supplied by interested parties.[136] There is no reason to think that the provision of such information is optimal. Indeed, there is little reason for the private sector to provide adequate information in government decisions, even when required by law to supply different kinds of information.[137]

As argued in Part III, information markets can help address some of these information gaps. We now turn to a qualitative

---

133. Some information aggregation can be achieved without disclosing sensitive information if an organization adopts an internal information market. Corporations, such as Hewlett-Packard and Eli Lilly, have already begun to adopt internal information markets for obtaining information that could confer advantages over their competitors.

134. For example, if only one individual has a key piece of information, then an information market would not necessarily help if that person were known beforehand.

135. *See, e.g.*, Francis M. Bator, *The Anatomy of Market Failure*, 72 Q.J. ECON. 351 (1958).

136. For example, automobile companies may lobby for pollution controls on oil companies and vice versa. *See, e.g.*, Bruce Yandle, *Bootleggers and Baptists: The Education of a Regulatory Economist*, REGULATION, May–June 1983, at 12, 13.

137. A key problem is that firms neither want to spend the resources nor risk giving valuable information to competitors. *See, e.g.*, BRUCE M. OWEN & RONALD BRAEUTIGAM, THE REGULATION GAME: STRATEGIC USE OF THE ADMINISTRATIVE PROCESS 4 (1978).

description of some of the benefits and costs of using these markets to inform and implement public policies.

There are three broad categories of benefits: a more informed assessment of policy proposals; greater transparency and accountability in the decision-making process; and greater availability of assets for efficient financing and risk sharing, which encourages the implementation of projects with positive net benefits.

A. A More Informed Assessment of Policy Proposals

One of the primary advantages of information markets is that they have the potential, if designed properly, to provide a much more informed assessment of policy proposals.[138] As noted above, they can be very helpful in determining net benefits, distributional implications, and the value of additional research in assessing net benefits. In this section, we explore some other benefits, including the potential for examining new policies, the role that markets can play as assimilators of information, and the benefits of real-time information.

Greater use of these markets could give rise to such proposals as well as a more informed assessment of new proposals. As better information is revealed about the potential impacts of policies, it could encourage policy makers to experiment and fine tune the menu of policies being considered. In addition, because the marginal cost of considering some new policies using information markets is likely to be modest, academics and private citizens may decide to advance proposals of their own and use information markets to estimate their effects.[139]

Similarly, interest groups could advance and analyze their preferred proposals using this approach. Disputes about the likely impact of different policies, such as a gasoline tax, a minimum wage, or a free trade area, could be more easily resolved with data from information markets.[140]

---

138. For a similar argument in support of information markets, see Hanson, *supra* note 7, at 1–9. *See also* Abramowicz, *supra* note 7, at 962–97.

139. *See infra* Part IV.E for a discussion of costs.

140. Disputes surrounding the effectiveness of policy are quite common when the economic stakes are large, with groups taking different sides on the impact of all sorts of proposed legislation and regulation. Some of these conflicts could be addressed by using information markets to measure the impact of various proposals on key variables, such as output and employment. The information from these markets could also be used more frequently by the media as a way to understand and resolve differences of opinion. It would provide a low-cost way

Still another benefit of information markets relates to assessing the effectiveness of different polices in something that approximates real time. Such information could be a potentially large benefit for policy makers who must make time-sensitive decisions. In some cases, agencies have a relatively short period of time to evaluate policy. Performing detailed analysis can often take months, while market prices can be generated from liquid information markets relatively quickly. For example, stock prices and derivative prices can reach equilibrium within minutes or even seconds.[141]

An additional advantage of nearly real-time information is that a comparison of policies can be made across time as key uncertainties are resolved. A major issue in cost-benefit analysis is the extent to which ex ante studies of a particular policy agree with ex post studies. With information markets, the policy maker could choose to examine the expected value of key variables at different points in time as the policy evolves and is implemented.

B. Greater Transparency and Accountability in Decision Making

One of the advantages of cost-benefit analysis is that it has the potential to make the analytical basis for public policy decisions more transparent.[142] The same argument can be made for information markets. By supplying information on key parameters that are involved in decision making, information markets make the public policy process more transparent.[143]

---

for journalists, who often operate under severe time constraints, to see how different groups are trying to shape issues.

141. *See* Joel Hasbrouck, *Intraday Price Formation in U.S. Equity Index Markets*, 58 J. FIN. 2375, 2387–88 (2003) (estimating that the prices of S&P 500 equity index futures respond to new information within roughly sixty seconds).

142. *See, e.g.*, ARROW ET AL., *supra* note 3, at 1–4; *see also* ROBERT W. HAHN & ROBERT E. LITAN, IMPROVING REGULATORY ACCOUNTABILITY (1997), *available at* http://aei-brookings.org/admin/pdffiles/phpf4.pdf; SUNSTEIN, *supra* note 2, at 106–08.

143. Indeed, if the government wished to go a step further, it could either directly or indirectly help make information market prices, volume, and order books available at very low cost or no cost. However, this information is probably most efficiently conveyed by the exchanges listing information market securities, which need not be subsidized by the government. These exchanges have obvious incentives to convey some, if not all, relevant information about buying and selling activity in their securities. *See* discussion *infra* Part IV.E.

Greater transparency is likely to increase accountability for lawmakers and regulators. Making the data more readily available means that the cost of monitoring regulators and lawmakers decreases. Interested parties, then, will find it easier to hold them accountable for decisions. Abramowicz makes a similar point, arguing that information markets have the potential to help discipline the agency decision-making process.[144] Indeed, even if the government does decide to pursue a more narrowly self-interested agenda, as public choice theory might predict,[145] the introduction of such markets could place constraints on what the government can do.[146]

Greater transparency could also reduce the scope for political manipulation. In traditional cost-benefit analyses, it is easy to imagine cases where the analyst may be biased or captured by a particular constituency.[147] Information markets, by improving the quality of information along with access to that information, reduce the potential for certain kinds of manipulation. It is possible, though by no means certain, that the level of rent-seeking would decrease if information markets for policy were used widely. This is because it would become more costly to change policy outcomes if something like our new approach were implemented.[148]

Consider an example: Suppose the government is considering a law that would regulate the prices of pharmaceuticals. Economists frequently argue that price controls would reduce incentives to develop new drugs and, therefore, reduce the number of new drugs that are introduced into the market. At the same time, others argue that such controls would make little difference. Well designed information markets could shed

144. *See* Abramowicz, *supra* note 7, at 934 ("Information markets could help constrain administrative decisionmaking and limit ideological decisionmaking.").

145. *See, e.g.*, JAMES M. BUCHANAN & GORDON TULLOCK, THE CALCULUS OF CONSENT: LOGICAL FOUNDATIONS OF CONSTITUTIONAL DEMOCRACY (1962).

146. In what follows, we generally presume that the government is interested in maximizing some measure of social welfare as it is typically operationalized. Even if it is trying to achieve some other goal, information markets could be helpful. We abstract from the general difficulties of defining social welfare. The seminal work here is KENNETH J. ARROW, SOCIAL CHOICE AND INDIVIDUAL VALUES (1951).

147. *See* Abramowicz, *supra* note 7, at 970 (discussing how information markets could reduce the potential for interest group manipulation).

148. We recognize that rent seeking is difficult to measure. The general idea is simply to try to reduce the returns on rent seeking over a wide range of activities. *See, e.g.*, Anne O. Krueger, *The Political Economy of the Rent-Seeking Society*, 64 AM. ECON. REV. 291 (1974).

light on this issue by estimating the rate of new drug approval under the two policies. If the public and the media came to see such markets as a trustworthy source of information, these markets could have a greater influence on policy, and the influence of interest groups on both sides of the issue could decrease.[149]

Information markets could also improve decision making by increasing the level of expected net benefits from policy choices. Regulators, lawmakers, and the courts might find it in their interest to make greater use of data on benefits and costs if those data were more accurate. These policy makers would then become more accountable in the sense that they would be more inclined to use data on costs and benefits to justify their decisions. In addition, lawmakers and regulators may implement procedural changes that increase the likelihood that data on costs and benefits will be considered seriously.[150]

Finally, the introduction of information markets could lead to a change in bureaucratic focus. Fewer resources may be allocated to traditional cost-benefit analyses and more to the construction of information markets to inform the process. Such a reallocation would not necessarily lead to a reduction in overall expenditures. That reduction depends on the value of the information that is produced from the information markets and the extent to which it compliments or substitutes for existing information. One could even imagine a case in which net time and effort spent on policy analysis actually increases, if time and effort spent analyzing data from information markets increases at a greater rate than time and effort spent on conventional cost-benefit analysis declines.

C. Greater Availability of Assets for Financing Projects and Spreading Risks

The creation of information markets for claims on the benefits from policy decisions has the potential to facilitate the im-

149. A current problem is that it is difficult for the public to know what may have occurred in situations that are not actually observed. Sam Peltzman argues this is one of the reasons that inefficient regulation persists. *See* SAM PELTZMAN, REGULATION AND THE NATURAL PROGRESS OF OPULENCE (2004). Information markets provide a solution to this problem.

150. Examples include a regulatory budget and regulatory net benefit accounts. *See* discussion *infra* Part V.B.

plementation of socially valuable projects. By aggregating information from diverse sources, the prices in these markets reveal the social value of policy proposals. Using these prices, the government can then monetize this value and package it in the form of a security. This security would be identical in its payoffs to the claim on project net benefits, but it would also confer to its owner the right to vote on project implementation decisions. The government could collect the surplus on projects by auctioning off the right to execute the project: in other words, by selling the security. The government's position in such a transaction is analogous to that of an entrepreneur auctioning his project to venture capitalists.[151]

In addition, as discussed above, it is also possible to design contracts that enable private sector participants to reduce or eliminate much of the risk associated with public policies. This applies to the firms actually implementing the policies, but it also extends to consumers and workers. Those parties affected by economic policies have an opportunity to mitigate the uncertain impact of the policy by purchasing information market contracts. For example, consumers concerned about price increases caused by a policy could purchase contracts that yield payments when the policy results in price increases. Similarly, workers concerned about job losses resulting from a policy could purchase contracts that yield payments when the policy creates more unemployment.

Finally, in addition to serving as a risk management tool for firms, information markets can help firms raise cash when they are constrained. While some firms have ready access to capital markets, others may not. The information market can serve as a source of project funding for both types of firms.

D. Cost of Information Markets

Several costs are related to information markets. These costs can be minimized by setting up an electronic exchange for the purchase and sale of contracts, similar to the electronic communication networks that facilitate trades in U.S. stock mar-

151. Robert Shiller identifies several markets that could be used to help increase efficiency and equity. These include markets for claims on individual and national income and human capital. While Shiller's proposals are not strictly information markets as we have defined them, information markets have the potential to improve risk management substantially. *See* ROBERT J. SHILLER, MACRO MARKETS: CREATING INSTITUTIONS FOR MANAGING SOCIETY'S LARGEST ECONOMIC RISKS (1993).

kets. The fixed costs of establishing this electronic exchange will include payments to programmers who develop the web exchange application and hardware costs for servers and storage devices necessary for the daily operation of the exchange.

The operating costs would, at a minimum, include monitoring, maintenance, and upkeep costs for addressing ongoing exchange issues. Legal procedures and controls are needed to maintain a reputation for the market so traders perceive that they are treated fairly. Operating costs could also include direct or indirect capital outlays that ensure sufficient liquidity in exchange contracts to retain traders' interest. Outlays could take the form of commission rebates or negative commission on trades, order flow, or some other variable of interest.

These operating costs may be offset somewhat by exchange revenues, which include commissions charged by the exchange. If the exchange also operates as a market maker, its revenue would be dependent upon the difference between the prevailing bid and ask prices on each contract. In this case, the exchange (or any market maker) also incurs the risk that the value of its contract inventory (net contract position) declines. Through careful risk management and research into exchange trading behavior, this risk can be kept to a minimum. Alternatively, the exchange could hire a market maker to share or completely eliminate this risk.

It would be desirable for an exchange to set up a mechanism, such as a public Internet forum, for proposing new contracts. The direct cost of listing new securities on the exchange for trading would be trivial, but indirect costs could be large. Adding too many new securities too quickly could detract from the liquidity in existing markets, thereby undermining the purpose of the exchange: facilitating the meeting of buyers and sellers by coordinating their demands and supplies in time and space. On the other hand, if adding a new security encourages new trading volume, meeting this volume could require the exchange to upgrade its hardware and increase the scale of its operations. Exchanges will need to balance these considerations in choosing which securities to approve for listing.

E. Comparison of Direct and Indirect Approaches for Information Markets

This paper has focused on a direct approach to estimating net benefits using information markets. An alternative to this direct approach is an indirect approach, where the market predicts what an expert will say about net project benefits at some time in the future. In other words, the market predicts the conclusion of an expert's cost-benefit analysis.[152]

Here, we will briefly compare the two approaches. A summary of these approaches is provided in Table 1. The table reveals that both approaches can estimate net social benefits and also address equity concerns. In general, however, the direct approach is preferable because it can perform just as well as the indirect approach. Furthermore, it has the advantage that it avoids introducing additional uncertainty related to the selection of an expert.

The reason the direct approach possesses most, if not all, of the capabilities of the indirect approach is because we allow the government to change its policy. The government is allowed to buy back contracts and issue new ones in light of new information or improved measures of benefits. Thus, it can redefine the benefit function, and hence, the policy, over time as circumstances warrant.

The flexibility of the decision rule when using an indirect approach has numerous disadvantages. Because there are few, if any, restrictions on what the analyst may include in a cost-benefit analysis, the price of an information market contract is extremely difficult to interpret. It is impossible to know what sort of cost-benefit analysis market participants anticipate will be used in the future. In other words, the information market's data will compound both the expected benefits and the expected way in which the analyst will conduct the benefits determination, producing data that is potentially less valuable.

Abramowicz argues that if the selection of the analyst occurs sufficiently far in the future, the market beliefs about the analyst's future cost-benefit analysis will yield an unbiased measure of the future expert's estimates of net benefits.[153] This may

152. *See, e.g.,* Abramowicz, *supra* note 7, (discussing the advantages of using information markets to forecast the retrospective cost-benefit analysis conclusions of such a future expert).

153. *See* Abramowicz, *supra* note 7, at 939, 991–92.

be true, but it overlooks some potential problems with the expert's estimate.

Although this procedure potentially overcomes expected political biases and policy preferences, it does not preclude biases introduced by the market's anticipation of the analyst's information access, information gathering, and information processing abilities.[154] For example, the market may believe that the future analyst will overlook certain types of economic benefits, perhaps because typical analysts performing cost-benefit analyses have tended to overlook these types of benefits in the past. More generally, there is no way to know whether the market expects the future analyst to omit certain types of benefits and costs that result from the policy. Essentially, the market's attempt to forecast the subjective decisions of the analyst, rather than objective measurements, needlessly introduces "noise" into the information conveyed by the market's prices. This noise reduces the value of the market as a general predictive mechanism.

The analyst is just one of many experts who could have different pieces of information that would inform a judgment of net benefits. The current price will reflect only the public's knowledge of the expected analyst's knowledge. The requirement that prices of information market contracts be based only on one imperfectly informed analyst's future judgment defeats much of the point of having information markets.

Also, the unspecified future analyst will not have a clear financial incentive to gather and process information accurately in her determination of net benefits.[155] In contrast, participants in information markets have strong monetary reasons to be thoughtful, resourceful, and meticulous in estimating future net benefits.

Finally, decision making based on the anticipation of an unspecified analyst doing an unspecified cost-benefit analysis is extremely opaque and, therefore, unlikely to inspire public confidence and trust in the decision-making process. The direct method, by comparison, provides a transparent way of valuing

---

154. This is true even for the case in which the analyst herself decides what should count as a benefit or cost. She is still not omniscient when it comes to measuring those benefits and costs.

155. In correspondence with the authors, Abramowicz notes that professional pride could serve as a partial substitute for financial incentives.

benefits and costs. Transparency is especially important to the success of an information market because traders are reluctant to participate in financial markets they do not trust or understand.

The indirect approach to estimating net benefits is inconsistent with the guiding concept of securities laws and rules. The SEC mission statement explains that "all investors . . . should have access to certain basic facts about an investment prior to buying it."[156] Designed to restore public confidence in financial markets, the Securities Exchange Acts of 1933 and 1934 stress disclosure of all material risks involved in investing.[157] A transparent explanation of risks is extremely difficult for a security whose value depends on the whims of an unspecified future analyst. By contrast, the risks in information market contracts based on transparent quantitative measures could be estimated using mathematical models.

V. SIGNIFICANCE FOR POLICY DESIGN AND EVALUATION

We have presented a framework for designing contracts that provides information on the costs, benefits, and net benefits of different policy proposals. In this section, we consider some broader issues: first, how this information could be used in the policy process; second, how it could substantially change the nature of legislative and regulatory oversight; and third, possible roles for the government in facilitating and evaluating information markets, and roles for researchers in evaluating such markets.

A. Information Markets in the Policy Process

First, the potential scope for information markets in decision making is enormous. In theory, information markets can address many gaps in our knowledge of both public and private

156. *See* U.S. Securities and Exchange Commission, The Investor's Advocate: How the SEC Protects Investors and Maintains Market Integrity, *available at* http://www.sec.gov/about/whatwedo.shtml.

157. The Exchange Act established the SEC in response to the investor disenchantment with securities markets that followed the stock market crash of 1929 and the Great Depression. *See* The Securities Act of 1933, 48 Stat. 74, (codified at 15 U.S.C. §§ 77a–bbbb); *and* The Securities Exchange Act of 1934, 48 Stat. 881, (codified as amended at 15 U.S.C. §§ 78a–*lll*).

decision making.[158] More generally, liquid information markets facilitate the allocation of capital to value-maximizing projects in the public sector, inform resource allocation decisions in the private sector,[159] and enable unique risk-sharing possibilities throughout the economy.[160]

The benefits of information markets are not limited to governments; there are a large number of potential applications to international institutions, non-governmental organizations, and foundations. It is quite common for these institutions to fund or design projects. The World Bank, for example, funds projects in developing countries. Information markets can help in the design of specific projects as well as with the ranking of those projects.[161]

Ventures in the private sector could also benefit from improved access to information and hedging vehicles. A natural extension of Germany's initial public offering (IPO) market would be an IPO market that conditions contracts on a firm's implementation strategy.[162] Hanson has described a similar secondary conditional market in which traders could directly bet on the effect of firing a company's CEO on the stock price of the company.[163]

There are also a number of potential applications of information markets to corporate finance. One of the current problems in corporate finance is delivering accurate information to the

---

158. *See* Abramowicz, *supra* note 7, at 962 ("Information markets in theory might be used in any decision making environment, and corporations have experimented to determine how effectively information markets aggregate information. . . . Information markets are best justified [for government] as a means for disciplining government decisionmaking.").

159. For more on applying information markets to both the private and public sectors, see Abramowicz, *supra* note 7; Hanson, *supra* note 7.

160. *See* SHILLER, *supra* note 151.

161. *See, e.g.*, Robert W. Hahn and Paul C. Tetlock, *Making Development Work*, POL'Y REV., Aug.–Sept. 2005, at 27, 32–34. One of the applications considered in that article in some detail is the Copenhagen Consensus, which aimed to rank approaches for addressing some of the world's most serious problems. Specifically, in 2004, a group of eight distinguished economists assembled in Copenhagen to see if they could achieve some consensus on the best ways of advancing global welfare if an additional $50 billion were made available to government. Information markets can serve as a valuable source of information in framing this important debate. *See id.*

162. For example, the when-issued contracts in benefits claims could be conditioned on alternative organizational structures or compensation systems.

163. *See* Hanson, *supra* note 7, at 25.

CEO so that she can decide which projects to fund. Self-interested project managers, desiring implementation of their own projects, do not necessarily have an incentive to supply accurate information. Information markets would be ideally suited to aggregating such information if it resided with several different employees in the company. This approach could reduce problems of rent seeking on the part of some key members in the firm.[164]

Athanasoulis et al. describe the importance of markets based on key macroeconomic indices.[165] In addition to conferring the benefits stressed by the authors, these markets would facilitate the spread of accurate economic statistics to agents whose decisions could benefit from this information. In this respect, macro markets are similar to information markets, but they serve another important purpose. As emphasized by Athanasoulis et al., individuals, firms, and even national governments could derive enormous benefit from the formation of liquid markets that allow them to minimize unnecessary fluctuations in their incomes.[166]

While there is a huge potential scope for information markets, the test will be in their application. One of the critical issues is how they are designed, researched, and applied in particular contexts.

A critical question is how information from these markets should be used in decision making. Hanson suggests that these markets should be decisive in certain contexts.[167] While we would not rule out making these markets decisive in certain settings, we think it is premature and that his proposal is overly ambitious.

Our general approach differs from Hanson's in a number of ways. First, we would allow decision makers to consider other factors in deciding whether to move forward with a project.

---

164. Managers and CEOs are sometimes modeled as having an interest in empire building, which is a form of rent seeking. *See, e.g.*, Jeremy C. Stein, *Internal Capital Markets and the Competition for Corporate Resources*, 52 J. FIN. 111 (1997). Information markets have the potential to reduce rent seeking in both private and public decisions: By giving principals more information, these markets expand the set of feasible incentive contracts, enabling improvements in firm efficiency.

165. *See* Stefano Athanasoulis et al., *Macro Markets and Financial Security*, ECON. POL'Y REV., Apr. 1999, at 21.

166. *See id.*

167. *See* Hanson, *supra* note 7, at 2 ("The basic rule of futarchy would then be this: when speculative markets clearly estimate that a proposed policy would increase expected national welfare, that policy becomes law.").

Second, we would not focus on GDP as a measure of welfare in most instances. Third, we believe it is important to look for proposals that are likely to maximize net benefits as opposed to implementing those that simply increase net benefits. Fourth, Hanson focuses solely on the role of information markets in improving policy, whereas we highlight the link between information markets and using pay-for-performance contracts to provide incentives for efficient implementation of projects.

More knowledge is needed about how these markets perform in practice in different settings. Furthermore, one needs to weigh the costs and benefits of using a decision rule as opposed to allowing decision makers to exercise judgment. This is a well studied problem in a number of contexts, including the choice of whether to have a monetary rule.[168]

In general, these markets will be most useful when they can closely approximate the variables of concern to interested parties and provide good measures of those variables. Thus, they should play a more prominent role when evidence indicates they can provide more useful information.

B. Solving Some Difficult Government Oversight Problems

One of the great problems facing lawmakers is how to design a legal and regulatory process that is transparent and also holds lawmakers and regulators accountable for the laws and regulations that they make. In addition to calls for greater transparency, several scholars have suggested mechanisms for gaining a better appreciation of the impacts of regulatory spending and overall spending. For example, Robert Crandall and Christopher DeMuth both propose a regulatory budget that would constrain the costs that an agency could impose on society with its regulations, or more generally with its policies.[169] Such a budget could encourage a focus on projects that have a high social payoff. Eric Posner suggests a similar idea,

---

168. *See, e.g.*, Finn E. Kydland & Edward C. Prescott, *Rules Rather than Discretion: The Inconsistency of Optimal Plans*, 85 J. POL. ECON. 473 (1977) (discussing the advantages and disadvantages of a monetary rule).

169. *See* Robert Crandall, *Is Government Regulation Crippling Business?*, SATURDAY REV., Jan. 20, 1979, at 31, 34–40; Christopher C. DeMuth, *Constraining Regulatory Costs—Part II: The Regulatory Budget*, REGULATION, Mar.–Apr. 1980, at 29; *see also* ROBERT E. LITAN & WILLIAM D. NORDHAUS, REFORMING FEDERAL REGULATION (1983); HAHN & LITAN, *supra* note 142, at 13–19.

but extends the concept to net benefits.[170] In his formulation, each agency could have a net benefit account based on the estimated net benefits of the regulations it promulgated.[171]

While scholars have advanced these proposals, both Congress and the executive branch have introduced a number of procedures and requirements aimed at improving regulation and getting a better understanding of its impacts.[172] For example, Congress now requires the Office of Management and Budget (OMB) to do an annual report on the costs and benefits of federal regulation.[173] In its report, OMB must estimate the total annual costs and benefits of federal rules, analyze the impacts of federal regulation on state, local and tribal government, small business, wages, and economic growth, and provide recommendations for reform.[174]

In addition, all presidents since Jimmy Carter have implemented some form of regulatory oversight.[175] Beginning with President Reagan, as a result of the regulatory impact analysis requirement in Executive Order 12,291,[176] agencies are providing more information than ever before about the benefits and costs of federal regulation.

---

170. *See* Eric A. Posner, *When Reforming Accounting, Don't Forget Regulation*, POL'Y MATTERS, Aug. 2002, *available at* http://www.aei-brookings.org/policy/page.php?id=104 (arguing that the institution of Net Benefit Accounts (NBA) will promote more efficient regulations and stating, "The well-run agency could build up a surplus in its NBA, allowing it to take risks with regulatory projects that might initially fail a cost-benefit analysis but, in the end, produce a net benefit for society").

171. *See id.*; *see also* Robert W. Hahn, *The Economic Analysis of Regulation: A Response to Critics*, 71 U. CHI. L. REV. 1021 (2004).

172. While the examples used here focus on regulation, this is done in the interest of avoiding abstraction. The applications to oversight could be much broader. The only requirement is that the government be able to specify some measure of the output or results of its resource allocation decisions.

173. *See* Robert W. Hahn & Mary Beth Muething, *The Grand Experiment in Regulatory Reporting*, 55 ADMIN. L. REV. 607 (2004), *available at* http://www.aei-brookings.org/publications/abstract.php?pid=314.

174. *See, e.g.*, OFFICE OF MGMT. & BUDGET, REPORT TO CONGRESS ON THE COSTS AND BENEFITS OF FEDERAL REGULATIONS (1998), *available at* http://www.whitehouse.gov/omb/inforeg/costbenefitreport1998.pdf.

175. President Carter created the Office with the Paperwork Reduction Act of 1980, Pub. L. No. 96-511, 94 Stat. 2812 (codified at 44 U.S.C. §§ 3501–3520) but the Act was not implemented until the end of 1980. Before Reagan introduced the Regulatory Impact Analysis requirement in 1981, data on the benefits and costs of regulation were sparse.

176. Exec. Order No. 12,291, 3 C.F.R. 127 (1981), *reprinted in* 5 U.S.C. § 601 (2000), *available at* http://www.archives.gov/federal-register/codification/executive-order/12291.html.

The effectiveness of many of these requirements and proposals rests on their ability to gather useful information on the costs and benefits of particular regulations. Take the regulatory budget as an example. The Achilles heel of the regulatory budget has always been the difficulty of getting reasonable estimates of the cost of regulation. Should the government use its own estimates, those of the affected businesses, or some other source? A similar problem plagues net benefit accounts, except information on both costs and benefits is needed. Even if using direct measures of costs and benefits, such as compliance costs and benefits such as lives saved, the problem can be formidable.

Information markets can aid substantially in solving the problems that have plagued the regulatory budget, net benefit accounts, and similar proposals. They do this by presenting reasonably unbiased estimates of the costs and benefits of particular regulations. As noted above, these estimates are not perfect.[177] They are likely, however, to be substantially better than those generated by conventional methods in many settings.[178]

Moreover, with the advent of widespread information markets, the nature of the regulatory oversight problem could change. It may no longer be necessary to have a regulatory budget if the expected net benefits of regulations can be readily estimated in a transparent manner that is viewed as legitimate.[179]

Net benefit accounts may not be needed for a similar reason. While net benefit accounts give agencies some incentive to promote economic efficiency, a simpler rule could be enacted. For example, if the net benefits of a particular rule did not pass a benefit-cost test according to data from information markets, then the agency could be asked to provide more justification for implementing the regulation.[180]

---

177. *See supra* Part III.D.4.

178. *See supra* Part II.

179. This presumes that such markets would be viewed as legitimate, which likely depends on how they evolve over time. In addition, if agencies still have a natural tendency to regulate in their area, perhaps due to "tunnel vision," a regulatory budget could be a helpful antidote. *See* BREYER, *supra* note 1, at 11–19 (discussing the problem of tunnel vision in some government agencies).

180. *See* ROBERT W. CRANDALL ET AL., AN AGENDA FOR FEDERAL REGULATORY REFORM 13 (2003).

In fact, greater use of information markets for assessing the benefits and costs of polices could fundamentally change the nature of the oversight process. Indeed, it could transform the process into something that is much more familiar to legislators: a budgetary process. At the core of contracts to elicit information on benefits and costs is a reward system. In the case of public policy, the government makes payments based on the actual benefits of a proposed policy. That is, the agency or the government is expected to pay out money in exchange for benefits.[181]

Thus, Congress, or other legislatures, may want to address the problem in a different way. Congress would need to allocate budgets so that agencies could essentially offer market-based contracts that increase expected social welfare or net benefits. Congress could, for example, pass budgets for particular policy areas such as telecommunications regulation, environmental, health and safety regulation, and health care. By imposing a budget constraint, Congress would be limiting the deadweight losses associated with raising revenues, and imposing discipline on agencies similar to a regulatory budget. Agencies could petition Congress in special cases if they could show that net benefits were particularly high for certain projects but they lacked adequate funds to support the projects. Alternatively, agencies with a very good track record of generating net benefits could be rewarded with larger budgets in subsequent budget cycles.[182]

Consider the case of environment, health and safety regulation, addressed in Justice Breyer's *Breaking the Vicious Circle*.[183] Justice Breyer, among others, has suggested trying to save more lives with a given level of resources by considering reallocating expenditures to activities that would save the most lives.[184] Information markets could easily accommodate this problem. If another objective were viewed as desirable, such as

---

181. It is not necessary that the money be paid by the government or come from any particular source, but some credible promise to pay needs to be made.

182. This is similar to Eric Posner's suggestions for net benefit accounts. *See, e.g.*, Eric A. Posner, *Controlling Agencies with Cost-Benefit Analysis: A Positive Political Theory Perspective*, 68 U. CHI. L. REV. 1137 (2001), *available at* http://www.law.uchicago.edu/Lawecon/WkngPprs_101-25/119.EP.Controlling%20Agencies.pdf.

183. *See* BREYER, *supra* note 1.

184. *See id.*; *see also* Tammy O. Tengs & John D. Graham, *The Opportunity Costs of Haphazard Social Investments in Life-Saving*, *in* RISKS, COSTS, AND LIVES SAVED: GETTING BETTER RESULTS FROM REGULATION 167 (Robert W. Hahn ed., 1996).

age-adjusted lives saved, quality-adjusted life years, or willingness to pay, information markets could be reconfigured to predict measurements of that objective under a variety of policy alternatives. Thus, if the value of a statistical life should differ across agencies,[185] contracts could be written to accommodate that characteristic by adjusting the payout depending on the lives expected to be saved.

Congress could apply the same logic to other areas of policy, such as public works projects. There is no reason, save perhaps politics, that Congress could not apply the same rigor to dam building or environmental restoration projects.[186]

In summary, we are suggesting that more direct measures of the benefits and costs of different policies can help administrators solve some vexing problems of legislative and regulatory oversight. In addition, these measures may change the nature of the oversight process to something that is more familiar to lawmakers: allocating government expenditures by agency.

C. The Potential Role for Government and Researchers

Because the potential benefits from the use of information markets are so large, government needs to examine its potential role in facilitating information markets for promoting better public policy. We suggest that government use an iterative approach in which it initially experiments with information markets on a small scale.[187] It should strive to obtain more information about the value of information markets and the impact of different policy interventions. After learning more about which strategies work through experimentation, the government could revise its strategy accordingly.

To assess the appropriate role for government, we need to know something about the potential net benefits that could

---

185. *See* Cass R. Sunstein, *Are Poor People Worth Less Than Rich People?: Disaggregating the Value of Statistical Lives* (AEI-Brookings Joint Center, Working Paper No. 04-05, 2004), *available at* http://www.aei-brookings.com/publications/abstract.php?pid=430.

186. For an early treatment of the politics of such issues, see JOHN A. FEREJOHN, PORK BARREL POLITICS: RIVERS AND HARBORS LEGISLATION, 1947–1968 (1974).

187. In many cases, it may be difficult to formulate an appropriate benefit function. Doing experiments on a small scale could help uncover problems in the specification of the benefit function, ensuring that policy makers get the benefits they intend. For example, municipalities eager to undertake policy experiments could serve as trial grounds for future statewide and nationwide policies.

flow from information markets. At this point, we know very little in a quantitative sense.[188] We also need information on the potential value of different methods of government intervention.

We offer one possible template for government involvement that consists of five parts. First, government could pursue a "no-regrets" policy in which it attempts to reduce inefficient regulatory barriers to information markets.[189] The current regulatory environment is highly uncertain for information market contracts, which is a barrier in itself. Neither the federal Commodity Futures Trading Commission (CFTC) nor the individual state gambling commissions have an obvious legal claim to regulating information markets. We recommend that the CFTC take decisive action to eliminate this uncertainty by assuming control of the review and approval process for newly introduced information market contracts.[190]

Second, the government could facilitate the development of an online exchange that specifically addresses these kinds of issues. The costs for setting up the exchange itself should be relatively modest. Subcontracting the job to an existing exchange (or exchanges) would be desirable to obtain economies of scale and expertise. In principle, this would be similar to the way the U.S. government hires private underwriters, such as Goldman Sachs or Merrill Lynch, to assist in the issuance of treasury securities. The exchange could also handle contract clearing and settlement, adjudicate trade disputes, and verify the events that form the basis of its listed contracts.

---

188. We would advocate the use of a "meta" information market to determine the expected benefits of the use of information markets. For example, establish several when-issued markets for U.S. GDP, conditional on different amounts of funds allocated to government-sponsored information market research. However, a meta-market might not satisfy those skeptical of the utility of information markets due to the circular nature of the assessment.

189. Currently, the Commodity Futures Trading Commission is considering whether event markets fall under its jurisdiction. If it decides to assert jurisdiction over event markets, even non-profit markets established for research purposes would be forced to adhere to the Commodity Exchange Act. The law would require all event markets to endure a long and costly designation process in order to demonstrate that market transactions are safe and secure. *See* 7 U.S.C. § 6 (2005).

190. For a more detailed discussion arguing that state regulation aimed at Internet gambling is generally not appropriate for regulation of information markets, see Robert W. Hahn & Paul C. Tetlock, *A New Approach for Regulating Information Markets*, J. REG. ECON. (forthcoming), *available at* http://www.aei-brookings.com/publications/abstract.php?pid=881.

The government should consider whether it wants to facilitate the coordination of contract supply and demand by providing liquidity—that is, the posting of bids and offers that enable buyers and sellers wishing to trade the option to do so.[191] By narrowing the bid-ask spread, the government would provide more profit opportunities for traders. This would encourage traders to reveal their existing private information through their trades and it would provide incentives for traders to acquire new information.[192]

Third, the government should also consider conducting or funding specific policy experiments, or both. These experiments should be selected so that they have a reasonable chance of success and are not controversial. For example, a controversial market related to predicting the probability of a terrorist event in some time frame would not make sense, at least at the outset. A market for claims on GDP conditional on the passage of a proposed tax cut might make sense. This market would potentially attract a great deal of expert, popular, private, and public interest. Such a market would also be unlikely to incite a public outcry.

Fourth, the government could experiment with stimulating new policy suggestions that have a potentially high payoff. It might, for example, provide a reward for specific policy suggestions that turn out to be "winners" and are adopted. The value of the reward could be equal to a fraction of the difference between the net benefits on the best proposal and the second-best proposal. Alternatively, the government could simply work to lower the cost of introducing new policy ideas into the debate by setting up and monitoring online idea forums for proposals.

---

191. Offering liquidity is analogous to selling an option to trade. *See* Thomas E. Copeland & Dan Galai, *Information Effects on the Bid-Ask Spread*, 38 J. FIN. 1457 (1983). The price of the option is equal to one half of the bid-ask spread.

192. The government should try to avoid holding contract inventory because it could send misleading signals to market participants and lead to avoidable ethical objections to market intervention. To meet this goal and still satisfy its market making function, the government could adjust the bid-ask prices downward or upward so that its net inventory balanced to zero. If the government wished to mimic the behavior of a competitive market maker, the bid-ask spread would widen or narrow depending upon equilibrium loss or profit for the government. The spread required to maintain zero profits will depend upon the number of informed and uninformed traders in the market. More informed traders implies a larger spread.

Fifth, the government could help make price and order book information from information markets available at a very low cost. Requiring exchanges that list information market contracts to release real-time order book information publicly (via the Internet, for example) would be a non-intrusive and virtually costless method of providing a public good. On the other hand, it makes sense to allow private parties an opportunity to supply the information before imposing this requirement.[193]

Even if there were widespread adoption of information markets to improve policy, the government would still need to consider how the new information should affect its analyses of costs, benefits, and other key factors. It is important to recognize that the process of assessing net benefits will not necessarily be automatic, even with extensive information markets. In many cases, the prices in information markets can only approximate the objective function of the decision maker due to the limitations of specifying contracts. The decision maker may want to include other factors or information in an analysis. For example, if only the direct cost to the government is estimated, some factor would need to be added for the deadweight cost of raising revenues.[194]

Because the process of assessing net benefits from a policy will not necessarily be automatic, there is likely to be an important role for government analysts. That role should depend on the quality and scope of information obtained from information markets. If the quality is high and the information essentially answers the desired question, the decision maker may decide there is less of a role for the kind of government policy analysis that is done today. At the same time, independent analysts may be more inclined to provide such analysis because the basic information inputs to this analysis are easier to obtain.[195]

---

193. It is unclear that private exchanges will display order book information to all parties whose decisions would benefit from this information. Not all such parties are prospective investors: liquidity-constrained consumers, for example.

194. *See, e.g.*, Feldstein, *supra* note 55; *see also* ANTHONY B. ATKINSON & JOSEPH E. STIGLITZ, LECTURES ON PUBLIC ECONOMICS (1980). In principle, this factor could be added to the benefit function.

195. We might expect experts, hired by those with a stake in the outcome, to develop and publicize "independent" reports about policy consequences in an attempt to influence the information market's evaluation of those consequences. Based on the minimal impact of stock analyst reports on prices, the impact of such experts is likely to be small. At the same time, such experts might affect the policy options considered by the government.

As with any problem involving new technology or information, the government will need to learn how best to use that information in conjunction with other sources. It may, for example, want to issue guidelines for cost-benefit analysis that highlight the importance of using information markets when they are useful. These could be included in economic guidance provided by the OMB.[196]

Both the government and researchers have a very important role to play in the design and evaluation of information markets. The government may choose to introduce procedures for evaluating the performance of these markets. An advisory board consisting of distinguished researchers could determine how to optimize contract structure, exchange rules, and government involvement to meet the objectives of information markets. Possible considerations include the ability of prices to forecast net benefits, the ex post net benefits of policies that are implemented, the costs of operating the markets, the scope for price manipulation, and the hedging opportunities afforded by contracts. The advisory board could examine, for example, how increases in market stakes and liquidity impact the forecasting accuracy of prices. Investigations such as this will be necessary to ensure that information markets are productively employed in public policy decision making.

Independent researchers can contribute by working to resolve other design and performance issues. First, much more thinking needs to be done about where such markets could be applied successfully. As there are a large number of potentially important applications, one of the problems in the short term may be to choose those with the highest potential payoff. For example, in the field of regulation, it probably makes sense to focus first on major regulations that have impacts on the economy of at least $100 million annually. Over time, analyses can help shed light on the optimum number of such markets, particularly for purposes of policy design. In addition, more thought needs to be given to the kind of markets that are most useful. For example, the potential value of using information markets to provide information on the equity implications of

196. *See, e.g.*, Office of Mgmt. & Budget, Appendix C: OMB Draft Guidelines for the Conduct of Regulatory Analysis and the Format of Accounting Statements, 68 Fed. Reg. 5513–14 (Feb. 3, 2003), *available at* http://www.whitehouse.gov/omb/fedreg/2003draft_cost-benefit_rpt.pdf.

different policies depends in part on how policy makers wish to use them in decision making. At one extreme, information markets can be used simply to generate information on the incidence of different policies without affecting the actual decision. At the other extreme, information markets could be used to block policies that have adverse distributional consequences, or at least require that those policies only be implemented when certain redistributive objectives are met.[197]

Second, research is needed to determine how such markets can inform important policy issues. For example, information markets could shed light on the nature of biases that may be introduced through ex ante analysis that does not take account of market data.[198] Similarly, one could attempt to identify patterns in prices and hence, net benefits, as information markets evolve over time. This analysis could reveal, for example, whether net benefits are likely to be overstated or understated when a policy is first proposed.

Researchers could also use information markets to compare different measures of net benefits.[199] Lawrence Goulder and Roberton Williams argue that in certain settings it is possible to estimate the difference in benefits resulting from partial and general equilibrium analysis.[200] In principle it would be possi-

---

197. For a novel approach to promoting equity, see SHILLER, *supra* note 82, Part 3.

198. *See, e.g.*, Winston Harrington et. al., *On the Accuracy of Regulatory Cost Estimates*, 19 J. POL'Y ANALYSIS & MGMT. 297 (2000) (examining trends in the biases of regulatory cost estimates). One of the problems with existing studies is that they often compare ex ante and ex post point estimates or make arbitrary assumptions about whether there is a significant difference between the estimates. Information markets can provide information on point estimates and variation that allow formal tests of significance. Another problem with many comparisons of ex ante and ex post point estimates is that they cannot easily address differences in the quality of the studies. Information markets could help address this deficiency by providing a common framework for making the comparison that is specifically aimed at aggregating diverse sources of private information. Finally, information markets could also eliminate problems that arise when different studies use different benchmarks for comparison.

199. In practice, all of the measures are likely to be imperfect for a variety of reasons. To take just one, the cost number used above only includes direct costs. It is well known that indirect costs and benefits could play an important role in certain policy arenas. *See, e.g.*, Michael Hazilla & Raymond J. Kopp, *Social Cost of Environmental Quality Regulations: A General Equilibrium Analysis*, 98 J. POL. ECON. 853 (1990).

200. *See* Lawrence H. Goulder & Roberton C. Williams III, *The Substantial Bias from Ignoring General Equilibrium Effects in Estimating Excess Burden, and a Practical Solution*, 111 J. POL. ECON. 898 (2003) (showing that under typical conditions the simple "excess-burden triangle" formula substantially underestimates the excess

ble to compare partial and general measures of welfare based on information markets. For example, one could compare the impact of the education reform project identified earlier in terms of its impact on GDP versus a narrower definition of net benefits. The difference in the prices of the contracts with payoffs dependent upon each measure of benefits would reveal the differences in the measures.

Finally, academics could study the potential impact and value of information from different sources, including information markets and conventional cost-benefit analyses done by experts.[201] The choice is not necessarily between using information markets or a trained analyst to provide most or all of the information on a cost-benefit analysis. Both sources of information could turn out to be useful. There is also an issue of how many information markets are necessary. For example, we argue that an information market that assesses benefits could be helpful in providing assessments of net benefits, and would possibly even raise revenues for the government, but that proposition needs to be examined more carefully.

The optimum utilization of different resources will likely vary with the situation. Where one is reasonably confident in the information provided by information markets, one could substitute markets for experts doing detailed cost-benefit analyses. Where one is less confident, one may want to use such experts or use them in tandem with information markets.

VI. CONCLUSION

This Article shows how a new decision framework for policy making can improve upon existing analytical tools, such as cost-benefit analysis. We explain how application of this framework can have important consequences for how government officials make public policy decisions. Using information markets in conjunction with pay-for-performance contracts can enhance both legislative and regulatory oversight activities. A

---

burden of commodity taxes, in some cases by a factor of 10 or more. This formula performs poorly because it ignores general equilibrium interactions, including interactions between the taxed commodity and the labor market.)

201. This could involve some combination of controlled experiments as well as natural experiments, as governments gain more experience with information markets.

key feature of our framework is that it would hold both lawmakers and regulators more accountable for their policies.

One might ask why a mechanism for implementing projects like the one proposed herein has not been designed or implemented already. The answer is likely that information markets are still relatively new and scholars and practitioners are just beginning to appreciate their potential. Further, there are strong forces that have an interest in maintaining the status quo, and the performance-based-policy paradigm could dramatically change the political landscape.

Our main contribution is to propose an efficient way to implement well informed policy decisions. We do this by linking and building upon the literatures on information markets and mechanism design. The prices in information markets can inform the mechanism design process, thereby making previously infeasible mechanisms feasible for the policy maker.

Information markets allow policy makers to estimate key quantities in the counterfactual world in which their policy is *not* implemented. This is an important advance. Among other things, it enables policy makers to assess the benefits and net benefits of a policy before it is actually implemented. This, in turn, allows the design of true pay-for-performance contracts in a wide array of policy domains.

Although we focus on public-sector decision making, the analysis is sufficiently general to apply to a wide range of problems in private-sector and not-for-profit decision making. The framework can be applied to any situation in which a decision maker has the resources, but not the necessary information and ability, to achieve her specified objective. There are a wide array of possible applications, including regulation, trade, taxation, and social policy. The private sector could use information markets to evaluate new investments, aid forecasting, improve decision making, and enhance risk-sharing opportunities. The not-for-profit sector could use these markets to evaluate new investments and inform decisions about how to design performance-based contracts.

To sum up the key points of our proposal: First, we show how it is generally possible to design contracts based on different contingencies whose prices will convey useful information on the costs and benefits of a number of policy choices, ranging from regulation to public works projects.

Second, we identify the strengths and limitations of using information markets to help improve cost-benefit analyses. With these markets, decision makers might do a much better job assessing critical public policy issues.

Third, we show how information markets can be used to provide a stronger foundation for implementing a variety of government oversight mechanisms that, up to this point, have been stymied because of difficulties in estimating costs and benefits. We consider two such mechanisms, a regulatory budget and the use of net benefit accounts, and show how our proposed mechanism can help make these concepts workable. We also show how legislators can use traditional budgetary controls in conjunction with information markets to exercise effective oversight and implement performance-based policies.

Fourth, we argue that information markets could promote greater transparency in governmental decision making, provide more accurate estimates of the efficiency and distributional impacts of different policies, provide a better understanding of uncertainties, help with sensitivity analysis, offer a low-cost way of assessing new policy proposals, finance government projects and regulations with positive net benefits, allow those affected by specific policies the opportunity to hedge risk, and aid in the design of policies. Information markets can also help assess the value of additional research on the decision to undertake a project. If these markets become more widely accepted as a way of informing policy decisions, they have the potential to change the political landscape and reduce rent seeking.

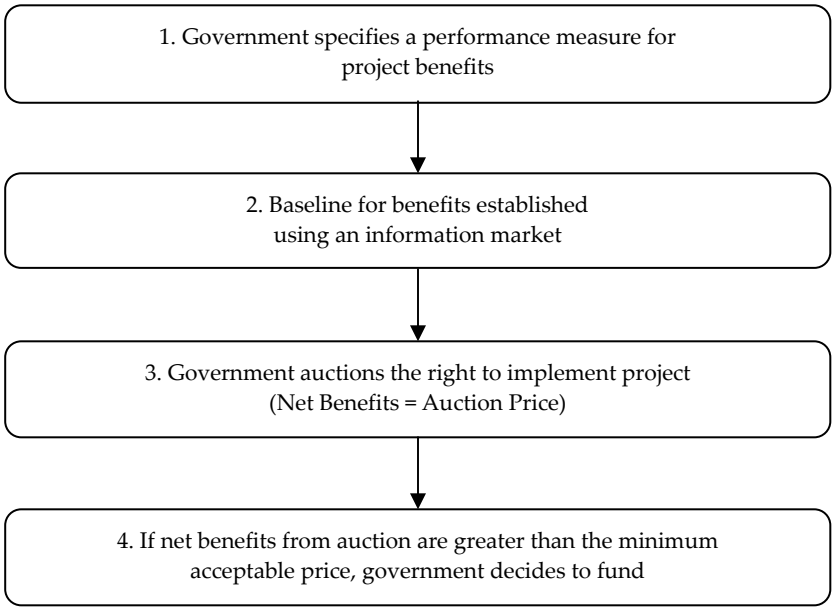
Fifth, we explain how government could play an important role in the expansion of information markets and how independent researchers could help in the development and assessment of these markets. Specifically, we recommend that government take a careful, iterative approach to facilitating the development of these markets.

Sixth, more research is needed to understand the theoretical potential of information markets as well as their practical relevance. The challenge for researchers will be to develop practical applications of these markets that can yield significant improvements in public policy. We think that goal is now within reach.

Finally, while we are enthusiastic about using information markets, they are only one of many tools available to policy makers for reaching decisions. The extent to which they can and should substitute for other tools should be examined carefully as their use in policy settings becomes more widespread. We would not advocate that information markets solely determine the outcome of key public policies at this point, but we would not rule out this possibility in the future.

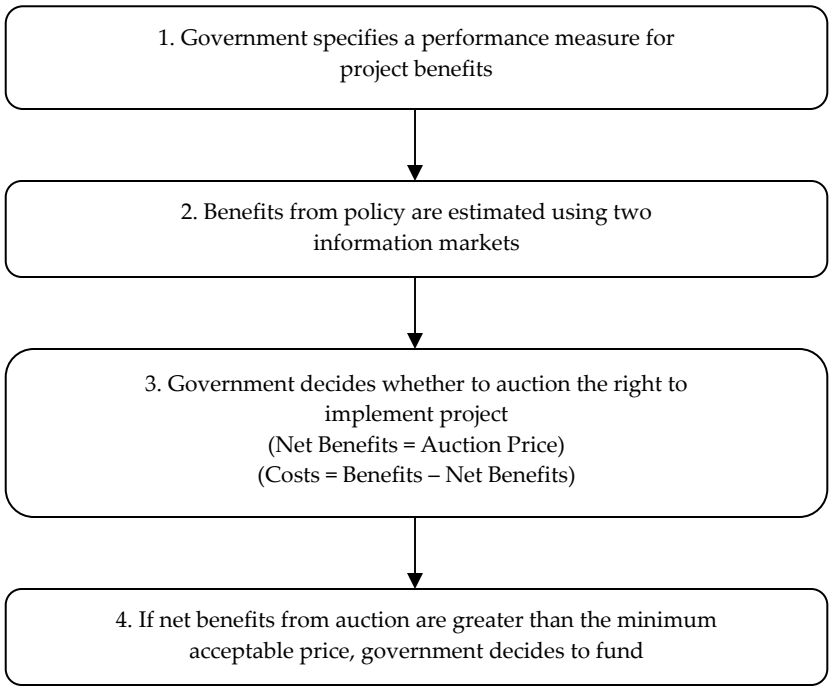
VII. APPENDIX

FIGURE 1.
Schematic for Project Selection



Note: The government must attach a monetary value to the performance measure. The baseline is the level of benefits expected in the absence of the project. In an auction with competitive bidding, the prevailing price will approximate expected net benefits from the project. The government's minimum acceptable price may or may not be publicly known to bidders. See text for additional details.

FIGURE 2.
Schematic for Project Selection with Early Information on Benefits



Note: The government must attach a monetary value to the performance measure. The estimate of project benefits equals benefits with policy minus benefits without policy. In an auction with competitive bidding, the prevailing price will approximate expected net benefits from the project. The government's minimum acceptable price may or may not be publicly known to bidders. See text for additional details.

TABLE 1.
Comparison of Direct and Indirect Cost-Benefit Methods Using Information Markets

Issue	**Direct Method**	**Indirect Method**
Net social benefits estimated	Yes, by market participants	Yes, market estimates hypothetical analyst's net benefit estimate
Measure policy equity	Yes (same as above)	Yes (same as above)
Price includes all information available to market participants	Yes, assuming markets are efficient	No, unless market participants anticipate the hypothetical analyst possesses all information
Information contained in prices clear	Yes	No, cannot distinguish between information about the analyst vs. net benefits
Adapt to new information	Yes, through issuing new contracts	Yes, based on hypothetical analyst
Potential for bias in prices	Low (possibility of inefficient prices)	Depends on expectations of the hypothetical analyst's identity and motives and possible inefficiencies in prices
Estimate the distribution of net benefits	Yes, using options or other derivatives	No, cannot distinguish between variance in prices caused by changes in beliefs about analyst and new information
Assess changes over time in information about net benefits	Yes, by following the path of price over time	No, price changes could be caused by changes in beliefs about the analyst
Assess the *social* value of new information	Yes, using options	No, can only assess the value to a hypothetical analyst
Can incorporate unverifiable benefits	Yes, the policy maker can explicitly employ different cutoffs in his decision rule	Yes, the hypothetical future analyst can employ any arbitrary decision rule
Contracts can serve as a hedging vehicle	Yes, perfect hedge for project implementer and others affected by the project	Yes, but imperfect hedge because of price fluctuations from changes in beliefs about the analyst

AGAINST FOREIGN LAW

ROBERT J. DELAHUNTY* AND JOHN YOO**

In recent years, several Supreme Court Justices have looked to the decisions of foreign and international courts for guidance in interpreting the U.S. Constitution. This practice has occurred in several controversial, high-profile cases. *Roper v. Simmons*[1] outlawed application of the death penalty to offenders who were under eighteen when their crimes were committed. *Lawrence v. Texas*[2] struck down a state law that criminalized homosexual sodomy. *Atkins v. Virginia*[3] held against the execution of mentally retarded capital defendants. All three cite foreign and international precedents.

In *Roper*, the Court, per Justice Kennedy, found it "proper that we acknowledge the overwhelming weight of international opinion against the juvenile death penalty . . . . The opinion of the world community, while not controlling our outcome, does provide respected and significant confirmation for our own conclusions."[4] The Court relied on a provision of the United Nations Convention on the Rights of the Child—a treaty the United States has not ratified—and on amicus briefs by the European Union and interested foreign observers.[5] In *Lawrence*, Justice Kennedy's majority opinion cited decisions of the European Court of Human Rights to conclude that prohibiting homosexual sodomy is at odds with the current norms of

---

* Associate Professor of Law, University of St. Thomas School of Law.

** Professor of Law, University of California, Berkeley School of Law (Boalt Hall); Visiting Scholar, American Enterprise Institute.

We received valuable comments on earlier versions of this paper at a faculty workshop at the William & Mary School of Law and at the Federalist Society's 2004 National Symposium. Jesse Choper and John McGinnis provided helpful comments.

1. 125 S. Ct. 1183 (2005) (plurality opinion).

2. 539 U.S. 558 (2003).

3. 536 U.S. 304 (2002).

4. *Roper*, 125 S. Ct. at 1200.

5. *Id.* at 1199 (citing United Nations Convention on the Rights of the Child art. 37, Nov. 20, 1989, 1577 U.N.T.S. 3, 28 I.L.M. 1448, 1468–70).

western civilization.[6] In *Atkins*, the majority opinion by Justice Stevens relied on an amicus brief filed by the European Union to assert that executing the mentally retarded is "overwhelmingly disapproved."[7] References to foreign decisions have appeared not just in cases expanding individual rights,[8] but also in dissents from federalism opinions.[9]

It is difficult to know how seriously to take this development. One possibility is that Justices cite foreign decisions merely as ornaments.[10] Under this theory, while providing aesthetic support, the practice of citing to international law may contribute little of analytical value and have no real effect on the actual course of judicial decisionmaking. In other words, even had the foreign decisions come out differently, the Supreme Court would have still reached its preferred result by the same primary mode of reasoning. In this situation, foreign precedents provide only tangential authority to bolster decisions reached on more traditional grounds.

There are, however, two main reasons to think that use of foreign or international decisions extends beyond mere ornamentation, one such reason grounded in the expressed views of several Justices and the other in academic opinion.

First, several Justices—most prominently Justice Breyer—have been explicit proponents of the use of foreign law. In a speech before the American Society of International Law, Justice Breyer enthusiastically claimed that a "global legal enter-

6. *Lawrence*, 539 U.S. at 572–73 (citing Dudgeon v. United Kingdom, 45 Eur. Ct. H.R. ¶ 52 (1981)).

7. *Atkins*, 536 U.S. at 316 n.21.

8. Other Justices have also relied on international agreements for support. *See, e.g.*, Grutter v. Bollinger, 539 U.S. 306, 344 (Ginsburg, J., joined by Breyer, J., concurring) ("The International Convention on the Elimination of All Forms of Racial Discrimination, ratified by the United States in 1994, see State Dept., Treaties in Force 422–423 (June 1996), endorses 'special and concrete measures to ensure the adequate development and protection of certain racial groups . . . .'").

9. *See* Printz v. United States, 521 U.S. 898, 976 (1997) (Breyer, J., dissenting) ("At least some other countries, facing the same basic problem, have found that local control is better maintained through application of a principle that is the direct opposite of the principle of the majority derives from the silence of our Constitution.").

10. Thus, Professor Alford suggests that the Court's use of international and foreign materials in constitutional cases may be mere "*bricolage*." *See* Roger P. Alford, *Misusing International Sources to Interpret the Constitution*, 98 AM. J. INT'L L. 57, 64–65 (2004).

prise . . . is now upon us."[11] His remarks built upon arguments he had previously made in judicial opinions.[12] Justices Stevens,[13] O'Connor,[14] and Ginsburg[15] have also advocated the use of foreign law in constitutional adjudication. In such cases, Justice Breyer argues that the "judgment of other nations" should help "guide this Court" in Eighth Amendment cases.[16] Justice O'Connor has gone even further in claiming that "we should not be surprised to find congruence between domestic and international values" in the Eighth Amendment context.[17]

---

11. Stephen Breyer, Assoc. Justice, Supreme Court of the U.S., The Supreme Court and The New International Law, remarks to the American Society of International Law (Apr. 4, 2003), http://www.supremecourtus.gov/publicinfo/speeches/sp_04-04-03.html.

12. *See, e.g.*, Nixon v. Shrink Mo. Gov't PAC, 528 U.S. 377, 403 (2000) (Breyer, J., concurring) (citing European and Canadian judicial decisions for their treatment of First Amendment issues); Foster v. Florida, 537 U.S. 990, 993 (2002) (Breyer, J., dissenting from denial of cert.) ("'[A]ttention to the judgment of other nations' . . . can help guide this Court when it decides whether a particular punishment violates the Eighth Amendment" (quoting THE FEDERALIST NO. 63 (James Madison))); Knight v. Florida, 528 U.S. 990, 995–97 (1999) (Breyer, J., dissenting from denial of cert.) (finding the opinions of former British Commonwealth nations "particularly instructive" in constitutional death penalty adjudication).

13. *See* Thompson v. Oklahoma, 487 U.S. 815, 830–31 (1988) (plurality opinion).

14. *See, e.g.*, Roper v. Simmons, 125 S. Ct. 1183, 1215–16 (2005) (O'Connor, J., dissenting) ("I disagree with Justice Scalia's contention . . . that foreign and international law have no place in our Eighth Amendment jurisprudence. Over the course of nearly half a century, the Court has consistently referred to foreign and international law as relevant to its assessment of evolving standards of decency . . . . [T]his Nation's evolving understanding of human dignity certainly is neither wholly isolated from, nor inherently at odds with, the values prevailing in other countries. On the contrary, we should not be surprised to find congruence between domestic and international values . . . ."); Sandra Day O'Connor, Keynote Address at the 96th Annual Meeting of the American Society of International Law (March 13–16, 2002), *in* 96 AM. SOC. INT'L L. PROC. 348, 350 (2002) ("Although international law and the law of other nations are rarely binding upon our decisions in U.S. courts, conclusions reached by other countries and by the international community should at times constitute persuasive authority in American courts."); Sandra Day O'Connor, *Broadening Our Horizons: Why American Lawyers Must Learn About Foreign Law*, FED. LAW., Sept. 1998, at 20.

15. *See, e.g.*, Grutter v. Bollinger, 539 U.S. 306, 344 (2003) (Ginsburg, J., concurring) (relying on international treatment of affirmative action); Ruth Bader Ginsburg, *Looking Beyond Our Borders: The Value of a Comparative Perspective in Constitutional Adjudication*, 22 YALE L. & POL'Y REV. 329 (2004); Ruth Bader Ginsburg & Deborah Jones Merritt, *Affirmative Action: An International Human Rights Dialogue*, 21 CARDOZO L. REV. 253, 282 (1999).

16. *Foster*, 537 U.S. at 993 (Breyer, J., dissenting from denial of cert.).

17. *Roper*, 125 S. Ct. at 1216 (O'Connor, J., dissenting).

Second, legal academics have urged the Supreme Court to engage in a "dialogue" with their foreign counterparts. Four academic projects are particularly noteworthy: Professor Bruce Ackerman has advocated "world constitutionalism";[18] Professors Vicki Jackson and Mark Tushnet have become interested in the possibilities of comparative constitutional analysis;[19] Professor Harold Koh has argued that the Court should look beyond American law when interpreting a constitutional term (like unreasonable search or due process) that "implicitly refers to a community standard";[20] and international law scholar Anne-Marie Slaughter has argued in favor of transnational communication between courts.[21] These academics appear to be attempting to construct an intellectual framework that could justify more extensive use of foreign judicial decisions in the future. This may presage further federal judicial reliance on foreign decisions for support.

This trend has not utterly swept the field, however. Several Justices have expressed hostility to the use of foreign or international decisions in interpreting the U.S. Constitution. In *Atkins*, Chief Justice Rehnquist declared, "I fail to see, however, how the views of other countries regarding the punishment of their citizens provide any support for the Court's ultimate determination."[22] After noting that such an approach had been rejected in earlier Eighth Amendment cases,[23] the Chief Justice argued that "[f]or if it is evidence of a *national* consensus for

---

18. *See* Bruce Ackerman, *The Rise of World Constitutionalism*, 83 VA. L. REV. 771 (1997); *see also* David S. Law, *Generic Constitutional Law*, 89 MINN. L. REV. 652 (2005).

19. *See Vicki C. Jackson, Ambivalent Resistance and Comparative Constitutionalism: Opening up the Conversation on "Proportionality," Rights and Federalism*, 1 U. PA. J. CONST. L. 583 (1999); Mark Tushnet, *The Possibilities of Comparative Constitutional Law*, 108 YALE L.J. 1225 (1999).

20. Harold H. Koh, *International Law as Part of Our Law,* 98 AM. J. INT'L L. 43, 46 (2004).

21. *See* Anne-Marie Slaughter, *A Typology of Transnational Communication*, 29 U. RICH. L. REV. 99 (1994); Anne-Marie Slaughter, *Judicial Globalization*, 40 VA. J. INT'L L. 1103 (2000). In addition to the scholars cited, other scholars, even while critical of aspects of the Court's use of foreign and international law sources, believe that those sources may be relevant to *some* questions of constitutional interpretation. *See, e.g.*, Michael Ramsey, *International Materials and Domestic Rights: Reflections on* Atkins *and* Lawrence, 98 AM. J. INT'L L. 69, 71 (2004) ("International sources are obviously relevant to the scope of the Constitution's structural provisions defining the international powers of the U.S. government.").

22. Atkins v. Virginia, 536 U.S. 304, 324–25 (2002) (Rehnquist, C.J., dissenting).

23. *Id.* at 325 (citing Stanford v. Kentucky, 492 U.S. 361, 369 (1989) (plurality opinion)).

which we are looking, then the viewpoints of other countries simply are not relevant."[24] Justices Scalia and Thomas have also argued that foreign precedents are irrelevant for constitutional interpretation because those decisions interpret other documents.[25] Most recently, Justice Scalia's *Roper* dissent contended that "the basic premise of the Court's argument—that American law should conform to the laws of the rest of the world—ought to be rejected out of hand."[26] Arguing that the Court's use of foreign law was inconsistent and unprincipled, as shown by the deviation between American and foreign case law on the exclusionary rule, church-state relations, and abortion, Justice Scalia concluded that "[t]he Court should either profess its willingness to reconsider all these matters in light of the views of foreigners, or else it should cease putting forth foreigners' views as part of the *reasoned basis* of its decisions. To invoke alien law when it agrees with one's own thinking, and ignore it otherwise, is not reasoned decisionmaking, but sophistry."[27]

It is essential to identify precisely what is and what is not in controversy here. Some Justices appear to believe that foreign and international legal practices and opinions can serve, at a minimum, to illuminate possible solutions to questions similar to those that U.S. courts must address, just as our federal courts may learn from our state courts, and vice versa. Foreign materials can provide relevant empirical information about the practical effects of particular social policies. Or, like law review articles, they can furnish original legal arguments. To the extent that foreign and international legal materials are used only for those purposes, such use is unobjectionable (assuming it is methodologically sound). However, some Justices may also think it desirable for U.S. constitutional law to converge with the constitutional law of European and other democratic legal systems. The desire to promote such convergence may lead

---

24. *Id.*

25. *See, e.g.*, Lawrence v. Texas, 539 U.S. 558, 598 (Scalia, J., dissenting) ("The Court's discussion of these foreign views . . . is therefore meaningless dicta."); Foster v. Florida, 537 U.S. 990, 990 n.* (Thomas, J., concurring in denial of cert.).

26. Roper v. Simmons, 125 S. Ct. 1183, 1226 (2005) (Scalia, J., dissenting).

27. *Id.* at 1228. Just two months before the *Roper* decision, Justices Breyer and Scalia debated the role of international law. The Relevance of Foreign Law for American Constitutional Adjudication, January 13, 2005, http://domino.american.eduAU/media/mediarel.nsf (follow "AU Media Press Releases" hyperlink; then locate the releases issued on January 14, 2005.)

those Justices to give *decisional* effect to foreign materials, thereby allowing them to be outcome-determinative in constitutional cases. So used, foreign or international law would have legal force in deciding important questions, including the rights of criminal defendants, the constitutionality of parental notification requirements for abortions, the reasonableness of searches by law enforcement, the extent of governmental leeway in religion cases, and the validity of various forms of capital punishment under the Eighth Amendment. It is this type of use of foreign law with which this Article will take issue.[28]

Foreign and international law cannot be legitimately used in an outcome-determinative way to decide questions of constitutional interpretation. A narrow exception embraces those cases where the text of the Constitution itself refers to international or foreign law, as it does, for example, when it vests in Congress the powers to "define and punish . . . Offences against the Law of Nations"[29] and to "declare War."[30] Under those clauses, the Constitution gives Congress the authority to promulgate rules of international law. As discussed further below, the Supreme Court held in the early days of the Republic that public international law (or more precisely, the Law of Nations) could be used to interpret the scope of Congress' power to "declare War."[31] That holding makes perfect sense, given that the power to "declare War" is correctly seen not as a matter of domestic separation-of-powers law enabling Congress to check the President's authority to engage in armed conflict, but rather as an exceptional grant of authority to Congress to make legal rules in the *international* sphere.[32] Likewise, the scope of Congress's power to define and punish offenses against the "Law of Nations" can be legitimately established by reference to that Law.[33] Such unusual cases aside, foreign and international law

28. No reason exists to distinguish between terms that appear in the constitutional text (like the term "unreasonable" in the Fourth Amendment) and terms that are the common coin of the Court's constitutional case law (like "undue burden" in the context of abortion). If foreign law is irrelevant to interpreting constitutional *text*, it should not be relevant to constitutional *doctrine*.

29. U.S. CONST. art. I, § 8, cl. 10.

30. U.S. CONST. art. I, § 8, cl. 11.

31. *See* Brown v. United States, 12 U.S. (8 Cranch) 110, 125 (1814).

32. *See* 1 JAMES KENT, COMMENTARIES 56–58 (Da Capo Press 1971) (1826) (explaining the meaning of the Declare War Clause). *See also* John C. Yoo, *War and the Constitutional Text*, 69 U. CHI. L. REV. 1639 (2004).

33. The Constitution also refers to international law in certain grants of power to the federal courts: for example, when it authorizes those courts to try "all Cases of

should not be available as a basis for interpreting the Constitution.

This Article makes four observations regarding the Supreme Court's practice of relying upon foreign and international decisions to support of its constitutional rulings. Part I argues that non-ornamental use of foreign decisions undermines the separation of powers and violates the constitutional rules against delegation of federal authority to bodies outside the control of the national government. Part II argues that use of foreign decisions undermines the limited theory of judicial review, as set out in *Marbury v. Madison*.[34] Chief Justice Marshall justified the federal courts' power to ignore enacted laws that were inconsistent with the Constitution on the ground that such statutes fell outside the delegation of authority by the people to the government, as expressed in the Constitution.[35] Relying on decisions that interpret a wholly different document runs counter to the notion that judicial review derives from the Court's duty to enforce the Constitution. Part III considers the relevance of the Constitution's Supremacy Clause and Law of Nations Clause to the Court's emerging use of foreign law. Part IV questions the Court's use of precedents that derive almost exclusively from Europe. We will suggest that Europe does not present the ideal model of constitutionalism for the United States to follow and that in fact deviation between the United States and Europe may significantly enhance global welfare.

---

admiralty and maritime Jurisdiction." U.S. CONST. art. III, § 2, cl. 1. Maritime and admiralty law "has strong roots in international custom. In extending judicial power 'to all cases of admiralty and maritime jurisdiction,' the framers of the Constitution apparently had English admiralty practice in mind, itself based on the civil law of continental Europe." James A.R. Nafziger, *The Evolving Role of Admiralty Courts in Litigation Related to Historic Wreck*, 44 HARV. INT'L L. J. 251, 266 (2003). *See also* WILLIAM BLACKSTONE, 4 COMMENTARIES ON THE LAWS OF ENGLAND at *67 (stating that the Law of Nations is "universal law" that encompasses "mercantile questions, such as bills of exchange and the like," "all marine causes," "the law-merchant," and "disputes relating to prizes, to shipwrecks, to hostages, and ransom bills."); Edward Dumbauld, *John Marshall and the Law of Nations*, 104 U. PA. L. REV. 38, 39 (1955) (quoting the statement of Lord Mansfield that "the maritime law is not the law of a particular country, but the general law of nations"). *See infra* text accompanying notes 120–22.

34. 5 U.S. (1 Cranch) 137 (1803).

35. *Id.* at 176 ("That the people have an original right to establish, for their future government, such principles as, in their opinion, shall most conduce to their own happiness, is the basis, on which the whole American fabric has been erected.").

I. FOREIGN LAW AND U.S. SEPARATION OF POWERS

As noted above, in *Roper*, *Atkins*, and *Lawrence*, the Court used foreign precedents to analyze the application of the Eighth Amendment and the Due Process Clause. All three cases involved measuring state action—execution of juveniles and the mentally retarded and the criminalization of homosexual sodomy—against social norms. To determine the content of such norms, the Court looked to European precedents as indicative of world opinion on the question.[36]

The Court's use of foreign law has the potential to turn into a standard of deference. There is support for such a development within the legal academy. Scholars have argued in favor of allowing customary international law to enjoy the status of federal common law,[37] and certain lower court interpretations of the Alien Tort Statute[38] threaten to do just that.[39] Other scholars have argued that American courts should defer to the decisions of foreign courts in interpreting treaties.[40]

---

36. Roper v. Simmons, 125 S. Ct. 1183, 1199 (2005) ("The United Kingdom's experience bears particular relevance here . . . . As of now, the United Kingdom has abolished the death penalty in its entirety; but, decades before it took this step, it recognized the disproportionate nature of the juvenile death penalty; and it abolished that penalty as a separate matter."); Lawrence v. Texas, 539 U.S. 558, 572 (2003) ("The sweeping references by Chief Justice Burger to the history of Western civilization and to Judeo-Christian moral and ethical standards did not take account of other authorities pointing in an opposite direction."); Atkins v. Virginia, 536 U.S. 304, 317 n.21 (2002) ("[W]ithin the world community, the imposition of the death penalty for crimes committed by mentally retarded offenders is overwhelmingly disapproved.").

37. *See, e.g.*, Harold H. Koh, *Is International Law Really State Law?*, 111 HARV. L. REV. 1824 (1998).

38. 1 Stat. 73 (1789) (codified at 28 U.S.C. § 1350 (2000)) (providing that federal district courts "shall have original jurisdiction of any civil action by an alien for a tort only, committed in violation of the law of nations or a treaty of the United States").

39. *See* Filartiga v. Pena-Irala, 630 F.2d 876 (2d Cir. 1980). *See also* Curtis A. Bradley, *The Alien Tort Statute and Article III*, 42 VA. J. INT'L L. 587 (2002) (arguing that the "law of nations" provision of the Alien Tort Statute was not originally understood as "creating a federal statutory cause of action for violations of the law of nations"); Julian Ku & John Yoo, *Foreign Competence: Formalism, Functionalism, and the Alien Tort Statute*, 2004 SUP. CT. REV. 153 (arguing that courts lack functional competence to promote foreign policy goals absent direct guidance from the political branches).

40. *See, e.g.*, Michael P. Van Alstine, *The Judicial Power and Treaty Delegation*, 90 CAL. L. REV. 1263, 1293 (2002) (responding to John Yoo, *Politics as Law?: The Anti-Ballistic Missile Treaty, the Separation of Powers, and Treaty Interpretation*, 89 CAL. L. REV. 851 (2001)); Michael P. Van Alstine, *Dynamic Treaty Interpretation*, 146 U. PA. L. REV. 687, 787–91 (1998).

Though the arguments for deferring to foreign precedents are not precise, analogy is possible to the standards of deference that courts apply to administrative agencies. The strongest of these standards, *Chevron* deference, requires courts to defer to agency interpretations of an ambiguous statutory provision if Congress's intent does not clearly dictate otherwise and if the interpretation is a permissible or not unreasonable reading of the provision.[41] "Deference under *Chevron* to an agency's construction of a statute that it administers is premised on the theory that a statute's ambiguity constitutes an implicit delegation from Congress to the agency to fill in the statutory gaps."[42] Under a weaker standard of deference, *Skidmore* deference, the weight given to an agency interpretation rests "upon the thoroughness evident in its consideration, the validity of its reasoning, its consistency with earlier and later pronouncements, and all those factors which give it power to persuade, if lacking power to control."[43] One of the authors has elsewhere criticized offering either form of deference to foreign law as a problematic delegation of authority from the federal government to foreign courts.[44]

That sort of deference to foreign decisions runs counter to the constitutional structure. It would subject American citizens to the judgments of foreign and international courts, and the Constitution makes no provision for the transfer of federal power to entities outside of our system of government. To the contrary, the Appointments Clause directly limits the transfer of federal power.[45] Much writing on this clause has focused on

---

41. Chevron U.S.A., Inc. v. Natural Res. Def. Council, Inc., 467 U.S. 837, 842–45 (1984). For a classic exposition of *Chevron*, see Laurence H. Silberman, Chevron—*The Intersection of Law & Policy*, 58 GEO. WASH. L. REV. 821 (1990). For a more recent exposition, see Jonathan Molot, *Reexamining* Marbury *in the Administrative State: A Structural and Institutional Defense of Judicial Power over Statutory Interpretation*, 96 NW. U. L. REV. 1239 (2002).

42. FDA v. Brown & Williamson Tobacco Corp., 529 U.S. 120, 159 (2000). On this point, see John Manning, *The Nondelegation Doctrine as a Canon of Avoidance*, 2000 SUP. CT. REV. 223.

43. Skidmore v. Swift & Co., 323 U.S. 134, 140 (1944); *see also* United States v. Mead Corp., 533 U.S. 218, 227, 234 (2001) (distinguishing *Chevron* deference from *Skidmore* deference).

44. *See* John C. Yoo, *Treaty Interpretation and the False Sirens of Delegation*, 90 CAL. L. REV. 1305 (2002).

45. John C. Yoo, *The New Sovereignty and the Old Constitution: The Chemical Weapons Convention and the Appointments Clause*, 15 CONST. COMMENT. 87, 96 (1998) ("[T]he Constitution erects limits on the ability of the federal government to

the balance of power between the President and Senate in the appointment of federal judges.[46] However, the Clause also functions as a mechanism to conserve federal power. In recent decisions, the Supreme Court has recognized that the Appointments Clause restricts the exercise of federal power to those officials appointed through the processes set out in the Clause.[47] This restriction ensures that federal power is limited to officials who are accountable solely to elected representatives, and thus ultimately to the American people. Plainly, international and foreign courts do not meet this standard.

The Court's discussions of the Appointments Clause in *Edmond v. United States*[48] and *Printz v. United States*[49] affirm the notion that Congress may not transfer responsibility for the execution of federal law to officers outside the control of the executive branch. In *Edmond*, the Court observed that the Appointments Clause "is among the significant structural safeguards of the constitutional scheme."[50] In *Printz*, the Court held that Congress could not delegate the power to enforce the Brady Act to state officials because such delegation would leave federal law enforcement free of "meaningful Presidential control" and would undermine the effectiveness of a unitary executive.[51] "That unity would be shattered, and the 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."[52] *Printz* made clear that the Appointments Clause would be offended not only if Congress sought to transfer federal law enforcement to officers of its own selection, but also if it attempted to delegate that

---

transfer or delegate power to entities that are not directly responsible to the American people.").

46. *See* John C. Yoo, *Choosing Justices: A Political Appointments Process and the Wages of Judicial Supremacy*, 98 MICH. L. REV. 1436, 1437 n.4 (2000) (collecting sources).

47. *See, e.g.*, Edmond v. United States, 520 U.S. 651, 663 (1997); Ryder v. United States, 515 U.S. 177, 180–84 (1995); Weiss v. United States, 510 U.S. 163, 169–76 (1994); Freytag v. Comm'r, 501 U.S. 868, 884 (1991); Buckley v. Valeo, 424 U.S. 1, 135 (1976) (per curiam).

48. 520 U.S. at 659.

49. 521 U.S. 898, 922–23 (1997).

50. 520 U.S. at 659.

51. 521 U.S. at 922.

52. *Id*. at 923.

power to officials outside the executive branch of the federal government.[53]

The Appointments Clause also concerns itself with the general scope and execution of national power. The Clause's requirement that all individuals who exercise significant federal authority become Officers of the United States, appointed pursuant to Article II, Section 2, ensures that the federal government cannot blur the lines of accountability between the people and their officials. As Chief Justice Rehnquist wrote for the Court in *Ryder v. United States*,[54] "The Clause is a bulwark against one branch aggrandizing its power at the expense of another branch, but it is more: it 'preserves another aspect of the Constitution's structural integrity by preventing the diffusion of the appointment power.'"[55] *Buckley v. Valeo*[56] first made clear the link between the Appointments Clause and the exercise of federal power. The *Buckley* Court rejected the proposition that Congress could appoint individuals to exercise federal power who were not Officers of the United States, observing, "We think . . . [that the Appointments Clause's] fair import is that any appointee exercising significant authority pursuant to the laws of the United States is an 'Officer of the United States,' and must, therefore, be appointed in the manner prescribed by [the Clause.]"[57] Individuals appointed by Congress, therefore, did not qualify as Officers of the United States and could only perform duties not involving the enforcement of federal law.

Two other elements of the Constitution's text and structure confirm the Appointments Clause's careful husbanding of federal power. First, Article III vests the federal judicial power "in one supreme Court, and in such inferior Courts as the Congress may from time to time ordain and establish."[58] This provision suggests that the federal judicial power, which includes the authority to decide cases or controversies under federal law, cannot be exercised by any other branch of the federal government, with the narrow and debatable exception of the

53. *Id.*
54. 515 U.S. 177 (1995).
55. *Id.* at 182 (quoting Freytag v. Comm'r, 501 U.S. 868, 878 (1991)).
56. 424 U.S. 1 (1996) (per curiam).
57. *Id.* at 126 (per curiam).
58. U.S. CONST. art. III, § 1.

Senate's role in trying cases of impeachment.[59] The logical implication is that no part of the Article III authority to decide federal cases and controversies, from which springs the judicial power to interpret the Constitution, can be delegated or transferred outside the United States government.[60]

Of course, Congress could have declined to create any lower federal courts. Furthermore, restrictions on the subject matter jurisdiction of federal courts cause many federal constitutional issues to arise first in state courts, whose judges are not members of the federal government.[61] Any resulting damage to the separation of powers is not, however, insurmountable. State judicial decisions can be reviewed by the Supreme Court, and state courts are still part of the American political system with judges who take an oath to uphold the Constitution. The Supremacy Clause specifically permits, even requires, the "commandeering" of state courts to enforce federal law. The potential violation of separation of powers is, therefore, greater when the courts defer to foreign laws or courts. This transfer of judicial power ignores the vesting of all judicial power in the Supreme Court and undermines the accountability of government. Members of the electorate cannot hold accountable officials who stand completely outside the structure of American government.

The second element of the Constitution's text and structure that confirms the Appointments Clause's conservation of power is the nondelegation doctrine, which prohibits Congress from delegating rulemaking authority to another branch unless

---

59. *See* Akhil Reed Amar, *A Neo-Federalist View of Article III: Separating the Two Tiers of Federal Jurisdiction*, 65 B.U. L. REV. 205 (1985); Steven G. Calabresi & Kevin H. Rhodes, *The Structural Constitution: Unitary Executive, Plural Judiciary*, 105 HARV. L. REV. 1153 (1992). Legislative attempts at jurisdiction-stripping, however, may obscure this concept. Commentators who believe that Congress has the authority to strip the federal courts of jurisdiction include Paul M. Bator, *Congressional Power over the Jurisdiction of the Federal Courts*, 27 VILL. L. REV. 1030, 1030–32 (1982); Gerald Gunther, *Congressional Power to Curtail Federal Court Jurisdiction: An Opinionated Guide to the Ongoing Debate*, 36 STAN. L. REV. 895, 898 (1984); and Henry M. Hart, Jr., *The Power of Congress to Limit the Jurisdiction of Federal Courts: An Exercise in Dialectic*, 66 HARV. L. REV. 1362, 1362–66 (1953).

60. *See* Julian G. Ku, *The Delegation of Federal Power to International Organizations: New Problems with Old Solutions*, 85 MINN. L. REV. 71 (2000).

61. The well pleaded complaint rule precludes almost all defendants in state courts from removing their cases to federal courts. Federal defenses to state law claims made by plaintiffs are therefore adjudicated almost exclusively by state judges. *See* Louisville & Nashville R.R. v. Mottley, 211 U.S. 149 (1908); 28 U.S.C. § 1331 (2001).

it has stated intelligible standards to guide administrative discretion.[62] This requirement ensures that the exercise of delegated power can be monitored and controlled, and even reversed when necessary. Although the Supreme Court has not invalidated a statute on nondelegation grounds since the New Deal period,[63] it remains an important structural principle that finds its expression in quasi-constitutional doctrines including canons of construction.[64] The delegation of lawmaking power outside the federal government would prevent lower courts, Congress, and the public from monitoring whether the delegated authority was being exercised consistent with legislative or constitutional standards.[65]

Providing any type of deference to foreign judicial decisions would cause considerable tension with these constitutional structures. Under the Appointments Clause, anyone who possesses the power to interpret and execute federal law must be an Officer of the United States. When the Court applies *Chevron* deference, or even the lesser form of *Skidmore* deference, it is providing that deference to officials who are appointed by the President or those responsible to him consistent with the Appointments Clause. Thus, those who make and interpret federal law, whether they are federal judges or federal agency officials, are still ultimately responsible to the American electorate. Foreign judges, in contrast, have been neither nominated by the President nor approved by the Senate, and thus should not exercise significant federal power by influencing the interpretation of a federal law. To the extent a foreign decision is outcome-determinative or triggers some type of deference, it would raise serious problems with the constitutional requirement that federal authority be exercised solely by federal officers.

---

62. *See* Whitman v. Am. Trucking Ass'ns, 531 U.S. 457 (2001); *cf.* Clinton v. City of New York, 524 U.S. 417 (1998).

63. *See* A.L.A. Schechter Poultry Corp. v. United States, 295 U.S. 495 (1935); Pan. Ref. Co. v. Ryan, 293 U.S. 388 (1935).

64. *See* Cass R. Sunstein, *Nondelegation Canons*, 67 U. CHI. L. REV. 315 (2000); Manning, *supra* note 42.

65. For recent discussion of the nondelegation doctrine, compare Eric A. Posner & Adrian Vermeule, *Interring the Nondelegation Doctrine*, 69 U. CHI. L. REV. 1721 (2002) with Larry Alexander & Saikrishna Prakash, *Reports of the Nondelegation Doctrine's Death Are Greatly Exaggerated*, 70 U. CHI. L. REV. 1297 (2003). *See also* Gary Lawson, *Delegation and Original Meaning*, 88 VA. L. REV. 327 (2002).

Reliance on foreign decisions also violates the policies animating the Appointments Clause and the nondelegation doctrine, namely, accountability and control. Delegation to federal agency officials seems tolerable because those officials are part of an executive branch responsible to the President, Congress, the courts, and the public. Though the Supreme Court may give it deference, an agency that goes well beyond its statutory mandate in interpreting or enforcing federal law is subject to checks, including congressional oversight, budget cuts, amended statutory requirements, presidential removal, public criticism and, ultimately, elections.[66] The same mechanisms do not constrain foreign judges, who are neither responsible to the American political system nor required to adapt their exercises of interpretive authority to federal constitutional or statutory principles. Reliance on foreign decisions would evade the Constitution's conferral of the power to implement and interpret federal law on officers of the United States who are accountable to the electorate.

II. FOREIGN LAW AND JUDICIAL REVIEW

Issues of delegation or deference aside, judicial reliance upon sources exogenous to the American political system in interpreting the Constitution undermines the textual and structural basis for judicial review. This criticism does not extend to all uses of foreign decisions by the federal courts. Such sources might be relevant to judicial interpretation of other types of federal law, such as treaties, statutes, and perhaps even federal common law. This Part argues, however, that when it comes to the Constitution, federal courts are limited to materials that derive from the American legal system.

The touchstone of this argument is the nature of judicial review. Some have argued that judicial review promotes certain functional goals, whether they be efforts to interpret the Constitution in light of contemporary values,[67] to protect minori-

66. Congressional attempts to shield itself from popular accountability have been struck down. *See* Metro. Wash. Airports Auth. v. Citizens for the Abatement of Aircraft Noise, 501 U.S. 252 (1991) (striking down a law that allowed a nine-member board of review made up of members of Congress to review the decisions of the airport authority).

67. *See, e.g.*, RONALD DWORKIN, FREEDOM'S LAW 2, 37 (1996) (articulating the view that judges should base judgments in part on public morality); ROBERT C. POST, CONSTITUTIONAL DOMAINS: DEMOCRACY, COMMUNITY, MANAGEMENT 36

ties,[68] to reinforce democratic representation and preserve space for democratic decisionmaking,[69] to promote the rule of law by interpretation along "common law" lines,[70] or to adhere solely to the Constitution's original understanding.[71] This Article offers a far more modest explanation of the origins of judicial review. Rather than deriving from any grand role of the judiciary within the constitutional system, judicial review originates in the nature of the Constitution as a document that delegates power from the people to the government, in the supremacy of constitutional law to statutory law, and in the duty of every federal officer to obey that higher law when confronted with the inconsistent actions of other branches of government.[72]

The structural foundation for judicial review lies in the nature of the Constitution and its relationship with the officers of the federal government. According to the theory of popular sovereignty prevalent at the time of ratification, the Constitution was a creation of the people of the several States.[73] Indeed,

---

(1995) (asserting that constitutional interpretation must be based on a common commitment to national values); LAURENCE H. TRIBE & MICHAEL C. DORF, ON READING THE CONSTITUTION 8–13 (1991) (arguing that each generation has a role in interpreting the Constitution).

68. *See, e.g.*, JOHN HART ELY, DEMOCRACY AND DISTRUST: A THEORY OF JUDICIAL REVIEW 87–88 (1980) (articulating a representation-reinforcing theory of judicial review); JESSE H. CHOPER, JUDICIAL REVIEW AND THE NATIONAL POLITICAL PROCESS: A FUNCTIONAL RECONSIDERATION OF THE ROLE OF THE SUPREME COURT (1980) (arguing that judicial review should protect minorities).

69. *See* CASS R. SUNSTEIN, ONE CASE AT A TIME: JUDICIAL MINIMALISM ON THE SUPREME COURT, at xiv (1999) (arguing that minimalist judicial decisions promote democratic deliberation).

70. *See, e.g.*, David A. Strauss, *Common Law Constitutional Interpretation*, 63 U. CHI. L. REV. 877, 879 (1996); HARRY H. WELLINGTON, INTERPRETING THE CONSTITUTION: THE SUPREME COURT AND THE PROCESS OF ADJUDICATION 78, 81–82 (1990).

71. *See, e.g.*, KEITH E. WHITTINGTON, CONSTITUTIONAL INTERPRETATION: TEXTUAL MEANING, ORIGINAL INTENT, AND JUDICIAL REVIEW 3 (1999); ANTONIN SCALIA, A MATTER OF INTERPRETATION 38–41 (1997); ROBERT H. BORK, THE TEMPTING OF AMERICA: THE POLITICAL SEDUCTION OF THE LAW 143 (1990).

72. *See* Saikrishna B. Prakash & John C. Yoo, *The Origins of Judicial Review*, 70 U. CHI. L. REV. 887 (2003).

73. *See* U.S. Term Limits, Inc. v. Thornton, 514 U.S. 779, 846 (1995) (Thomas, J., dissenting) (noting that the Constitution was ratified by the consent of the people of individual states and not by the consent of "the undifferentiated people of the Nation as a whole"). *See also* Gordon S. Wood, *The Origins of Judicial Review Revisited, or How the Marshall Court Made More out of Less*, 56 WASH. & LEE L. REV. 787, 792–94 (1999) (describing the development of the idea that "judges, though

"We the People of the United States" speak at the start of the Constitution in an authorial voice.[74] This understanding of government power indicates a rejection of the notion that sovereignty itself lodged in the government or monarch. Necessarily, the government exercises power only because it serves as the agent of the people's will. As James Madison wrote in *The Federalist No. 46*, "The federal and State Governments are in fact but different agents and trustees of the people, constituted with different powers and designed for different purposes."[75] Madison reminded critics of the proposed Constitution that "the ultimate authority, wherever the derivative may be found, resides in the people alone . . . ."[76]

It follows from this logic that the government may exercise only the power that the people have delegated to it. A written constitution serves to specify and limit those powers. Any exercise of authority beyond the grant of power in the written constitution is therefore invalid because it goes beyond the delegation from the people and undermines popular sovereignty. As Alexander Hamilton explained in *The Federalist No. 78*, "every act of a delegated authority, contrary to the tenor of the commission under which it is exercised, is void."[77] A written constitution would prove inconsequential if its agents could simply exercise the powers that they saw fit, regardless of the will of the people. As Chief Justice John Marshall declared in *Marbury v. Madison*, "The distinction between a government with limited and unlimited powers is abolished, if those limits do not confine the persons on whom they are imposed, and if acts prohibited and acts allowed, are of equal obligation."[78] For the Constitution successfully to establish written limitations on the powers of the branches of government, it must establish a rule of decision that places itself above the organs it creates.

---

not elected, resembled the legislators and executives in being agents or servants of the people with a responsibility equal to that of the other two branches of government"); Akhil Reed Amar, *Of Sovereignty and Federalism*, 96 YALE L.J. 1425, 1443 (1987) (describing the Federalists' identification of the executive and judiciary as "agents of the People" in curbing legislative power).

74. U.S. CONST. pmbl.

75. THE FEDERALIST NO. 46, at 294 (James Madison) (Clinton Rossiter ed., 1961).

76. *Id.*

77. THE FEDERALIST NO. 78 (Alexander Hamilton), *supra* note 75, at 467.

78. 5 U.S. (1 Cranch) 137, 176–77 (1803).

The task of policing the other federal and state governmental actors has largely fallen to the federal courts.[79] In *The Federalist No. 78*, Hamilton specifically emphasized the role of the federal courts as "an intermediate body between the people and the legislature in order, among other things, to keep the latter within the limits assigned to their authority."[80] Following Hamilton, Chief Justice Marshall in *Marbury* spelled out, through a series of *reductio ad absurdum* arguments, the necessity of such a judicial role. The view, he argued, that the courts could not examine the constitutionality of a statute before them, or pronounce it void if they found it repugnant to the Constitution, "would subvert the very foundation of all written constitutions."[81] Marshall explained:

> It would declare that an act which, according to the principles and theory of our government, is entirely void, is yet, in practice, completely obligatory. It would declare that if the legislature shall do what is expressly forbidden, such act, notwithstanding the express prohibition, is in reality effectual. It would be giving to the legislature a practical and real omnipotence, with the same breath which professes to restrict their powers within narrow limits."[82]

Neither *The Federalist* nor *Marbury*, however, makes the claim that it is *solely* the function of the judiciary to decide whether the acts of the other branches of government are unconstitutional, and hence ought not be obeyed.[83] Rather, popular sovereignty theory suggests that each branch has an obligation to refuse to obey commands that violate the Constitution. Judicial

---

79. This is in part because of the phenomenon of "agency costs," which can make it effectively impossible for a principal to monitor the performance of its agents and so ensure that they are not acting outside the scope of the authority delegated to them. In order to maintain control over its agents, a principal may therefore delegate to one of those agents a lead role in monitoring the actions of the others. To be sure, such a delegation may itself be abused, precisely because close monitoring of *that* agent may also be unduly costly for a principal.

80. THE FEDERALIST NO. 78 (Alexander Hamilton), *supra* note 75, at 467.

81. 5 U.S. (1 Cranch) at 178.

82. *Id.*

83. Indeed, Marshall later went to some length to specify a class of "political" cases in which the constitutionality of the actions of another branch is not open to judicial review. *See, e.g.*, Foster v. Neilson, 27 U.S. (2 Pet.) 253, 307–09 (1829) (international boundaries a political question); United States v. Palmer, 16 U.S. (3 Wheat.) 610, 634–35 (1818) (recognition of international status a political question); Rose v. Himely, 8 U.S. (4 Cranch) 241, 272 (1808) (independence of French colony a political question).

review arises from the principle that each branch of government is coordinate, independent, and responsible for interpreting and enforcing the Constitution when fulfilling its particular constitutional role. While the federal judiciary enjoys no constitutional authority to force the other branches to adopt its interpretations of the Constitution in the performance of their unique functions, neither can the other branches dictate constitutional meaning to the judiciary when it decides cases or controversies.[84] The Constitution's separation of powers necessarily requires judicial review.

There has been much debate between those who interpret separation of powers formally and those who interpret it functionally.[85] Nevertheless, both sides should concede several points that follow. The Constitution makes clear that the three branches are coordinate, in the sense that they are equal to each other. As James Madison wrote in *The Federalist* No. 49, "The several departments being perfectly co-ordinate by the terms of their common commission, neither of them, it is evident, can pretend to an exclusive or superior right of settling the boundaries between their respective powers . . . ."[86] Each branch exercises authority granted directly by the people through the Constitution, and no branch is subordinate to the others. In addition to being coordinate, the branches are separate. While some powers are shared, such those regarding treaties and appointments, each branch executes certain core functions that belong to it alone. Only Congress can enact legislation within the field of authority vested in it by Article I, Section 8 and the Reconstruction Amendments; only the President may execute federal laws; and only the judiciary may decide Article III cases or controversies. This basic structure gives birth to judicial review. In the course of performing its constitutional responsibility to decide cases or controversies, the judiciary must give primacy to the Constitution over other actions of

84. *See, e.g.*, City of Boerne v. Flores, 521 U.S. 507 (1997); United States v. Klein, 80 U.S. (13 Wall.) 128 (1872).

85. *See generally* Erwin Chemerinsky, *Formalism and Functionalism in Federalism Analysis*, 13 GA. ST. U. L. REV. 959 (1997); William N. Eskridge, Jr., *Relationships Between Formalism and Functionalism in Separation of Powers Cases*, 22 HARV. J.L. & PUB. POL'Y 21 (1998); Laura E. Little, *Envy and Jealousy: A Study of Separation of Powers and Judicial Review*, 52 HASTINGS L.J. 47 (2000).

86. THE FEDERALIST NO. 49 (James Madison), *supra* note 75, at 314.

the federal or state governments.[87] This requires federal judges to interpret the Constitution in the course of resolving conflicts that arise between federal or state law and the Constitution.

Judicial review's roots in the Constitution's text and structure explain why reliance on foreign decisions creates severe constitutional difficulties. Judicial review operates because the Court, in carrying out its Article III duties, must follow the higher law of the Constitution above any inconsistent federal or state statutes. The Constitution is higher law, as Chief Justice Marshall observed in *Marbury*, because it represents the delegation of power from the people to their government.[88] When interpreting the scope and meaning of that delegation of power to the federal government, the federal courts should have no recourse to foreign decisions. Those, after all, interpret documents entirely different from the United States Constitution. The European Court of Human Rights (ECHR), for example, interprets and applies the European Convention on Human Rights of 1950, which was created by the member states of the Council of Europe.[89] The beliefs of the member states of the Council of Europe about the scope of various individual rights circa 1950 have little to do with the extent of the powers the American people delegated to their government in 1788, 1791 or 1865–1870. The European Convention and the ECHR decisions interpreting it do not even purport to relate to our Constitution's delegation or power; rather, the European Convention is an international agreement in which the party states committed themselves to abide by certain requirements.[90] By relying

87. Marbury v. Madison, 5 U.S. (1 Cranch) 137, 178 (1803) ("So if a law be in opposition to the constitution; if both the law and the constitution apply to a particular case, so that the court must either decide that case conformably to the law, disregarding the constitution; or conformably to the constitution, disregarding the law; the court must determine which of these conflicting rules governs the case. This is of the very essence of judicial duty.").

88. *Id.* at 176 ("That the people have an original right to establish, for their future government, such principles as, in their opinion, shall most conduce to their own happiness, is the basis, on which the whole American fabric has been erected. The exercise of this original right is a very great exertion; nor can it, nor ought it be frequently repeated. The principles, therefore, so established, are deemed fundamental. And as the authority, from which they proceed, is supreme, and can seldom act, they are designed to be permanent.").

89. Convention for the Protection of Human Rights and Fundamental Freedoms, Nov. 4, 1950, 213 U.N.T.S. 221.

90. The agreement's preamble opens, "The Governments signatory hereto, *being Members of the Council of Europe . . . .*" *Id.* at 221 (emphasis added).

on foreign sources of law to interpret the Constitution, the Court undermines the very delegation of authority that gives it the power of judicial review.

Moreover, "the People" are the originator of the Constitution and the delegator of all powers under it. The states that are parties to the European Convention, no matter how worthy or progressive in their approach to human rights, are not part of the American polity and were not in 1787 or 1791. At those times, many of the nations that constitute the state parties to the various European treaties on economic integration or human rights were not even democracies that protected individual liberties, but were monarchies without representative governments. If anything, we enjoy our current Constitution precisely because the Americans of the late Eighteenth Century rejected their relationship with Europe, against the efforts of the British Empire.[91] The Union (embodied by the Constitution) was largely designed to serve as an "effective barrier against the Europeanization of American politics."[92] This is not to say that interpreters should not refer to eighteenth-century understandings of British constitutional terms that the Framers borrowed in writing the Constitution. But it cautions against relying on the decisions of postwar Europeans who were neither part of the American polity that drafted and ratified the U.S. Constitution and Bill of Rights, nor descended from ancestors who shared a common political system with the Framers.

Even before the Framing, the Declaration of Independence had denounced King George III for "subject[ing] us to a jurisdiction foreign to our constitution, and unacknowledged by our laws," and for "taking away our Charters, abolishing our most valuable Laws, and altering fundamentally the Forms of our Government."[93] If the Framers considered English law oppressive, it is unlikely that they would have wanted the Constitution's interpreters to defer to the laws and practices of Bourbon France or Spain, Habsburg Austria, Hohenzollern

---

91. *See* Larry D. Kramer, *The Supreme Court, 2000 Term—Foreword: We the Court*, 115 HARV. L. REV 4, 73–74 (2001) ("The colonial experience of resisting King and Parliament served as the model from which the Founders constructed their theories, and the Revolution itself, beginning with the Stamp Act protests, provided their blueprint for opposing a government that exceeded its constitutional authority.").

92. PETER ONUF & NICHOLAS ONUF, FEDERAL UNION, MODERN WORLD: THE LAW OF NATIONS IN AN AGE OF REVOLUTIONS, 1776–1814, at 176 (1993).

93. THE DECLARATION OF INDEPENDENCE paras. 17, 25 (U.S. 1776).

Prussia, or the Papal States. Nor is there any indication that the American people, at any point in their history, have wanted their delegation of authority to the national government to be construed to match the constitutions of foreign nations. Certainly the framers of those documents would not have thought any European treaties in existence at that time should provide a model for constitutional rights; indeed, the framework of international human rights that we have today would have been foreign to them.[94]

Furthermore, a "living constitution" approach to constitutional interpretation—the conviction that the Constitution should be interpreted in light of contemporary attitudes and values—does not justify reliance on foreign precedents. If a living Constitution approach counsels in favor of interpreting the Constitution according to the meaning that "we the People" today would give it, then at the very least, it should be determined how "we the People" think. But there is no indication that the American people today believe that their constitutional rights and distribution of powers should be interpreted in light of foreign judicial decisions. In fact, American attitudes toward international human rights indicate the opposite. The United States has entered into relatively few human rights treaties, and those agreements to which it has consented have been ratified only with significant reservations, understandings, and declarations (RUDs).[95] The RUDs usually contain provisions making clear that the United States considers its existing laws to meet the requirements of the treaty, and that the treaty is non-self-executing.[96] Such a practice undermines any argument that inter-

---

94. How attentive an ear the Court should give to the intent of the Constitution's Framers is often disputed. There is little doubt, however, that the Court itself often claims to place at least some importance on the intended design of the Framers.

95. Many scholars who complain of the United States's unwillingness to join fully the international community readily admit that the cause may be a reluctance to do so on the part of American elected officials. *See, e.g.*, Harold Hongju Koh, *On American Exceptionalism*, 55 STAN. L. REV. 1479 (2003); Johan D. van der Vyver, *American Exceptionalism: Human Rights, International Criminal Justice, and National Self-Righteousness*, 50 EMORY L.J. 775 (2001).

96. At least three modern treaties to which the United States is a party have been qualified with RUDs: the Convention Against Torture and Other Cruel Inhuman or Degrading Treatment or Punishment, Dec. 10, 1984, 1465 U.N.T.S. 85; the International Covenant on Civil and Political Rights, Dec. 19, 1966, 999 U.N.T.S. 171 [hereinafter Political Rights Covenant]; and the International Convention on the Elimination of All Forms of Racial Discrimination, Dec. 21,

national human rights agreements, even those to which the United States is a party, should be given domestic effect. Certainly, the argument for judicial deference to international agreements to which the United States is not a party is even weaker.

Nevertheless, the majority opinion in *Roper* discounted the reservation, proposed by the President and accepted by the Senate, to the International Covenant on Civil and Political Rights (ICCPR).[97] This reservation was addressed to article 6(5) of the ICCPR, which prohibits the juvenile death penalty, and was intended to enable the United States and its component States to retain that form of punishment if they saw fit. The reservation stated, "The United States reserves the right, subject to its Constitutional constraints, to impose capital punishment on any person (other than a pregnant woman) duly convicted under existing or future laws permitting the imposition of capital punishment, including such punishment for crimes committed by persons below eighteen years of age."[98] Despite this recent agreement by the nation's treaty-making branches on the matter, the Court effectively stripped the RUD out of the treaty.[99] The Court cited article 6(5) of the ICCPR to support its conclusion that the juvenile death penalty is unconstitutional.[100] Moreover, the majority opinion gave considerable weight to the prohibition on the juvenile death penalty in article 37 of the Convention on the Rights of the Child, despite the United States's failure to ratify that treaty.[101] An exasperated Justice Scalia was led to wonder whether "the Court ha[d] added to its arsenal the power to join and ratify treaties on behalf of the United States . . . ."[102]

1965, 660 U.N.T.S. 195. On RUDs, see John Yoo, *Treaties and Public Lawmaking: A Textual and Structural Defense of Non-Self-Execution*, 99 COLUM. L. REV. 2218 (1999).

97. Political Rights Covenant, *supra* note 96, 999 U.N.T.S. at 171.

98. *International Covenant on Civil and Political Rights: Hearing Before the Committee on Foreign Relations of the United States Senate*, 102d Cong. 109 (1991) (reservation proposed by the administration).

99. Roper v. Simmons, 125 S. Ct. 1183, 1194 (2005) ("This reservation at best provides only faint support for [the state] petitioner's argument.").

100. *Id.*

101. *Id.* at 1199.

102. *Id.* at 1226 (Scalia, J., dissenting).

III. FOREIGN LAW, THE SUPREMACY CLAUSE, AND THE LAW OF NATIONS

The Supremacy Clause does not mention the law of nations as a form of supreme law. By implication, this omission seems to exclude international law, absent congressional implementation, as a form of law enforceable in federal courts. However, the law of nations has played a role in American jurisprudence from the earliest days of the Republic. This Article considers the two competing formulations of the law of nations—as a form of natural law and as a form of positive law—and concludes that in neither case does it generally provide a valid basis for using foreign materials in constitutional adjudication. Only in the few cases involving constitutional provisions that themselves refer to international law is reliance on foreign sources appropriate.

On its face, the Supremacy Clause[103] does not include foreign or international law as a legal basis for judicial decisions.

> The Constitution, and the Laws of the United States which shall be made in Pursuance thereof; and all Treaties made, or which shall be made, under the Authority of the United States, shall be the supreme Law of the Land; and the Judges in every State shall be bound thereby, any Thing in the Constitution or Laws of any State to the contrary notwithstanding.[104]

The Clause identifies three, and only three, kinds of supreme law within the United States: (1) the Constitution itself; (2) acts of Congress enacted in accordance with the procedures prescribed in Article I, Section 7; and (3) treaties ratified in accordance with the procedures prescribed in Article II, Section 2, Clause 2, together with treaties pre-existing the Constitution that had been ratified "under the authority of the United States" during the period of the Articles of Confederation. Foreign and international laws, other than treaties ratified by the United States, are not enumerated among the three kinds of law that can be "the supreme Law of the Land." Therefore, they should not be treated as outcome-determinative in constitutional adjudication.

---

103. U.S. CONST. art. VI, cl. 2.
104. *Id.*

The purposes of the Supremacy Clause confirm this textual analysis. The Clause helped to cement a federal union. The Framers were gravely concerned that claims of state sovereignty would undermine the nation's independence, threaten the survival of republican forms of government, and undo the achievements of the Revolution. In their view, the Articles of Confederation reinforced the unfortunate centrifugal tendencies of the post-revolutionary period. Article II of the Articles even enshrined the principle that "[e]ach state retains its sovereignty, freedom, and independence, and every Power, Jurisdiction, and right, which is not by this confederation expressly delegated to the United States, in Congress assembled."[105] Thus, to some of the Framers, the Articles amounted to little more than a treaty among sovereign states that could be abrogated by any one of them at will. Given the sharp conflicts of interest between states and regions, the eventual dissolution of the Union was a real possibility. And the Framers feared that the Union's dissolution would result in the emergence of an unstable, unrepublican and war-prone international system in North America that would resemble the uneasy balance of power in Europe. They were convinced that a fragmented nation would inevitably be swept up in European power politics, just as statesmen on both sides of the Atlantic had, prior to the Revolution, predicted that it would.[106]

James Madison recognized the danger:

> The [con]federal system being devoid of both [sanction and coercion], wants the great vital principles of a Political Cons[ti]tution. Under the form of such a Constitution, it is in fact nothing more than a treaty of amity of commerce and alliance, between so many independent and Sovereign States. From what cause could so fatal an omission have happened in the articles of Confederation? from a mistaken confidence that the justice, the good faith, the honor, the sound policy, of the several legislatures would render superfluous any ap-

105. ARTICLES OF CONFEDERATION art. II.

106. "[W]ith the growing importance of commerce and naval strength as elements of national power and their recognition in the theories of mercantilism, public opinion and leading men in France, England, and America realized, in the words of Franklin's opponent in the Canada-Guadeloupe controversy, William Burke, that 'there is a Balance of Power in *America* as well as in *Europe*.'" GERALD STOURZH, BENJAMIN FRANKLIN AND AMERICAN FOREIGN POLICY 113 (1954). *See generally* FREDERICK W. MARKS III, INDEPENDENCE ON TRIAL: FOREIGN AFFAIRS AND THE MAKING OF THE CONSTITUTION (1973) (discussing the foreign affairs concerns of Framers).

> peal to the ordinary motives by which the laws secure the obedience of individuals.[107]

The Supremacy Clause addressed these deficiencies by clearly subordinating state sovereignty to the authority of the federal government where the federal government acts within its proper sphere. The Clause operates, in effect, as a sharp and definitive conflict-of-laws rule, determining which of various laws is to be enforced in the event that a conflict arises. This rule has two aspects: It establishes the priority of the Constitution itself over all other law, whether federal or state; and the priority of federal law, whether in the form of the Constitution, an act of Congress, or a treaty, over any form of state law, including state constitutions.

The Supremacy Clause does not explicitly address the question of conflicts between domestic law and foreign and international law. This omission likely reflects the Framers' understanding that such laws could not serve as the basis for decisions by American courts, absent the consent of the American sovereign. As Chief Justice Marshall later explained in *The Schooner Exchange v. McFaddon*:[108]

> The jurisdiction of the nation within its own territory is necessarily exclusive and absolute. It is susceptible of no limitation not imposed by itself. Any restriction on it, deriving validity from an external source, would imply a diminution of its sovereignty to the extent of the restriction, and an investment of that sovereignty to the same extent in that power which could impose such restriction. All exceptions, therefore, to the full and complete power of a nation within its own territories, must be traced up to the consent of the nation itself. They can flow from no other legitimate source.[109]

The second purpose of the Supremacy Clause is to link the forms of federal law to specific processes of enactment, which require majorities or even supermajorities. To stand as the "supreme Law of the Land," a measure must either be the Constitution itself, a federal law that represents the outcome of the constitutional processes prescribed in Article I, or a treaty rati-

---

107. JAMES MADISON, *Vices of the Political System of the United States*, *in* 9 THE PAPERS OF JAMES MADISON 348, 351 (Robert A. Rutland et al. eds., 1975).

108. 11 U.S. (7 Cranch) 116 (1812).

109. *Id.* at 133–34.

fied under Article II. By requiring that any supreme source of legal authority meets these standards, the Supremacy Clause guarantees that such laws will ultimately stem only from the preferences of political majorities or super-majorities. All forms of federal law under the Supremacy Clause must receive at least the approval of the Senate, if not also the House and the President.[110] Moreover, by subordinating state law to federal law, the Clause ensures that local majorities will not prevail over national majorities when their preferences conflict. The Supremacy Clause makes clear that sources of law that are not enacted through the appropriate federal political process cannot be the Law of the Land.

The Supremacy Clause's fundamental purposes are undermined if foreign and international laws are treated on par with, or as one of, the explicit sources of authority under the Clause. The Framers sought to establish a federal union in large part to stave off "the imminent Europeanization of American politics" and used "modern Europe . . . as a conceptual antitype and foil for their energized, extended republic."[111] The Supremacy Clause furthered that purpose by guaranteeing "that the corporate interests of state governments would always be subordinate to the rights of the sovereign people . . . and that the states would not arbitrarily interfere with the free movement of trade and people across state boundaries."[112] The Clause was thus integral to the creation and maintenance of a viable American nationhood, distinct from, but as strong as, that of any other nation. At the same time, the specific type of union that the Framers envisaged by no means encompassed the consolidation and destruction of the States. They remained free to "flourish as republics, the primary locus of self-government."[113] To invalidate state laws because they do not conform to the laws of foreign legal systems therefore contradicts the policies of the Supremacy Clause in two ways: It undermines the distinctiveness of American law, and it frustrates and discourages the exercise, by the citizens of the States, of the republican freedoms that constitutionally belong to them.

---

110. Bradford R. Clark, *The Supremacy Clause as a Constraint on Federal Power*, 71 GEO. WASH. L. REV. 91 (2003)

111. ONUF & ONUF, *supra* note 92, at 131.

112. *Id.* at 132.

113. *Id.*

This interpretation of the Supremacy Clause is subject to challenge, however. This analysis assumes that the Clause's enumeration of the possible forms of "supreme Law" is exhaustive: that international and foreign law, as such, cannot be "supreme Law" because they are nowhere mentioned in the Clause. But longstanding historical practice, it might be said, contradicts that interpretation.[114] Though neither the "Law of Nations" nor common law is explicitly mentioned in the Supremacy Clause, American courts have long recognized such law. Indeed, the Constitution itself, in Article I, Section 8, Clause 10, refers expressly to the "Law of Nations," thus suggesting that such law can provide a rule of decision for our courts. And if the Law of Nations can provide a rule of decision, then, it may be argued, foreign and international law may legitimately provide a source of law in constitutional cases, notwithstanding the absence of any explicit reference to such law in the Supremacy Clause.

An assessment of this critique requires a determination whether the early understanding of the "Law of Nations" provides any support for treating contemporary foreign and international law as a source of law in constitutional adjudication. Early understanding does not support such a notion.

Some of the earliest commentators on the Constitution appear to have believed that the United States, merely by constituting itself as a nation among other nations in 1776, had assumed an obligation to comply with and enforce the law of nations. In *Chisholm v. Georgia*,[115] Chief Justice John Jay stated that the United States "had, by taking a place among the nations of the earth, become amenable to the law of nations."[116] Likewise, in *Ware v. Hylton*,[117] Justice James Wilson, an influential and prominent Founder from Pennsylvania, wrote that when "the United States declared their independence, they were bound to receive the law of nations, in its modern state of purity and refinement."[118] And President George Washington's

---

114. *See* Stephen G. Calabresi & Stephanie Zimdahl, *The Supreme Court and Foreign Sources of Law: Two Hundred Years of Practice and the Juvenile Death Penalty Decision*, 47 WM. & MARY L. REV. (forthcoming Dec. 2005) (extensive survey of Supreme Court constitutional caselaw relying on foreign law).

115. 2 U.S. (2 Dall.) 419 (1793).

116. *Id.* at 474.

117. 3 U.S. (3 Dall.) 199 (1796).

118. *Id.* at 281.

Attorney General Edmund Randolph, another Framer from Virginia, opined in 1792 that the "law of nations, although not specifically adopted by the constitution or any municipal act, is essentially a part of the law of the land."[119] On this view, the constitutive act of founding the nation in itself subjected the United States to the law of nations; and the change in the nation's constitutional arrangements in 1787 did nothing to alter that subjection.

Furthermore, the Constitution elsewhere authorizes federal courts to apply, as rules of decision, laws that are not the law of the United States under the Supremacy Clause. Article III, Section 2 of the Constitution states that the judicial power shall extend to "all Cases, in Law and Equity, arising under this Constitution, the Laws of the United States, and Treaties made, or which shall be made, under their Authority." That language corresponds closely to the categorization of laws found in the Supremacy Clause. In addition, however, Article III, Section 2 states that the federal judicial power shall extend to other matters, including, for example, "all Cases of admiralty and maritime Jurisdiction." In *American Insurance Co. v. Cantor*,[120] Chief Justice Marshall ruled that although a territorial court of Florida had been vested by an act of Congress with such jurisdiction as arose under the Constitution and laws of the United States, it was not on that account also vested with the authority to decide cases in admiralty.[121] "A case in admiralty does not, in fact, arise under the Constitution or the laws of the United States. These cases are as old as navigation itself; and the law of admiralty and maritime, as it has existed for ages, is applied by our courts to the cases as they arise."[122] The Constitution itself, therefore, appears to license the federal courts, sitting in admiralty, to apply a species of federal common law that represents more than simply filling in gaps in a federal statute. If so, this Article's argument that foreign law is not "law" as specified in the Supremacy Clause, and thus cannot be applied to constitutional interpretation in an outcome-determinative way, may seem to fail.

That line of attack, however, does not lead to the conclusion that contemporary foreign and international law can be applied

119. 1 Op. Att'y Gen. 26, 27 (1792).
120. 26 U.S. (1 Pet.) 511 (1828).
121. *Id.* at 545–46.
122. *Id.* at 545.

comprehensively to decide constitutional issues. Although the original meaning of the "Law of Nations" is uncertain and disputable, it does not appear to be equivalent to the later term "international law" (which was first introduced in 1780 by Jeremy Bentham).[123] Still less did the eighteenth-century law of nations embrace anything resembling contemporary international human-rights law.

There were at least two main points of view expressed by leading American statesmen and jurists during the founding period concerning the foundations or sources of the law of nations. One view held that the law of nations was, at least primarily, an instance of the law of nature as applied to nations. On this approach, the law of nations (or such elements of it as were directly founded on the law of nature) was binding on all nations and could not be varied by local law, including domestic constitutional law. On the alternative view, the law of nations was primarily a matter of convention and derived its force from the consent of each nation, either through treaty or by practice, to be governed by that law.

James Wilson, a proponent of the former view, explained: "The law of nature, when applied to states or political societies, receives a new name, that of the law of nations."[124] Wilson argued that the binding force of the law of nations does not depend on the consent of the nations subject to it:

> I freely admit that there are laws of nations, which are founded altogether upon consent. National treaties are laws of nations, obligatory solely by consent. The customs of nations become laws solely by consent. Both kinds are certainly voluntary. But the municipal laws of a state are not more different from the law of nature, than those voluntary laws of nations are, in their source and power, different from the law of nations, properly so called. Indeed, those voluntary laws of nations are as much under the control of the law of nations, properly so called, as municipal laws are under the control of the law of nature. The law of nations, properly so called, is the law of nature applied to states and sovereigns. *The law of nations, properly so called, is the law of states and sovereigns, obligatory upon them in the same manner, and for the same reasons, as the law of nature is obligatory upon individuals.*

123. *See* ANTONIO CASSESE, INTERNATIONAL LAW IN A DIVIDED WORLD 46 (1986).

124. JAMES WILSON, *Of the Law of Nations*, *in* 1 THE WORKS OF JAMES WILSON 148 (Robert Green McCloskey ed., 1967).

> *Universal, indispensable, and unchangeable is the obligation of both.*[125]

This understanding of the law of nations as the application, in the sphere of nations, of a "[u]niversal, indispensable, and unchangeable" law of nature is of course the latent premise behind Wilson's statement in *Ware v. Hylton* that the United States received the law of nations upon becoming independent. Views akin to Wilson's persisted for decades afterwards in American law. For instance, Justice Story's opinion (as a Circuit Justice) in *United States v. La Jeune Eugenie*,[126] held that the African slave trade was inconsistent with the law of nations because it was "founded in a violation of some of the first principles, which ought to govern nations," that is, a violation of the law of nature.[127] Accordingly, he considered himself bound "in an American court of judicature" to hold that the slave trade was "an offence against the universal law of society" and to subject those who engaged in it to "the penalty of confiscation."[128]

An alternative, more positivistic view of the law of nations was articulated by Secretary of State James Madison in 1806.[129] Madison identified five sources of the law of nations, including the work of publicists and the evidence of treaties, but *not* including the law of nature.[130] For Madison, the law of nations was to be defined instead as consisting of "those rules of conduct which reason deduces, as consonant to justice and common good, from the nature of the society existing among independent nations; with such definitions and modifications as may be established by general consent."[131] Madison gave particular weight to the general treaty practice of nations: "Treaties can be sufficiently general, sufficiently uniform, and of sufficient duration, to attest that general and settled concur-

---

125. *Id.* at 150 (emphasis added). Wilson added that there was also a "voluntary" part of the law of nations, "founded on the principle of consent," which consisted mainly of the "publick compacts and customs received and observed by civilized states." *Id.* at 165.

126. 26 F. Cas. 832 (D. Mass. 1822) (No. 15,551).

127. *Id.* at 846.

128. *Id.* at 847.

129. JAMES MADISON, *An Examination of the British Doctrine, Which Subjects to Capture a Neutral Trade, Not Open in Time of Peace*, *in* 7 THE WRITINGS OF JAMES MADISON, 204 (Gaillard Hunt ed., 1908). For a discussion of the circumstances of this writing, see ONUF & ONUF, *supra* note 92, at 201–11.

130. MADISON, *supra* note 129, at 208.

131. *Id.* at 238.

rence of nations in a principle or rule of conduct among themselves, which amounts to the establishment of a general law."[132]

Under this approach, the law of nations does not depend on, and indeed may be inconsistent with, the law of nature. This was the view of Chief Justice Marshall, writing for the Court in *The Antelope*.[133] That the African slave trade was inconsistent with the law of nature, Marshall conceded, could "scarcely be denied" because it was "generally admitted" that "every man has a natural right to the fruits of his own labour."[134] But Marshall nonetheless also ruled that the slave trade could not be condemned as contrary to the law of nations, because it had been universally practiced by nations since antiquity: "This, which was the usage of all, could not be pronounced repugnant to the law of nations, which is certainly to be tried by the test of general usage. That which has received the assent of all, must be the law of all."[135] Although Madison and Marshall differ in their emphases—the former relying more heavily on treaty law, the latter on state practice—they appear to agree that the law of nations is ultimately founded on state consent, not on the law of nature.

Distinguishing between these two early understandings of the law of nations clarifies the question of the legitimacy of relying on international materials. As both Justice Wilson and Justice Story argued, insofar as the law of nations is a species of the law of nature, *all* positive domestic law, including the Constitution, must be subordinated to it. As Wilson wrote, the law of nations, considered as falling under the law of nature, is "of obligation indispensable" and "of origin divine."[136] If the law of nations has those characteristics, then plainly it may and must

---

132. *Id.*

133. 23 U.S. (10 Wheat.) 66 (1825).

134. *Id.* at 120.

135. *Id.* at 120–21.

136. WILSON, *supra* note 124, at 149. Likewise, Sir William Blackstone, in a passage quoted by Alexander Hamilton, affirmed that the law of nature, "being coeval with mankind and dictated by God himself, is of course superior in obligation to any other. It is binding over all the globe, in all countries, and at all times: No human laws are of any validity, if contrary to this." 1 BLACKSTONE, *supra* note 33, at *41, *quoted in* ALEXANDER HAMILTON, THE FARMER REFUTED (1775), *reprinted in* 1 THE WORKS OF ALEXANDER HAMILTON 53, 62 (Henry Cabot Lodge ed., 1904). *See generally* GERALD STOURZH, ALEXANDER HAMILTON AND THE IDEA OF REPUBLICAN GOVERNMENT 59–61 (1970) (exploring but not deciding the question of whether Hamilton regarded the Constitution as subordinate to natural law).

guide constitutional interpretation. However, from its decision nearly seventy years ago in *Erie Railroad Co. v. Tompkins*[137] through its decision last year in *Sosa v. Alvarez-Machain*,[138] the Supreme Court has rejected the idea of law "as a discoverable reflection of human reason," regarding it instead "in a positivistic way, as a product of human choice."[139] The law-of-nations-as-natural-law defense of the use of foreign and international law is thus preempted, at least within the Court's current jurisprudential horizon.

Nor is the law of nations relevant to constitutional interpretation even as a type of positive law, based only on the consent of sovereigns. In eighteenth-century America as in eighteenth-century England, the law of nations was part of the "general common law," and was not the law of any particular national or local jurisdiction.[140] Like other kinds of common law, however, it could be overruled by statute.[141] The subordination of the law of nations to national statutory law cuts against its use to decide constitutional questions. It would be incongruous if a court could invoke the law of nations to strike down an act of Congress that sought to overrule a tenet of the law of nations if the act itself has preemptive force. Thus, Congress could undoubtedly have banned the importation of African slaves (after 1808),[142] even if the African slave trade remained permissible under the law of nations at the time; and the legality of the slave traffic under international law could hardly have been cited to undermine the constitutionality of such a statute. To be

---

137. 304 U.S. 64 (1938).

138. 542 U.S. 692 (2004).

139. *Id.* at 729; *see also id.* at 744 (Scalia, J., concurring in part and concurring in the judgment) ("The Court recognizes that *Erie* was a 'watershed' decision heralding an avulsive change.").

140. *Id.* at 739 (Scalia, J., concurring in part and concurring in the judgment) ("The law of nations . . . at the time [of the Framing was] part of the so-called general common law" (citing Edward Young, *Sorting Out the Debate Over Customary International Law,* 42 VA. J. INT'L L. 365, 374 (2002); Curtis Bradley & Jack Goldsmith, *Customary International Law: A Critique of the Modern Position,* 110 HARV. L. REV. 815, 824 (1997))). For English law at the time, see 4 BLACKSTONE, *supra* note 33, at *67 ("[T]he law of nations (wherever any question arises which is properly the object of it's [sic] jurisdiction) is here adopted in it's [sic] full extent by the common law.").

141. *See* 1 BLACKSTONE, *supra* note 33, at *42–55 (doctrine of parliamentary sovereignty); *see also* FORREST MCDONALD, ALEXANDER HAMILTON: A BIOGRAPHY 58 (1980) (discussing Blackstone's views on parliamentary sovereignty in relation to the law of nature).

142. *See* U.S. CONST. art. I, § 9, cl. 1.

sure, the Supreme Court held in *Murray v. The Schooner Charming Betsy*[143] that acts of Congress should be construed, if reasonably possible, to avoid conflicts with international law.[144] However, that international or foreign law is relevant when determining the intent of Congress does not mean that they are relevant to determining the constitutional validity of that statute.

Moreover, during John Marshall's lengthy tenure as Chief Justice, the Court appears to have relied on the law of nations only once to decide a question of constitutional interpretation.[145] That case, *Brown v. United States*,[146] addressed whether a congressional declaration of war in 1812 had authorized the President to seize enemy property, found on land at the commencement of hostilities, which U.S. citizens had purchased before the outbreak of the war.[147] The Court held the seizure invalid on the grounds that further legislative authorization was needed before the President could seize enemy property.[148] Relying on treatises by several noted European writers and publicists, the Court found,

> [T]he modern rule [of the law of nations] would then seem to be, that tangible property belonging to an enemy and found in the country at the commencement of war, ought not to be immediately confiscated . . . . It may be considered

143. 6 U.S. (2 Cranch) 64 (1804).

144. *Id.* at 118. For different views of the *Charming Betsy* canon, compare Curtis Bradley, *The* Charming Betsy *Canon and Separation of Powers: Rethinking the Interpretive Role of International Law*, 86 GEO. L. J. 479 (1998) (arguing that the canon is justified as a means to preserve the separation of powers, to shift certain decision making away from the courts to Congress and the President, and to prevent Congress from unintentionally interfering with the diplomatic prerogatives of the President), with Ingrid Brunk Wuerth, *Authorizations for the Use of Force, International Law, and the* Charming Betsy *Canon*, 46 B.C. L. REV. 293 (2005) (arguing that the canon not only shifts certain decisionmaking from the courts to Congress and the President but also prefers legislative over presidential decisionmaking).

145. *See* Calabresi & Zimdahl, *supra* note 114. This excludes the Marshall Court's decisions in admiralty or maritime law, which did indeed frequently cite foreign or international law, but which did not implicate any constitutional questions (other than the scope of the Court's jurisdiction) to which that foreign law was relevant. As noted above, Article III, Section 2 seems to make express provision for the Court to employ foreign and international law in particular categories of cases, including admiralty and maritime law.

146. 12 U.S. (8 Cranch) 110 (1814).

147. *Id*. at 121–22.

148. *Id.* at 127–29.

> as the opinion of all who have written on *jus belli,* that war gives the right to confiscate, but does not itself confiscate the property of the enemy.[149]

Further, the Court stated,

> The constitution of the United States was framed at a time when this rule, introduced by commerce in favor of moderation and humanity, was received throughout the civilized world. *In expounding that constitution, a construction ought not lightly to be admitted which would give to a declaration of war an effect in this country it does not possess elsewhere,* and which would fetter that exercise of entire discretion respecting enemy property, which may enable the government to apply to the enemy the rule that he applies to us.[150]

Brown falls into a narrow and unusual category of cases in which it can be conceded that the use of international law to construe the Constitution is valid: cases in which the constitutional text being construed itself refers to international law. The Declare War Clause is such a clause. It confers on Congress the authority to create and transform relationships existing in international law. Because a purported exercise of that power necessarily implicates international law, it is appropriate for a reviewing court to consult international law in reviewing the scope and effect of Congress's action. To that extent alone is the use of foreign and international law in deciding constitutional issues legitimated by early judicial practice. And even then, *Brown* should not be read to suggest that international law prevented the United States from seizing enemy property; rather, it found that Congress simply had to pass a statute after the declaration of war.

Apart from the Declare War Clause, a few constitutional provisions may arguably permit a court to weigh foreign or international law.[151] A court might properly consult foreign law in determining whether a federal office holder had accepted a "title" of nobility from a foreign state within the meaning of Article I, Section 9, Clause 8. Or a court might examine the content of the "Law of Nations" in determining whether Congress had validly exercised its power under Article I, Sections 8 and 10 to "define and punish . . . Offences" against that law. But

149. *Id.* at 125.

150. *Id.* (emphasis added).

151. Again, this excludes Article III, Section 2.

these provisions are isolated, and are of at most marginal relevance to contemporary constitutional law.

To summarize: The Supremacy Clause makes no reference to foreign or international laws. This is persuasive evidence that such foreign or international laws do not form part of the "supreme Law" of the United States. Although the Constitution makes various references to foreign and international law, including the "Law of Nations," those references do not open the door to the general use of foreign law in interpreting the Constitution. International and foreign law had a constitutionally recognized place within early American jurisprudence, and even provided rules of decision in the non-constitutional areas of admiralty and maritime law. Foreign law might also be relevant to questions of statutory construction. It does not follow, however, that international or foreign law can be controlling in questions of constitutional interpretation. That consequence would be defensible if the law of nations was taken to be part of the law of nature; but such natural-law thinking has been eclipsed in American jurisprudence by *Erie*. The only place that remains for using foreign or international law to decide constitutional questions are those pockets of constitutional text that themselves refer to such law.

IV. EUROPEAN SOURCES OF LAW

The Court's recent turn toward foreign law in constitutional adjudication involves reliance primarily on European decisions. Perhaps it is too early to make this generalization; decisions from Asia, Africa, and Latin America may appear over the next few years. For now, however, the Court's use of foreign decisions appears to be a European phenomenon.[152] This part argues that Europe may not be the appropriate model for American constitutional interpretation.

Europe and the United States share different political histories. While the United States continues to exist in a Lockean framework in which government derives from a social contract with the American people, at least according to some sociolo-

152. When Justices have indicated a desire to incorporate foreign law into American decisions, they have specifically mentioned European law. *See* Anne-Marie Slaughter, *Court to Court*, 92 AM. J. INT'L L. 708, 710–11 (1998) (describing interest by Justices O'Connor and Breyer in applying European Community law).

gists,[153] Europe has been given to fluctuations of ideological extremes. In the Nineteenth Century, many European nations still considered monarchy the best system of government. Indeed, the other European powers intervened after the French Revolution to restore the Bourbon dynasty to power. In the Twentieth Century, monarchy was followed by fascism, socialism, and communism.[154] As history has demonstrated, the performance of these regimes has been less than exemplary. In particular, fascism and communism, which were once viewed by some as advanced, modern ideologies, were adopted by regimes that murdered millions. Should the Supreme Courts of the 1930s or the 1950s have looked to the decisions of Nazi or Soviet courts for guidance?[155] While the relative stability or gradual change in American political philosophy may have prevented the United States from adopting programs or policies viewed by some as progressive or enlightened, it may also have kept the nation from pursuing ideological extremes that resulted in disaster for European nations. Some attribute moderation in American politics, in part, to our written Constitution.[156] Separation of powers and federalism make it difficult to enact any sweeping, ideologically inspired legislation, and the Bill of Rights curtails government action that infringes individual liberties. Appealing to European decisions evades these structural checks on federal lawmaking because Supreme Court decisions are not subject to strict restraints of bicameralism, presentment, and federalism that apply to Congress and the President.

It is therefore doubtful that current European attitudes are superior to those of Americans. While current European consti-

---

153. *See* LOUIS HARTZ, THE LIBERAL TRADITION IN AMERICA: AN INTERPRETATION OF AMERICAN POLITICAL THOUGHT SINCE THE REVOLUTION (1955).

154. *See* PAUL JOHNSON, MODERN TIMES: A HISTORY OF THE WORLD FROM THE 1920S TO THE YEAR 2000 (1999).

155. For a discussion of both jurisprudence and legal scholarship during the Nazi era, see H.W. KOCH, IN THE NAME OF THE VOLK: POLITICAL JUSTICE IN HITLER'S GERMANY (1989); MICHAEL STOLLEIS, THE LAW UNDER THE SWASTIKA: STUDIES ON LEGAL HISTORY IN NAZI GERMANY (Thomas Dunlap trans., 1998). For analysis of the Soviet legal system, see GORDON B. SMITH, REFORMING THE RUSSIAN LEGAL SYSTEM 60–61 (1996) (noting that the social order under Bolshevism, as dictated by the Communist Party, was always the principal concern of Soviet courts).

156. Several authors have discussed the unique impact that the separation of powers doctrine has on constraining each actor and maintaining policy moderation. *See, e.g.*, CRAIG BRADY & DAVID W. VOLDEN, REVOLVING GRIDLOCK: POLITICS AND POLICY FROM CARTER TO CLINTON (1998); KEITH KREHBIEL, PIVOTAL POLITICS: A THEORY OF U.S. LAWMAKING (1998).

tutional schemes may appear to protect individual liberties more effectively, or better balance the tension between government power and individual rights, it is difficult to predict whether history will vindicate the choices that Europe has made. Some Americans once thought that fascism and communism were progressive ideologies from which the United States could learn, but history has demonstrated otherwise. Those ideologies resulted in the oppression of domestic populations, inter-European warfare, and the deaths of millions.[157]

Not only do their histories differ, but the United States and Europe face social and political circumstances so different as to counsel against any attempt to transplant constitutional values from one to the other. Europe has spent the last sixty years turning away from great power conflict and forging a cooperative enterprise that has solved the problem of German ambition and melded former enemies into a broad economic common market.[158] The tools for this amazing integration have not been military power and conquest, but rather supranational institutions, international law, and diplomacy. As Robert Kagan explains, "Europe is turning away from power, or to put it a little differently, it is moving beyond power into a self-contained world of laws and rules and transnational negotiation and cooperation."[159] The United States, on the other hand, relies on power rather than international law, employs military force as much as persuasion, and sees a world threatened by terrorist organizations, rogue nations, and the proliferation of weapons of mass destruction.[160]

The difference between European and American attitudes has promoted the integration of Europe and permitted Europeans to attempt a new experiment in political organization.[161]

---

157. JOHNSON, *supra* note 154, at 261–308, 398–431.

158. ROBERT KAGAN, OF PARADISE AND POWER: AMERICA AND EUROPE IN THE NEW WORLD ORDER (2003).

159. *Id.* at 3.

160. *See id.* ("[O]n major strategic and international questions today, Americans are from Mars and Europeans are from Venus." While Europe "is entering a post-historical paradise of peace and relative prosperity, the realization of Immanuel Kant's 'perpetual peace,'" the United States "remains mired in history, exercising power in an anarchic Hobbesian world where international laws and rules are unreliable, and where true security and the defense and promotion of a liberal order still depend on the possession and use of military might.").

161. *Id.* at 59 ("The integration of Europe was not to be based on military deterrence or the balance of power . . . . [T]he end of the Cold War, by removing

The ability of European nations to put aside their historical animosities and engage in integration may be the result of an American security guarantee. The North Atlantic Treaty Organization and heavy American military presence in Western Europe deterred the Soviet Union and allowed European integration to proceed. As Lord Ismay, the first secretary general of NATO, famously quipped, the purpose of the Atlantic alliance was "to keep the Americans in, the Russians out, and the Germans down."[162] Existing disparities in defense spending have only grown since the end of the Cold War. In the 1990s, Europeans discussed increasing collective defense expenditures from $150 billion to $180 billion a year while the United States was spending $280 billion a year.[163] Ultimately, the Europeans could not, and had no political desire to, emulate high U.S. defense spending. The United States has become the "indispensable nation," without which Europe cannot handle even civil wars along its borders. Only the United States has the ability to project power globally.[164]

Without the United States's willingness to engage in power politics, Europe would not have had the luxury to integrate. If this is correct, then European constitutional values are inappropriate for the United States. These values were developed because European governments enjoyed a different tradeoff between national security and individual liberties and economic prosperity. The United States, which has greater responsibility for keeping international peace and for guaranteeing stability in Europe, faces a different balance between the demands of national security and constitutional liberties.

---

even the external danger of the Soviet Union, allowed Europe's new order, and its new idealism, to blossom fully into a grand plan for world order.").

162. W.R. SMYSER & PAUL H. NITZE, FROM YALTA TO BERLIN: THE COLD WAR STRUGGLE OVER GERMANY 135 (1999).

163. Post-Cold War developments may be largely responsible for European opposition to defense spending increases. *See* Office of Technology Assessment, *European Defense Industries: Politics, Structure, and Markets*, *in* GLOBAL ARMS RACE: COMMERCE IN ADVANCED MILITARY TECHNOLOGY AND WEAPONS 63–82 (1991), *available at* http://www.wws.princeton.edu/ota/disk1/1991/9122/912206.pdf; KAGAN, *supra* note 158, at 22–23 ("Under the best of circumstances, the European role was limited to filling out peacekeeping forces after the United States had, largely on its own, carried out the decisive phases of a military mission and stabilized the situation.").

164. *See* KAGAN, *supra* note 158, at 25 ("Not only were [European voters] unwilling to pay to project force beyond Europe, but, after the Cold War, they would not pay for sufficient force to conduct even minor military actions on their own continent without American help.").

V. CONCLUSIONS

This Article concludes by speculating on *why* the Supreme Court is so attracted to the use of foreign and international law in constitutional adjudication. Two explanations seem plausible.

First, the Court may be seeking to augment its own power. The use of foreign law in constitutional cases makes constitutional law more uncertain and unpredictable, and hence more malleable. The use of foreign law enables the Court to impose the results it wants in any given case, with fewer constraints from the written text of the Constitution, the Constitution's structure and purposes, or even the Court's own past precedents. The Court is also less fettered by specifically American traditions of law or social practice, and freer to adopt European models and customs, if it finds them compelling. This has the effect of expanding the Court's power to pursue different policies that would be foreclosed were it limited to relying on purely American practices or doctrines.

A second possible explanation hypothesizes that the Justices seek not power, but rather reputation and esteem, whether it be the esteem of their peers, or of particular foreign or domestic audiences. Globalization has abetted the emergence of a variety of "global networks," including networks of governmental power-holders who, over time, may form a common outlook based on personal ties, shared experiences, and the like. Such global networks seem to be resulting in the creation of a transnational class of judicial and regulatory elites who are increasingly freed from the constraints of territoriality, national sovereignty, and domestic political constituencies,[165] and whose judicial and administrative decisions reflect an increasingly harmonized outlook.[166] As Professor Jonathan Macey has observed, "[m]any people, particularly those active in the foreign policy community, view regulatory cooperation as an end in itself. Cooperation in the international sphere is a form of

---

165. *See* Wolfgang H. Reinecke, *Global Public Policy*, 56 FOREIGN AFF. 127, 137 n.6 (1997).

166. *See* Kenneth Anderson, *Squaring the Circle? Reconciling Sovereignty and Global Governance Through Global Government Networks,* 118 HARV. L. REV. 1255, 1272, 1302 n.74 (2005) (reviewing ANNE-MARIE SLAUGHTER, A NEW WORLD ORDER (2004)).

global social norm."[167] The phenomenon has deeper and broader implications than the mere engineering of channels for institutional cooperation across national boundaries. The phenomenon appears to be linked to the emergence of what can be called a deterritorialized, "cosmopolitan" moral sensibility, generally shared by governing elites of the advanced nations. The Justices, who stand at the very apex of these elites, seem particularly likely to share this sensibility and to express it in their constitutional decisions.

167. Jonathan A. Macey, *Regulatory Globalization as a Response to Regulatory Competition*, 52 EMORY L.J. 1353, 1373–74 (2003).

RECENT DEVELOPMENTS

IMPOSING NECESSARY BOUNDARIES ON JUDICIAL DISCRETION IN BALLOT ACCESS CASES: *Clingman v. Beaver*, 125 S. Ct. 2029 (2005)

In recent years, when the Supreme Court has reviewed a state's regulation of political parties' ability to define their membership, it has done so under the framework of the First Amendment right to freedom of association.[1] Although the Constitution grants state legislatures broad power to regulate the manner in which elections are conducted,[2] the Court has held that regulations imposing severe burdens on associational rights are subject to strict judicial scrutiny.[3] Lesser burdens, however, are upheld as long as they advance important state interests in a reasonable, nondiscriminatory manner.[4] Some academics have criticized this approach, arguing that when the Court analyzes the validity of electoral regulations through the lens of associational rights, it will grapple inadequately with underlying issues of democratic openness and competition.[5]

Last Term, in *Clingman v. Beaver*,[6] the Supreme Court, again gazing through the lens of associational freedom, upheld an Oklahoma statute that prevented a political party from allowing voters registered with other parties to vote in its primary, dismissing arguments that the statute burdened the associational freedoms of both the party and the voters.[7] In upholding the statute, the Court rightly declined to use the First Amend-

---

1. *See, e.g.*, Cal. Democratic Party v. Jones, 530 U.S. 567, 571 (2000); Timmons v. Twin Cities Area New Party, 520 U.S. 351, 357 (1997); Tashjian v. Republican Party of Conn., 479 U.S. 208, 214 (1986).

2. U.S. CONST. art. I, § 4, cl. 1.

3. *See Timmons*, 520 U.S. at 358.

4. *See id.*

5. *See, e.g.*, Richard H. Pildes, *The Supreme Court, 2003 Term—Foreword: The Constitutionalization of Democratic Politics*, 118 HARV. L. REV. 28, 101–02 (2004); *Samuel Issacharoff, Gerrymandering and Political Cartels*, 116 HARV. L. REV. 593, 645 (2002).

6. 125 S. Ct. 2029 (2005).

7. *Id.* at 2034.

ment to impose upon Oklahoma its own ideas of democracy, and thus avoided the precarious conclusion that voters may be constitutionally entitled to simultaneously associate with multiple political parties. *Clingman* therefore supports the notion of a limited role for the First Amendment in judicial enforcement of political competition.

Under section 1-104 of Oklahoma's election laws, a political party can open its primary only to its registered members[8] and, if it chooses, registered independents.[9] In May 2000, the Libertarian Party of Oklahoma (LPO) made public its intent to open its primary to all registered voters in Oklahoma, including registered members of other parties.[10] Oklahoma's secretary of state denied the LPO's request for members of other parties,[11] and the LPO, along with several Republican and Democratic voters, filed suit in the United States District Court for the Western District of Oklahoma.[12] The plaintiffs sought declaratory and injunctive relief, alleging that section 1-104 unconstitutionally burdened their First Amendment freedoms of political association and speech.[13] They argued that Oklahoma's "semi-closed" primary system should be replaced by a "party-option open primary" system in which each political party could elect to open its primary to voters registered with other parties.[14]

The district court declined to offer relief, finding that section 1-104 did not severely burden the plaintiffs' associational rights, and that any burden it did impose was justified by Oklahoma's important regulatory interest in "matters of structure, stability and party integrity."[15] The court noted that a party-option open primary system would directly impose on the state's other parties by placing them "in the position of con-

8. OKLA. STAT. ANN., tit. 26, § 1-104(A) (West 1997) ("No registered voter shall be permitted to vote in any Primary Election or Runoff Primary Election of any political party except the political party of which his registration form shows him to be a member, except as otherwise provided by this section.").

9. *Id.* § 1-104(B)(1) ("A recognized political party may permit registered voters designated as Independents . . . to vote in a Primary Election or Runoff Primary Election of the party.").

10. *See* Beaver v. Clingman, No. CIV-00-1071-F, 2003 WL 745562, at *3 (W.D. Okla. Jan. 24, 2003).

11. *See id.*

12. *Id.* at *1–2.

13. *Id.* at *2.

14. *Id.*

15. *Clingman*, 2003 WL 745562, at *22.

tributing voters to the [LPO] while maintaining a home for them on their registration rolls";[16] such "poaching" by the LPO would diminish the meaning of all political party designations within the state[17] and likely would affect the results of a number of primary elections.[18] The court notably rejected as paternalistic and unfounded justifications based on preventing harms such as "electoral raiding,"[19] "swamping,"[20] voter confusion, administrative convenience, and poor electoral strategies by the parties.[21]

A unanimous panel of the Court of Appeals for the Tenth Circuit reversed.[22] The court advanced the principle that "a state generally may not restrict the ability of a political party to define the group of citizens that will choose its standard-bearer";[23] accordingly, it found that section 1-104 imposed a severe burden on the plaintiffs' associational rights and demanded strict scrutiny.[24] Employing that standard, the court rejected the state interests that the district court had found determinative. It was unconvinced by the argument that section 1-104 was necessary to promote stability in the political process, noting that both Utah and Alaska use party-option open primaries without suffering political instability.[25] Neither was the court persuaded that Oklahoma had a compelling interest

16. *Id.* at *17.

17. *Id.*

18. *Id.* at *12. The court found that because many recent Oklahoma primary elections were decided by narrow margins (twenty-four percent of the primary elections in 2000 were decided by margins of five percent or less), the crossing over of voters from one party to another would probably affect the outcomes of some close races. *Id.*

19. "Electoral raiding" occurs when voters associated with one party temporarily align themselves with another party to influence the results of the other party's primary, possibly to nominate the candidate that the voters believe will be less likely to triumph in the general election. *See id.* at *18; *Clingman*, 125 S. Ct. at 2034.

20. "Swamping" is the process by which a party's philosophical or political identity (or perhaps both) is diminished or lost as a result of primary election participation by non-party members in numbers high enough to drown out the voices of voting party members. *See Clingman*, 2005 WL 745562, at *18.

21. *Id.* at *18–19.

22. Beaver v. Clingman, 363 F.3d 1048, 1051 (10th Cir. 2004).

23. *Id.* at 1057.

24. *Id.* at 1057–58.

25. *Id.* at 1060. The court did not explain what might constitute political instability. *See id.*

in preventing electoral "poaching."[26] Indeed, the court argued that by finding that a party had a right to not have its members poached, the district court had invented a new right: "the ability of a group to harness and control the associational opportunities of its members."[27]

The Supreme Court reversed.[28] Writing for the Court, Justice Thomas[29] concluded that any burden imposed by section 1-104 was minor and justified by legitimate state interests.[30] First, in a section joined only by a plurality, Justice Thomas expressed a skepticism regarding the existence of a right to associate with one party by voting in its primary, while associating with another party as a registered member.[31] Moreover, he maintained that even if section 1-104 did burden an associational right, the burden imposed was less severe than the burden imposed by the law upheld in *Timmons v. Twin Cities Area New Party*.[32] Next, writing for the Court, Justice Thomas distinguished the case at hand from *Tashjian v. Republican Party of Connecticut*,[33] where the Court held that a law requiring voters to register with a party before voting in the party's primary imposed a severe burden on the rights of both the Republican Party and the Independent voters who wished to associate with it.[34] Jus-

---

26. *Id.* at 1061.

27. *Clingman*, 363 F.3d at 1061.

28. *Clingman*, 125 S. Ct. at 2034.

29. Chief Justice Rehnquist and Justices Scalia and Kennedy joined the opinion in full. Justices O'Connor and Breyer joined except as to the characterization of the burden imposed by section 1-104. *See infra* text accompanying notes 30–31.

30. *Clingman*, 125 S. Ct. at 2035.

31. *Id.* at 2036 ("[A] voter who is unwilling to disaffiliate from another party to vote in the LPO's primary forms little 'association' with the LPO—nor the LPO with him. That same voter might wish to participate in numerous party primaries, or cast ballots for several candidates, in any given race. The issue is not 'dual associations,' but seemingly boundless ones." (citations omitted)).

32. 520 U.S. 351, 359 (1997). In *Timmons*, the Court held that a party's associational rights are not violated when it is prohibited from selecting as a candidate an individual who is also the candidate of another party. *Id.* at 353–54. Justice Thomas reasoned that if Minnesota could constitutionally prevent a party from associating with the candidate of its choice because that candidate refused to disaffiliate from another political party, then a party could also be prevented from associating with a voter who refused to do the same. *Clingman*, 125 S. Ct. at 2036–37.

33. 479 U.S. 208 (1986).

34. *Id.* at 216–17. The *Tashjian* Court went on to determine that the interests advanced by the state—administrative convenience, prevention of raiding, avoidance of voter confusion, and protecting the integrity and stability of the two-party system—were not compelling, and, in the case of raiding, inapplicable to the independent voters at issue. *Id.* at 217–25.

tice Thomas argued that the Oklahoma statute imposed a less severe burden than the Connecticut statute,[35] noting that, whereas the law in *Tashjian* burdened Independent voters by requiring them to affiliate publicly with a party to vote in any primary election, section 1-104 allowed parties to open their primaries to Independent voters. Thus, the voters at issue did not face the burden of having to affiliate publicly with a party, since they had already done so.[36] Justice Thomas therefore concluded that strict scrutiny did not apply.[37]

Justice Thomas next determined that section 1-104 advanced three important state interests.[38] First, it "preserv[ed] [political] parties as viable and identifiable interest groups."[39] Second, it aided parties in their electioneering and party-building efforts by preserving the integrity of party membership lists as indicators of who will vote in their primaries.[40] Third, it helped prevent parties from raiding voters of other parties.[41] As a final matter, Justice Thomas declined to respond to the argument that section 1-104, when taken together with several other of the state's election laws, severely burdened the respondents' associational rights by effectively preventing voters from changing their party affiliations before a primary election.[42] Re-

---

35. *Clingman*, 125 S. Ct. at 2038.

36. *Id.*

37. *Id.* at 2039. Justice Thomas argued that "[t]o deem ordinary and widespread burdens like these severe would subject virtually every electoral regulation to strict scrutiny, hamper the ability of States to run efficient and equitable elections, and compel federal courts to rewrite state electoral codes." *Id.*

38. *Id.*

39. *Id.* at 2039 (citing Nader v. Shaffer, 417 F. Supp. 837, 845 (D. Conn. 1976)). Justice Thomas argued that this interest was important partly because it helped prevent primary election outcomes that would confuse voters about the ideology of the candidate nominated, and partly because it would make every registered party affiliation in the state significantly less meaningful as an actual index of the registered voter's ideology. *See id.* at 2039–40.

40. *Clingman*, 125 S. Ct. at 2040.

41. *Id.* at 2040–41.

42. *See id.* at 2042. If a party's candidate does not receive at least ten percent of the votes cast for President or governor in the general election, it loses recognized party status. *See* OKLA. STAT. ANN. tit. 26, § 1-109(A) (West 1997). When a party loses recognized status, those individuals registered with that party become listed as Independents. *See id.* § 110. Because no third party in recent history has been able to meet the ten percent threshold, voters who sympathize with third parties but do not wish to be hassled by repeatedly having to switch their affiliations back from Independent status may simply affiliate with one of the two major parties out of convenience, despite their third-party sympathies. *See Clingman*, 125

spondents raised the argument for the first time before the Supreme Court, and Justice Thomas found no reason to depart from the Court's ordinary practice of not considering arguments so raised.[43]

Justice O'Connor issued a concurring opinion.[44] She argued that although the burden imposed by section 1-104 was not sufficiently severe to trigger strict scrutiny,[45] it was not as trivial as Justice Thomas suggested.[46] She argued that a voter may form a cognizable association with a political party merely by voting in its primary[47] and that a voter may form "dual associations" with two different political parties.[48] A voter may wish to associate with one party by voting in its primary, yet retain his or her affiliation with another party, because of an abiding partisan commitment or the objective costs of switching parties.[49] Justice O'Connor would therefore have set forth a rule that "where a party invites a voter to participate in its primary and the voter seeks to do so, we should begin with the premise that there are significant associational interests at stake."[50] Finally, writing only for herself, Justice O'Connor noted that although the Court was correct not to consider whether section 1-104 imposed a severe burden when taken together with Oklahoma's other election laws, she was seriously troubled by the entirety of Oklahoma's primary system and might have ruled otherwise had the argument been properly raised.[51]

S. Ct. at 2046 (O'Connor, J., concurring). Respondents also argued that Oklahoma's registration deadlines leave an extremely narrow window between the time that a party regains recognized status and the time that voters must affiliate with a party, thus making affiliation with a third party particularly difficult. *See id.*

43. *See Clingman*, 125 S. Ct. at 2041–42.

44. *Id.* at 2044 (O'Connor, J., concurring). Justice Breyer joined the opinion except as to the discussion of the interaction of section 1-104 with Oklahoma's other election laws. *See infra* text accompanying note 51.

45. *Id.* at 2044–45.

46. *Id.* at 2042.

47. *Id.* at 2043.

48. *Id.* at 2043–44. Justice O'Connor further explained that

> The validity of voters' and parties' interests in dual associations seems particularly clear where minor parties are concerned. For example, a voter may have a longstanding affiliation with a major party that she wishes to maintain, but she may nevertheless have a substantial interest in associating with a minor party during particular election cycles or in elections for particular offices.

Id. at 2043.

49. *Clingman*, 125 S. Ct. at 2043–44.

50. *Id.* at 2044.

51. *Id.* at 2046–47. Justice Breyer did not join this portion of the opinion.

Justice Stevens dissented.[52] He argued that section 1-104 severely burdened the voters' First Amendment right to vote for the candidate of their choice.[53] Indeed, he argued that the burden was so great that, in effect, it amounted to a prohibition, because, for many voters, changing parties can be a "significant decision"[54] and "the denial of the right to vote cannot be cured by the ability to participate in a subsequent or different election."[55] Additionally, he asserted that section 1-104 placed a severe burden on the LPO's ability to select its standard-bearer and shape its message to the public.[56]

Justice Stevens went on to dismiss each of the asserted state interests as "either irrelevant or insignificant."[57] First, he argued that the preservation of the parties as identifiable interest groups was an illegitimate interest because it is improper for the state to tell a political party what its message should be.[58] Second, he found the asserted interest in avoiding voter confusion to be both paternalistic and contrary to the district court's findings.[59] Third, he argued that the state had no valid interest in ensuring that party labels remain rigid identifiers of voter preferences because a state has no valid interest in defining what it means to be a member of a particular party.[60] Fourth, he maintained that the state's asserted interest in avoiding the "administrative inconvenience" of inaccurate voter roles was outweighed by the associational interests burdened by the statute.[61] Last, he argued that any interest in avoiding electoral

---

52. *Id.* at 2048 (Stevens, J., concurring). Justice Ginsburg joined the dissent in full. Justice Souter joined except as to the discussion of the effect that the Court's opinion would have on the two-party system. *See infra* text accompanying notes 63–65.

53. *See Clingman*, 125 S. Ct. at 2048 (Stevens, J., dissenting).

54. *Id.* at 2049 n.1.

55. *Id.*

56. *Id.* at 2049–50.

57. *Id.* at 2051. Justice Stevens determined that "[n]o matter what the standard, [the asserted interests] simply do not outweigh the interests of the LPO and its voters," *id.* at 2054, thus suggesting that he did not believe that Oklahoma's law was supported by even a rational basis. In effect, Justice Stevens may have been arguing that the State's asserted interests do not "bear some plausible relationship to the burdens [which they] place[] on political parties." *See* Timmons v. Twin Cities Area New Party, 520 U.S. 351, 375 (1997) (Stevens, J., dissenting).

58. *Clingman*, 125 S. Ct. at 2051.

59. *Id.* at 2052.

60. *Id.*

61. *Id.* at 2052–53.

raiding was based on speculation and, in the case of harms that a party might inflict on itself, outside of the state's power to regulate.[62] Justice Stevens concluded that the Court had given undue deference to the state's supposed interest in preserving the two-party system.[63] Noting the recent proliferation of "safe" congressional districts, he argued that primaries are replacing general elections as the most common method of determining the composition of legislatures,[64] and that because the Court upheld a statute that limited political competition at the primary stage, all participants in the political market were harmed.[65]

Clingman's addition to the caselaw of ballot access lays bare two fundamental points. First, the inherent subjectivity of the First Amendment "severity of the burden" test allows for the Justices' competing views about the role of the courts in promoting an open and competitive democracy to drive what is purportedly an analysis of associational freedoms. Second, whereas individual voters may develop complex associational preferences within the political marketplace, those preferences simply are not mirrored by the current judicial understanding of First Amendment protections.

Employing a "severity" test is an inherently subjective enterprise: Different members of the Court described the same burden on associational rights as either a "minor barrier[]," [66] a "modest and politically neutral burden [that was] not altogether trivial," [67] or a "heavy burden" [68] that might be best classified as a "prohibition" on protected activity.[69] Adding to the confusion, the position taken by each Justice appears inconsistent with the position each took five years ago in *California Democratic Party v. Jones*.[70] In *Jones*, the Court, in an opinion authored by Justice Scalia and joined by every member of the *Clingman* majority (plus Justice Souter), struck down California's blanket primary law, which had allowed a voter to participate in the primary of whatever party she chose, regardless

62. *Clingman*, 125 S. Ct. at 2053.
63. *Id.* at 2054. Justice Souter did not join this portion of the opinion.
64. *Id.*
65. *See id.*
66. *Id.* at 2039 (majority opinion).
67. *Clingman*, 125 S. Ct. at 2045 (O'Connor, J., concurring).
68. *Id.* at 2049 (Stevens, J., dissenting).
69. *See id.*
70. 530 U.S. 567 (2000).

of whether the party wanted to open its primary to nonmembers.[71] The Court found that the law placed a severe burden on a party's right not to associate with the members of other parties in its primary,[72] and that the law was not narrowly tailored to advance a compelling state interest.[73] In dissent, Justice Stevens[74] argued that a political party has no First Amendment right to exclude nonmembers from its primary, particularly given the strong public function that the primary serves.[75] Why then, in *Clingman*, did all the Justices save Souter take stances on the associational autonomy of parties that seem to counter those they took in *Jones*?

Justices Stevens and Ginsburg seem to maintain that, all else equal, the Court should favor results that expand opportunities for voters to participate in the selection of candidates and elected officials.[76] In *Jones*, Justice Stevens wrote, "States should be free to experiment with reforms designed to make the democratic process more robust by involving the entire electorate in the process of selecting those who will serve as government officials."[77] In contrast, Justice Stevens read the statute in *Clingman* as effectively prohibiting voters from casting ballots for the candidates of their choice in the primary,[78] thereby curtailing their ability to participate in the selection process. Under this view, the associational rights that matter most are those that promote inclusion, opportunity, and choice;[79] the right to

---

71. *Id.* at 569.

72. *Id.* at 581–82.

73. *Id.* at 585.

74. *Jones*, 530 U.S. at 590 (Stevens, J., dissenting). Justice Ginsburg joined the portion of the opinion discussing the First Amendment.

75. *Id.* at 594–95 ("[P]rimary elections, unlike most 'party affairs,' are state action. The protections that the First Amendment affords to the 'internal processes' of a political party do not encompass a right to exclude nonmembers from voting in a state-required, state-financed primary election.").

76. *See, e.g.*, *id.* at 595–96 ("When a State acts not to limit democratic participation but to expand the ability of individuals to participate in the democratic process, it is acting not as a foe of the First Amendment but as a friend and ally.").

77. *Id.* at 601.

78. *Clingman*, 125 S. Ct. at 2049.

79. *See Jones*, 530 U.S. at 601 (Stevens, J., dissenting) ("I would also give some weight to the First Amendment associational interests of nonmembers of a party seeking to participate in the primary process, to the fundamental right of such nonmembers to cast a meaningful vote for the candidate of their choice, and to the preference of almost 60% of California voters—including a majority of registered

exclude, although part of the freedom of association that most expressive organizations enjoy,[80] limits the ability of voters to select their candidates and officials.[81] As Justices Stevens and Ginsburg made clear in *Clingman*, they believe that the associational interests in any given case are of secondary concern relative to the Court's main function in expanding the franchise and increasing electoral competition.[82]

One therefore might be tempted to infer that for the Justices composing the Thomas plurality in *Clingman*, the consistent theme in their rulings is a raw antipathy toward minor parties. In *Jones*, they rejected the notion that party affairs are wholly public affairs, free of First Amendment protections,[83] and they championed the freedom of the parties to select their standard-bearers without being forced to associate with particular individuals.[84] Then, in *Clingman*, when robust associational freedom would have benefited a minor party, the Justices found the burden on association to be less than severe.[85] Indeed, when

---

Democrats and Republicans—for a blanket primary." (footnotes and citations omitted)).

80. An expressive organization may generally decline to associate with those individuals whose participation in that organization would interfere with its expressive activity. *See*, *e.g.*, Boy Scouts of Am. v. Dale, 530 U.S. 640, 644 (2000); Hurley v. Irish-American Gay, Lesbian and Bisexual Group of Boston, 515 U.S. 557, 581 (1995).

81. *See Jones*, 530 U.S. at 595 (Stevens, J., dissenting) ("The so-called 'right not to associate' that the Court relies upon, then, is simply inapplicable to participation in a state election. A political party, like any other association, may refuse to allow nonmembers to participate in the party's decisions when it is conducting its own affairs . . . . But an election, unlike a convention or caucus, is a public affair." (footnote and citation omitted)).

82. *See Clingman*, 125 S. Ct. at 2048 (Stevens, J., dissenting) ("While the voters in this case certainly have an interest in associating with the LPO, they are primarily interested in voting for a particular candidate, who happens to be in the LPO. Indeed, I think we have lost sight of the principal purpose of a primary: to nominate a candidate for office." (citation omitted)); s*ee also* Timmons v. Twin Cities Area New Party, 520 U.S. 351, 381–82 (1997) (Stevens, J., dissenting) ("Indeed, it is a central theme of our jurisprudence that the entire electorate . . . will benefit from robust competition in ideas and governmental policies . . . ."); Pildes, *supra* note 5, at 123–24.

83. *Jones*, 530 U.S. at 573.

84. That freedom would still be cabined by other provisions of the Constitution, such as the Equal Protection Clause. *See id.* at 573 n.5.

85. *See Clingman*, 125 S. Ct. at 2053 n.11 (Stevens, J., dissenting) ("In *Jones,* the Court concluded that the associational interests of the parties trumped state interests that were much more compelling than those asserted in this case. Here, by contrast, where the associational interests are being asserted by a minor party rather than by one of the dominant parties, the Court has reversed course and

one considers that the Justices in the Thomas plurality also voted in *Timmons* to uphold a statute that impinged on parties' associational interests[86] and potentially reduced the ability of minor parties to thrive,[87] the thesis of raw antipathy appears sound. Along this line of thought, Professor Richard Pildes has posited that the majority opinions in *Timmons* and *Jones* reflect a strong preference for electoral stability, which is threatened by the ascendance of third parties.[88]

Although the Justices in the Thomas plurality certainly feared the creation of ripples in the electoral system,[89] a fair reading of *Clingman* reveals a careful attempt to advance not only stability, but also the competing values of free association, a robust franchise, and deference to legislatures. A deep sense of balancing is conveyed where, in both *Clingman* and *Timmons*, the Court went out of its way to note the various activities that each state law did *not* prohibit.[90] In both cases, the Court seems to have said that whatever burdens were imposed, those burdens were not severe because voters remained free to choose the party with which they would affiliate, and each party retained the same legal right to attempt to win over voters (in the traditional, exclusive sense).[91]

---

rejected those associational interests as insubstantial compared to the interests asserted by the State.").

86. *See Timmons*, 520 U.S. at 363.

87. *See id.* at 361–62.

88. *See* Pildes, *supra* note 5, at 126–30.

89. For instance, the *Clingman* Court argued that Oklahoma had a valid interest in protecting the parties' electioneering and party-building efforts. *See* 125 S. Ct. at 2040.

90. *See id.* at 2035–36 ("Oklahoma has not sought through its electoral system to discover the names of the LPO's members; to interfere with the LPO by restricting activities central to its purpose; to disqualify the LPO from public benefits or privileges; or to compel the LPO's association with unwanted members or voters. The LPO is free to canvass the electorate, enroll or exclude potential members, nominate the candidate of its choice, and engage in the same electoral activities as every other political party in Oklahoma. . . . Nothing in § 1-104 prevents members of other parties from switching their registration to the LPO or to Independent status." (footnote and citations omitted)); *Timmons*, 520 U.S. at 363 ("Minnesota's laws do not restrict the ability of the New Party and its members to endorse, support, or vote for anyone they like. The laws do not directly limit the party's access to the ballot. They are silent on parties' internal structure, governance, and policymaking.").

91. Under such a reading, the Court's holding in *Tashjian* may be reconciled as a necessary step to expand the franchise: If parties had not been required to associate with Independents, then voters who chose not to affiliate with a party

Further, the plurality's decision to limit associational freedom need not be seen as inconsistent with *Jones*, as it reflects the discrepancy between the associational preferences that voters may form and the associational protections that are feasible. In her concurrence, Justice O'Connor emphasized the concept of "dual associations,"[92] whereby a voter might have a recognized associational interest with more than one political party. That voters might desire to associate with multiple parties is hardly disputed; in American congressional elections, significant numbers of voters choose to associate with one party by voting for its candidate while formally affiliating with a different party.[93] In declining to grant First Amendment protections to dual associations as described by Justice O'Connor, the Thomas plurality did not reject the idea that complex attachments may develop within a voter's mind or that an individual American might have the capacity to "contain multitudes";[94] rather, the plurality understood that, although parties certainly have associational rights as recognized in *Jones*, a jurisprudence that recognizes a right to dual associations will invite challenges to a wide swath of election laws.[95] The plurality's concern is hardly alarmist, given that Justice Stevens embraced this expanded role for the Court, arguing in his *Clingman* dissent that voters' associational interests should overcome a party's interest in defining its membership whenever the electoral district is "safe" enough for one party that the primary becomes

---

would not be entitled to vote in *any* primary (which is a much greater burden on the right to vote than is not being able to vote in the primary of one's choice).

92. *Clingman*, 125 S. Ct. at 2043–44 (O'Connor, J., concurring).

93. *See, e.g.*, GARY C. JACOBSON, THE POLITICS OF CONGRESSIONAL ELECTIONS 107–10 (5th ed. 2001) (describing the fluctuation in partisan defection rates in U.S. House and Senate elections). *See also* Tashjian v. Republican Party of Conn., 479 U.S. 208, 215 (1986) ("[T]he act of formal enrollment or public affiliation with the Party is merely one element in the continuum of participation in Party affairs, and need not be in any sense the most important.").

94. WALT WHITMAN, *Song of Myself*, *in* LEAVES OF GRASS 22, 72 (Bantam Books 1983) (1892) ("Do I contradict myself? / Very well then I contradict myself, / (I am large, I contain multitudes.)").

95. The plurality feared that, rather than simply recognizing voters' interest in associating with two parties, it would be opening the door to "seemingly boundless" associations. *See Clingman*, 125 S. Ct. at 2036. That the LPO wanted to be a party of second association for those voters is beside the point. Because a regulation that imposes a severe burden on a voter's associational freedom compels strict scrutiny, *see id.* at 2035, an expansion of recognized associational interests would mean that strict scrutiny would be employed in more cases.

the only true contest.[96] In rejecting the concept of protected dual associations, the plurality limited the possibility of the Court taking a more proactive stance in rewriting primary laws across the nation.[97]

Ultimately, a wide variety of electoral regulations, down to single member districts and first-past-the-post voting, work to limit competition and voter choice.[98] If courts are to respect legislatures' constitutional role in determining the manner in which elections will be conducted,[99] then the Court must be able to limit the path to strict scrutiny of such regulations. If the extent to which a statute restricts competition and choice is to be the test for whether the legislature has overreached, what is to separate permissible from impermissible restrictions?[100] Alternatively, if the Court seeks to strike down statutes on the basis of an impermissible intent to entrench incumbents,[101] it will find the analysis to be extraordinarily messy.[102] Given that

---

96. *See id.* at 2049 n.2 (Stevens, J., dissenting). In effect, Justice Stevens has argued for taking the competition that gerrymandering has sapped from the general election and transplanting it into the primary of whichever party is dominant in a particular district.

97. As further evidence that the plurality sought to limit any associational right that might be used in a later case to expand the role of the Court, the plurality did not reach the issue of whether political parties have an associational right to not have their members poached. *Compare* Beaver v. Clingman, No. CIV-00-1071-F, 2003 WL 745562, at *17 (W.D. Okla. Jan. 24, 2003) *with* Beaver v. Clingman, 363 F.3d 1048, 1061 (10th Cir. 2004).

98. *See, e.g.*, Timmons v. Twin Cities Area New Party, 520 U.S. 351, 362 (1997); GARY W. COX, MAKING VOTES COUNT: STRATEGIC COORDINATION IN THE WORLD'S ELECTORAL SYSTEMS ch. 2 (1997).

99. *See* U.S. CONST. art. I, § 4, cl. 1.

100. Justice Stevens has suggested that regulations might be differentiated based on the extent to which they burden associational rights and the extent to which overturning the statute would affect political stability. *Timmons*, 520 U.S. at 379–80. In form, this test may not be terribly different from the balancing done by the Justices composing the Thomas plurality; after all, these same Justices struck down a statute that burdened association in *Jones* and championed stability in *Timmons*. *See supra* note 87 and accompanying text. In substance, however, the test advanced by Justice Stevens appears to have few teeth, given his implicit willingness to accept dual associations in *Clingman*. Moreover, on what basis should a judge decide that certain electoral structures tend too much toward political instability?

101. *See, e.g.*, Pildes, *supra* note 5, at 137–38; Michael J. Klarman, *Majoritarian Judicial Review: The Entrenchment Problem*, 85 GEO. L.J. 491, 528–30 (1997).

102. Professor Klarman suggests that a legislature's illicit intent to stifle competition through a self-entrenching ballot access restriction might be determined by examining the timing of the legislative enactment and the relative stringency of the restriction. *See* Klarman, *supra* note 101, at 535–36. The ability of courts to

the regulation of elections inherently involves the balancing of competing democratic values,[103] the danger of judicial overreaching by an overly zealous court is particularly acute. If the preexisting and highly subjective severity test had been combined with constitutional protection for multiple associations, courts would have found themselves in the possession of great freedom to second-guess legislatures but with little to no guidance for how to do it. Such an untenable result suggests that when a jurisprudence of democracy masquerades under the trappings of the First Amendment, appropriate concerns over judicial manageability will rightly cabin the result.[104]

Lowell J. Schiller

---

objectively analyze the relative "stringency" or "severity" of a restriction should be questioned in light of the Court's widely divergent assessments of severity in *Clingman*. *See supra* text accompanying notes 66–69. Additionally, when, as in *Clingman*, a regulation has been in place for over thirty years and has previously been amended to expand opportunities for voters to participate, *see* S.B. 149, 41st Leg., 1st Reg. Sess. (Okla. 1987), it becomes nearly impossible to determine whether the regulation's *continued* existence is the product of illicit motives of legislative self-entrenchment.

103. *See, e.g.*, Richard H. Pildes, *Democracy and Disorder*, *in* THE VOTE: BUSH, GORE AND THE SUPREME COURT 151 (Cass R. Sunstein & Richard A. Epstein eds., 2001).

104. *Cf.* Vieth v. Jubelirer, 541 U.S. 267, 278 ("One of the most obvious limitations imposed by [Article III, § 1 of the U.S. Constitution] is that judicial action must be governed by *standard,* by *rule*. . . . [L]aw pronounced by the courts must be principled, rational, and based upon reasoned distinctions.").

PLACING UNNECESSARY LIMITS ON VOTING AND ASSOCIATIONAL FREEDOMS: *Clingman v. Beaver*, 125 S. Ct. 2029 (2005)

The right to vote candidates of one's choice into office is at the core of representative democracy. The ability to organize on a large scale with willing groups of voters to elect a mutually desirable candidate is prerequisite to any enjoyment of this right. If the ability to appeal to fellow voters is too heavily restricted, then society will travel down the perilous path of having its leaders selected by entrenched party leaders rather than by the considered consensus of the people. Last Term, in *Clingman v. Beaver*,[1] the Supreme Court acquiesced in a state's decision to hamper voters' freedom to associate for political purposes and their ability to vote for the candidate of their choice. This decision, which authorized states to deny parties and voters registered with other parties the right to come together for primary elections, serves, in Justice Stevens's words, "a naked interest in protecting the two major parties."[2]

The plaintiffs in the trial court were a group of Oklahoma voters, registered with various political parties, as well as the Libertarian Party of Oklahoma (LPO), and several Libertarian candidates for office.[3] The LPO expressed an interest in opening its primary to voters enrolled in other parties,[4] and the voters desired to participate in the Libertarian primary without abandoning their principal party affiliation.[5] All plaintiffs sued in the United States District Court for the Western District of Oklahoma under 42 U.S.C. § 1983 and other statutes, claiming

---

1. 125 S. Ct. 2029 (2005).

2. *Id.* at 2055 (Stevens, J., dissenting).

3. Beaver v. Clingman, No. CIV-00-1071-F, 2003 WL 745562, at *1–2 (W.D. Okla. Jan. 24, 2003).

4. The specific reasons advanced for such interest by the LPO included the ability to "reach out to Libertarian-oriented voters of other political affiliations," produce candidates with a greater likelihood of being elected, garner positive publicity, and generally further Libertarian ideology. *Id.* at *5.

5. *Id.* at *2.

that Oklahoma's semi-closed primary system[6] violated their rights to free political association and free speech under the First and Fourteenth Amendments to the U.S. Constitution.[7] The plaintiffs requested that the district court declare unconstitutional the Oklahoma statutes restricting parties from opening their primaries to persons registered with other parties,[8] and that the court enjoin the defendants—the Oklahoma State Election Board and its members—from enforcing those laws.[9]

The district court denied the plaintiffs' request for declaratory and injunctive relief, but not before making several factual findings that were favorable to the plaintiffs. The court first determined that there would be no significant voter confusion from granting the plaintiffs' request.[10] Second, the court dismissed the state's concerns about "raiding" of minor parties, in part because "[t]he reality of raiding as a phenomenon existing other than in the minds of supporters of closed primaries is a matter of conjecture."[11] While the district court acknowledged that "swamping," or "the control of a minor party's candidate selection by voters affiliated with other parties,"[12] might well

---

6. As explained by the district court, there are "four general types of political primary": closed, semi-closed, open, and blanket. *Id.* at *7. Only party members can vote in closed primaries, although Independents can also vote in semi-closed primaries. Voters in an open primary may select in which party's primary they vote; in contrast, blanket primaries allow the voter to vote in a different party's primary for each office. *Id.* Oklahoma law required primaries to be closed unless the party indicated a preference for semi-closed primaries. *See* OKLA. STAT. ANN. tit. 26, § 1-104 (West 1997).

7. *Clingman*, 2003 WL 745562, at *2.

8. OKLA. STAT. ANN. tit. 26, § 1-104 (West 1997) reads, in relevant part:

> A. No registered voter shall be permitted to vote in any Primary Election or Runoff Primary Election of any political party except the political party of which his registration form shows him to be a member, except as otherwise provided by this section.
>
> B. 1. A recognized political party may permit registered voters designated as Independents pursuant to the provisions of Section 4-112 of this title to vote in a Primary Election or Runoff Primary Election of the party.

9. *Clingman*, 2003 WL 745562, at *2.

10. *Id.* at *9 (rejecting this concern because "very simple rules for voting eligibility can be posted at polling places when the primary and runoff elections are conducted").

11. *Id.* at *8 & n.5; *see also id.* at *18 (dismissing the raiding argument as "at bottom, another paternalistic argument of the kind which has found little support from the Supreme Court"). Raiding is "malicious crossover voting." *Id.* at *8 n.5 (quoting Cal. Democratic Party v. Jones, 530 U.S. 567, 579 (2000)). According to the defendants, it would "give non-members of a minor party the ability to overtake the minor party as a flag of convenience, thus affecting the ideology of the ticket representatives of the party." *Id.* at *8.

12. *Id.* at *10.

occur, given the extremely small number of Libertarian voters,[13] it did not view this phenomenon as adequate reason for upholding Oklahoma's election statutes.[14] However, the court, applying a balancing test,[15] held the electoral laws constitutional because it found that allowing voters registered with one party to vote elsewhere could alter the primary results of the party of original registration.[16]

The Tenth Circuit reversed. It held that the burden imposed on the plaintiffs by the Oklahoma statute was severe, and thus applied strict scrutiny.[17] The court justified its decision by integrating the Supreme Court's holdings in *California Democratic Party v. Jones*[18] and *Tashjian v. Republican Party of Connecticut*[19] to create "the clear and unavoidable implication . . . that a state generally may not restrict the ability of a political party to define the group of citizens that will choose its standard-bearer."[20] After quickly disposing of the swamping and raiding justifications on grounds similar to those employed by the trial court,[21] the appellate court found Oklahoma's interest in political stability to be not compelling "*in the circumstances of*" the

---

13. *Id.* at *11 (noting that defendants' expert witness suggested that swamping was already occurring with the semi-closed primary, as many more Independents than Libertarians were voting in the LPO primary); *see also id.* at *4 (finding that at least ninety-five percent of voters in the Libertarian primary for United States Senate were non-Libertarian). Of course, there were very few Libertarian voters in large part because Libertarian voters were returned to Independent status after each gubernatorial or presidential election in which the LPO candidate failed to receive ten percent of all votes cast. *See id.* at *6 (citing OKLA. STAT. ANN. tit. 26, § 1-109 (West 1997)).

14. *See id.* at *12 (minimizing the importance of raiding and swamping "in the classic examples" because they "affect the minor party which chooses to open its primary" rather than others); *see also id.* at *18 (finding swamping, like raiding, is an "essentially paternalistic justification").

15. *Id.* at *14; *see also id.* at *14 n.6 (characterizing the case as involving "*competing* First Amendment interests" of the LPO and other parties).

16. *Id.* at *12, *23.

17. Beaver v. Clingman, 363 F.3d 1048, 1057–58 (10th Cir. 2004).

18. 530 U.S. 567 (2000) (holding that forcing parties to open their primaries to all voters in a blanket primary is unconstitutional).

19. 479 U.S. 208 (1986) (holding that, when a political party wishes to open its primaries to Independents, a state cannot constitutionally prohibit them from doing so).

20. *Clingman*, 363 F.3d at 1057. As an additional justification, the court noted that strict scrutiny is usually applied to "almost all state regulations that disrupt the decision making processes of the parties." *Id.* at 1057 n.6.

21. *See id.* at 1058.

case at bar.[22] Granting that "some election results might change if Oklahoma switches from a semi-closed primary system," the state did not show why this "would undermine the integrity of the political process, or . . . diminish [the state's] political stability."[23] Regarding the trial court's assertion that the First Amendment interests of other political parties should be considered, the Tenth Circuit found that this would create a novel associational right to "harness and control the associational opportunities of . . . members."[24]

The Supreme Court reversed. Writing for himself and three other Justices,[25] Justice Thomas found that the burdens on the plaintiffs were no more restrictive than burdens that had previously been held nonsevere.[26] Writing for the Court, Justice Thomas criticized the *Tashjian* court for "appl[ying] strict scrutiny with little discussion of the magnitude of the burdens imposed," but found the prior precedent distinguishable in any case, because the Oklahoma statute did not require Independents to affiliate with a party to vote in its primary.[27] Finding only "minor barriers between voter and party," the Court chose to apply a standard from *Timmons v. Twin Cities Area New Party*, that "a State's important regulatory interests will usually be enough to justify reasonable, nondiscriminatory restrictions."[28]

The Court then found several of Oklahoma's justifications sufficiently important to justify the restrictions on parties and voters. First, the majority agreed that Oklahoma had an important interest in avoiding the confusion that might result if candidates were nominated who did not share the ideology that voters would expect from a particular party.[29] Second, the

22. *Id.* at 1059 (quoting *Jones*, 530 U.S. at 584).

23. *Id.* at 1060. Indeed, the court pointed out, in neither of two states that grant parties the right to open their primaries has there been any evidence of increased instability in the system. *Id.*

24. *Id.* at 1061.

25. Justice Thomas delivered the opinion of the Court except as to Part II.A. The Chief Justice and Justices Scalia and Kennedy joined the opinion in its entirety. Justices O'Connor and Breyer joined in the opinion except as to Part II.A.

26. The discussion principally focused on *Timmons v. Twin Cities Area New Party*, 520 U.S. 351 (1997), where the Court upheld the constitutionality of a law forbidding so-called "fusion" candidacies, where multiple parties would endorse the same candidate for office. Clingman v. Beaver, 125 S. Ct. 2029, 2036–37 (2005).

27. *Clingman*, 125 S. Ct. at 2038.

28. *Id.* at 2039 (quoting *Timmons*, 520 U.S. at 358).

29. *Id.*

Court noted that allowing open primaries "would undermine the crucial role of political parties in the primary process."[30] The Court further found that Oklahoma had an interest in "aid[ing] in parties' electioneering and party-building efforts," which "depend in large part on accurate voter registration rolls."[31] Finally, the six Justices in the majority viewed the discouragement of raiding as an interest to be counted in favor of the challenged statutes' constitutionality.[32]

The Court refused to consider whether Oklahoma's election laws, taken as a whole, created an unconstitutional burden on the LPO or voters, because the argument had not been raised below.[33] Justices O'Connor and Kennedy, however, expressed some concern in their concurrence that the "cumulative effects of the State's overall scheme governing primary elections" might be too much for the state interests involved to bear.[34] Specifically, they worried that a major-party voter would have to disaffiliate from his party before even knowing if the LPO would be recognized by the state for that election cycle to be eligible to vote in the Libertarian primary.[35]

In dissent, Justice Stevens argued that the burden on the plaintiffs' associational rights was substantial and that all of the interests Oklahoma advanced were "either irrelevant or insignificant," if not "[il]legitimate."[36] Unlike circumstances in cases

---

30. *Id.* at 2040. It is not clear how the majority would respond to the example of Utah and Alaska given by the Tenth Circuit; presumably politicians from the major parties would have terminated the party-option open primaries if those parties' vitality was being threatened. *See id.* at 2044 (O'Connor, J., concurring) (noting that "the State is itself controlled by the political party or parties in power, which presumably have an incentive to shape the rules of the electoral game to their own benefit"). Of course, the LPO's proposed course of action may lead to its own implosion, but the state does not have a sufficient interest in protecting the LPO from its own "self-destructive acts." *See* Eu v. S.F. County Democratic Cent. Comm., 489 U.S. 214, 227 (1989).

31. *Clingman*, 125 S. Ct. at 2040.

32. *Id.*

33. *Id.* at 2041.

34. *Id.* at 2042 (O'Connor, J., concurring).

35. *Id.* at 2046–47. Because both major parties held closed primaries, the hypothetical voter would be relinquishing his right to vote in the major-party primary even if registering as an Independent. *Id.* at 2047. Indeed, it is possible that by disaffiliating, a voter might be unable to vote in any primary, especially if no parties requested semi-closed primaries. *Cf.* Tashjian v. Republican Party of Conn., 479 U.S. 208 (1986) (declaring closed primaries unconstitutional even when voters could change party allegiance the day before the primary).

36. *See Clingman*, 125 S. Ct. at 2048, 2051 (Stevens, J., dissenting).

such as *Jones*, the interests of the voters and of the party coincided in the case at bar, making for "a decision [that] should be easy."[37] It is not entirely clear what standard of review Justice Stevens would propose. On the one hand, he speaks of requiring a showing of a "state interest of overriding importance,"[38] which implies a level of scrutiny more searching than the majority's requirement of "important" interests. Yet he never explicitly writes in terms of strict scrutiny, and, in the end, he concludes that "the balance of interests clearly favors the LPO."[39]

Justice Stevens's criticism of the majority is most pronounced when he accuses it of considering interests that are "not legitimate," namely, interests "in manipulating the outcome of elections, in protecting the major parties from competition, [and] in stunting the growth of new parties."[40] Like the trial and intermediate appellate courts, Justice Stevens and the other dissenters agree that state concerns about raiding are paternalistic and intrude upon affairs constitutionally left to the party's judgment.[41] The dissent also notes the district court's factual finding that voter confusion was unlikely.[42] Rejecting the goal of ensuring "the political orthodoxy of party members simply for the convenience of those parties" as a legitimate interest, the dissent saw the "State's interest in properly classifying candidates and voters" as fairly minimal.[43] Finally, the minority viewed the interest in protecting the major parties' "electioneering and party-building efforts" as not substantial, being directed only at saving the parties from "administrative inconvenience."[44] Even that "interest," according to the dissent, was masquerad-

37. *Id.* at 2048.

38. *Id.* at 2049.

39. *Id.* at 2054. Notably, Justice Stevens does not quibble with Justice Thomas's assertion that the burden imposed on the LPO and voters is less severe than the burden on parties and candidates interested in fusion candidacies in *Timmons*, although one might expect him to if he were arguing for strict scrutiny. Instead, Justice Stevens contents himself with lamenting the "misguided" nature of the *Timmons* decision. *Id.*

40. *Id.* at 2048.

41. *See Clingman*, 125 S. Ct. at 2051–53.

42. *Id.* at 2052.

43. *Id.* Even assuming that crossover voters are less committed to their parties of registration, Justice Stevens argues that "the State does not have a valid interest in defining what it means to be a Republican or a Democrat." *Id.*

44. *Id.* at 2052–53.

ing as the "impermissible" state interest of dissuading major-party voters from participating more actively in the LPO.[45]

This Note will not challenge Justice Thomas's refusal to apply strict scrutiny, but will advance the case that the majority opinion's scrutiny of the challenged legislation was not searching enough. It will then explain why, under a regime of intermediate scrutiny, the legislation might not have been upheld. Finally, in light of Justice O'Connor's wary concurrence, it will discuss the future of semi-closed primary laws, cautioning that all is not well in semi-closed lands.

Despite the Court's frequent explanation that it is using a specific test to determine a law's constitutionality, it is probably better to conceptualize the different standards the Court employs as existing on a continuum.[46] Given this understanding, the Court's review of the Oklahoma statutes should not be located near the deferential pole on the continuum. First, applying more deferential review in this case would be hard to explain given the application of strict scrutiny in *Tashjian*. Admittedly, the Oklahoma statutes are somewhat less restrictive in that they do not require affiliation with the party to vote, only public disaffiliation with whichever party with which the voter had been registered. Yet the Court would be ill-advised to draw an arbitrary line between "severe" restrictions on the ballot, which would be subject to strict scrutiny, and slightly less restrictive regulations, which would be subject to minimal scrutiny. Such an approach more closely resembles empty for-

---

45. *See id.* at 2053.

46. *See, e.g.*, ERWIN CHEMERINSKY, CONSTITUTIONAL LAW: PRINCIPLES AND POLICIES § 6.5 (2d ed. 2002) (noting normative criticism by Justices of the levels of scrutiny regime); Jeffrey M. Shaman, *Cracks in the Structure: The Coming Breakdown of the Levels of Scrutiny*, 45 OHIO ST. L.J. 161, 183 (1984) (arguing that "the multi-level system [of scrutiny analysis] has reached a point of substantial disarray [and] deviation from the system has become so common as to render constitutional adjudication a desultory affair"). Some language in prior election law cases suggests that the Court operates on a continuum. *See, e.g.*, Timmons v. Twin Cities Area New Party, 520 U.S. 351, 358 (1997) ("When deciding whether a state election law violates First and Fourteenth Amendment associational rights, we weigh the character and magnitude of the burden the State's rule imposes on those rights against the interests the State contends justify that burden, and consider the extent to which the State's concerns make the burden necessary.") (citations and internal quotation marks omitted); Burdick v. Takushi, 504 U.S. 428, 434 (1992) (describing the Court's inquiry as "flexible"); Anderson v. Celebrezze, 460 U.S. 780, 789 (1983) (describing the Court's approach to the adjudication of claims of denial of the right to vote).

malism than the judging of factually complex cases or controversies.[47]

The more logical approach is to reduce the level of scrutiny in a gradual fashion as the level of infringement on the fundamental right is reduced. This is exactly what the Court has done with other fundamental rights. For instance, hampering the First Amendment right to free speech generally requires a "compelling interest" wedded to the "least restrictive means" of achieving that interest.[48] In contrast, when a reasonable and content-neutral regulation of "time, place, and manner" is issued, it only needs to meet an intermediate level of scrutiny.[49] *Clingman* certainly stands in relation to *Tashjian* in the same way that cases like *Ward v. Rock Against Racism*[50] stand against free speech cases applying strict scrutiny.

A more searching scrutiny is also appropriate in cases where semi-closed primaries are evaluated because of the importance that primary elections have in the process of determining who shall govern. The right to vote was enshrined in the ideology of the American Revolution[51] and has been the subject of several amendments to the Constitution.[52] And, as Justice Stevens notes, the primary election has increasingly become more important than the general election held a few months afterward.[53]

Finally, courts must apply more heightened levels of scrutiny in cases where even facially neutral regulations favor major

---

47. One might argue that a binary test serves the salutary purpose of cabining judicial discretion. This might well be true if the test for determining severity were more objective. With the definition of a "severe" violation open to interpretation, judges are allowed to treat fairly similar cases near the line of severity quite differently. In election law cases, there is "no substitute for the hard judgments that must be made." Storer v. Brown, 415 U.S. 724, 730 (1974). Indeed, to the extent that judicial decision making is cabined by precedent, a continuum approach may be preferable because it is more difficult to authorize a large break from a prior decision with similar facts.

48. *See, e.g.*, Sable Commc'ns of Cal., Inc. v. F.C.C., 492 U.S. 115, 126 (1989).

49. *See* Ward v. Rock Against Racism, 491 U.S. 781, 798–99 (1989) (holding that restrictions must be "narrowly tailored" and advance "substantial" government interests).

50. 491 U.S. 781 (1989).

51. Of course, the right to vote was not understood to extend to all groups at the time of the Founding, such as women, slaves, and those not holding property.

52. *See* U.S. CONST. amends. XV, XIX, XXVI.

53. *See* Clingman v. Beaver, 125 S. Ct. 2029, 2049 n.2 (2005) (Stevens, J., dissenting).

parties because (as Justice O'Connor hints in her concurrence)[54] those who currently run a state have an interest in keeping power.[55] Even when upholding regulations, the Court has noted that states cannot "completely insulate" the major parties from outside influences.[56] If the courts were to uncritically accept rationales given by the major political parties, speaking through the state, then meaningful electoral change would be impossible.

Given an intermediate standard of review, the Court should have demanded a stronger factual basis for the parade of horribles that Oklahoma's election board used as its defense. Justice Souter's concurrence and dissent in *City of Erie v. Pap's A.M.* compiles several key statements the Court has made about the evidentiary quality required before First Amendment rights can be overridden, even under intermediate levels of scrutiny.[57] For example, the Court has "never accepted mere conjecture as adequate to carry a First Amendment burden"; instead, it looks in part to the "evidence . . . cited by the lower courts" to determine whether satisfactory proof has been made.[58] In the trial court's decision, continual reference was

---

54. *Id.* at 2044 (O'Connor, J., concurring).

55. There is a growing literature in election law emphasizing the need to prevent "partisan lockups" of the electoral system. *See generally* Samuel Issacharoff & Richard H. Pildes, *Politics as Markets: Partisan Lockups of the Democratic Process*, 50 STAN. L. REV. 643 (1998). This article critiques the individual-rights framework used by the Court (and in this Note), *see id.* at 645, and instead argues that courts should take a more structural view to "ensure an appropriately competitive partisan environment." *Id.* at 648. The Court's embrace and reaffirmation of the traditional individual-rights framework in *Clingman* suggests that this alternative perspective is not likely to carry the day. *But see* Heather Gerken, *Lost in the Political Thicket: The Court, Election Law, and the Doctrinal Interregnum*, 153 U. PA. L. REV. 503, 539 (2004) (arguing that the Supreme Court is "casting about for a new strategy for deciding election law cases," which may well be the adoption of a more structural approach).

56. Timmons v. Twin Cities Area New Party, 520 U.S. 351, 367 (1997).

57. City of Erie v. Pap's A.M., 529 U.S. 277, 310–18 (2000) (Souter, J., concurring in part and dissenting in part). Of course, to the extent that *Clingman* is about voting rather than associational rights, the analogy to First Amendment cases is not, strictly speaking, appropriate. However, voting, like the First Amendment rights, is among the most fundamental rights guaranteed under the Constitution. *See* Reynolds v. Sims, 377 U.S. 533, 555 (1964) (proclaiming that "[t]he right to vote freely for the candidate of one's choice is of the essence of a democratic society, and any restrictions on that right strike at the heart of representative government").

58. *City of Erie*, 529 U.S. at 312 (quoting Nixon v. Shrink Mo. Gov't PAC, 528 U.S. 377, 392 (2000)).

made to how "speculative" or conjectural the effects claimed by the state of Oklahoma were.[59] The Tenth Circuit noted that the experience of states that allowed parties to elect open primaries was, indeed, inconsistent with Oklahoma's claims.[60] Yet the Court accepted rationales that the district and intermediate appellate court had found factually lacking, without explaining why the lower courts' conclusions were wrong or even what facts supported the state's contentions. To allow speculation to trump two of the most fundamental rights in the panoply of rights that Americans enjoy—free association and voting—is to do a disservice to the Constitution.

Second, the Court failed to consider alternatives that may have been just as effective in curtailing the evils it found in the LPO's request for an open primary. For instance, several alternate means could be devised for meeting the state's interest that political party membership lists be up to date.[61] Although less exacting forms of scrutiny do not require as strong a nexus as strict scrutiny, there still must be substantial congruence between the means employed and the ends sought.[62] There might not be an easy alternative in the present case if one accepts all the interests credited by the majority, but the lack of a search does not speak well to the Court's respect for voting and associational rights.[63]

---

59. *See, e.g.*, Beaver v. Clingman, No. CIV-00-1071-F, 2003 WL 745562, at *8 n.5 (W.D. Okla. Jan. 24, 2003).

60. *See supra* note 23 and accompanying text.

61. For instance, state law might direct that a voter is deemed to resign from her political party upon casting a ballot in another party's primary. Or, that fact might simply be reported to the political party, which could then choose to disassociate the "disloyal" voter. Of course, it will not always be possible to find alternatives that essentially satisfy the both claimed governmental interest as well as the challenged provision and do not interfere with other important state objectives. The governmental party should not be required to rule out all plausible alternatives, but the existence of a readily available and less restrictive alternative should cut against a finding of constitutionality.

62. *See* United States v. O'Brien, 391 U.S. 367, 377 (1968) (including in the eponymous test the requirement that a "restriction on alleged First Amendment freedoms is no greater than is essential to the furtherance of that interest"); *see also* Anderson v. Celebrezze, 460 U.S. 780, 789 (1983) (observing that a court "must consider the extent to which those interests make it *necessary* to burden the plaintiff's rights") (emphasis added).

63. This is especially true when one considers that a majority of the Justices found the restrictions imposed by the legislation to be fairly substantial. *See* Clingman v. Beaver, 125 S. Ct. 2029, 2041–44 (2005) (O'Connor, J., concurring); *id.* at 2048–51 (Stevens, J., dissenting).

Advocates of semi-closed primaries should not breathe easy. Although *Clingman* settles the constitutionality of semi-closed primaries in the abstract, a question explicitly left open in *Tashjian*, it does not speak to the details that are likely to determine whether a given legislative scheme, taken as a whole, passes constitutional muster.[64]

To some extent, designing a system that will meet a properly conducted test is an exercise in paradox. A scheme like Oklahoma's would doubtless be more effective in minimizing raiding, swamping, and the like, but might (if the electoral regime as a whole were challenged) fall as unconstitutional. On the other hand, a more minimal scheme—for example, one that allowed voters to disaffiliate with their old party the day before the election—would be less burdensome, but might not substantially advance state interests.[65] It is very probable that a constitutionally satisfactory scheme exists under *Clingman*, but legislatures may need to use more "narrowly tailored"[66] means to bring about their desired results.

Legislatures might choose, in the wake of judicial preclusion of more restrictive semi-closed primary schemes, to place the responsibility for policing the loyalty of political party members where it belongs: with the parties themselves. The rights of a political party clearly include the "freedom to identify the people who constitute the association and to limit the association to those people only."[67] States could authorize parties to exclude persons from voting in their primaries who had been "disloyal" to the party (that is, had voted in another party's

64. It is easy to gloss over Part III of the majority opinion as a technical lecture on how to draft pleadings at the trial stage. However, the Court's refusal to consider the nuances of Oklahoma electoral law as a whole means that only the basic idea of a semi-closed primary system was examined. Justices O'Connor and Breyer appear to go out of their way to comment on problems in sections of the Oklahoma statutes that they agree are not properly before the Court. *See id.* at 2042–47. This can reasonably be read as a warning to states that would otherwise see the decision as placing the Court's imprimatur on electoral schemes as restrictive as Oklahoma's.

65. Although some cases applying intermediate scrutiny do not speak in terms of balancing interests, a balancing may well be implicit in the inner workings of the tests, given the vagueness of the words used to constitute the test. In any case, both the majority, *see id.* at 2035, and the dissent, *see id.* at 2054, seem to agree that balancing concerns are implicated in the question of semi-closed primaries.

66. Clark v. Cmty. for Creative Non-Violence, 468 U.S. 288, 293 (1984).

67. Democratic Party of the U.S. v. Wisconsin *ex rel.* La Follette, 450 U.S. 107, 122 (1981).

primary while enrolled as a member of the party) within a specific period of time.[68] Whether states themselves could do this is questionable;[69] but a political party's right to exclude is broader than a state's. A political party generally owes no constitutional duties to the voters; however, voters are owed constitutional duties by the state.[70] A scheme wherein political parties police member loyalty would have positive policy implications. It would be more transparent and improve accountability because the force truly responsible for a voter's inability to participate in her old party's primary would be apparent.

The sky will not fall if the Court decides to take a harder look at semi-closed primary schemes in the context of the electoral laws in which they are placed. By doing so, the Court can defend the historic place that American political experience grants to the important rights of voting and free political association.

M. Jason Scoggins

68. Whether parties could do this on their own, based on their First Amendment free association rights, is beyond the scope of this Note.

69. *See* Kusper v. Pontikes, 414 U.S. 51 (1973) (declaring unconstitutional a law barring a voter from participating in a party's primary if he had participated in another party's primary within a period of time almost two years before the election).

70. *See generally* Cal. Democratic Party v. Jones, 530 U.S. 567 (2000).

PAYING FOR THE SINS OF THEIR USERS: LIABILITY AND GROWING UNCERTAINTY IN A DIGITAL AGE: *Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*, 125 S. Ct. 2764 (2005)

Although space may represent the final frontier for man, the realm of cyberspace embodies an equally untamed wilderness. With more similarities to the Wild West than to a civilized modern society,[1] the Internet defies previously established principles of law and order, shunning the constructs of reality in favor of a virtual existence where there are few recognized legal conventions. Since the Supreme Court's decision in *Sony Corp. of America v. Universal City Studios Inc.*,[2] copyright holders and developers of emerging technology alike have struggled with the laws regarding the liability of distributors whose products are used by third parties for the purposes of copyright infringement.[3] The continuing creation and distribution of new software through the use of the Internet, however, has led the courts deeper into a legal quagmire. Gone are the days of Betamax[4] and cassette tapes; high-quality pirated music and video are now available for free download through online services. In

---

1. For a discussion of cyberspace and the Wild West metaphor, see Shamoil Shipchandler, Note, *The Wild Wild Web: Nonregulation as the Answer to the Regulatory Question*, 33 CORNELL INT'L L.J. 435, 436 (2000) ("Today's Wild West is the Internet."); Raymond Ku, Foreword, *A Brave New Cyberworld?*, 22 T. JEFFERSON L. REV. 125, 125–26 (2000). *But see* Alfred C. Yen, *Western Frontier or Feudal Society?: Metaphors and Perceptions of Cyberspace*, 17 BERKELEY TECH. L.J. 1207, 1207–08 (2002) (concluding that the Internet is more comparable to the structure of a medieval feudal society than that of the lawless Wild West).

2. 464 U.S. 417 (1984).

3. The modern-day legal dilemmas regarding copyright infringement and online music trace back to the Court's decision in *Sony*. Both developers and copyright holders assert that the decision supports their arguments regarding online piracy. For a discussion of the many facets of *Sony* and the growing debate, see Brandon Michael Francavillo, *Pretzel Logic: The Ninth Circuit's Approach to Contributory Copyright Infringement Mandates that the Supreme Court Revisit* Sony, 53 CATH. U. L. REV. 855, 885 (2004). *See also infra* note 25 and accompanying text.

4. Betamax was a propriety video format, similar to what is now known as VHS, developed by Sony in the 1980s and the subject of the Court's *Sony* decision. *See Sony*, 464 U.S. at 422–23.

the face of this emerging technology, courts and legal scholars are actively debating the merits of holding liable the distributors of products that enable the infringement of copyrights.[5]

Last Term, in *Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*,[6] the Supreme Court held that "one who distributes a device with the object of promoting its use to infringe copyrights, as shown by clear expression or other affirmative steps taken to foster infringement, is liable for the resulting [third-party] acts of [copyright] infringement."[7] Despite the Court's attempt to develop new rules for a digital age, the numerous opinions by the Justices in *Grokster* left a murky standard with little application outside the particular facts of that case. Presenting a unanimous opinion solely on the facts of the case, the Justices then splintered on the application of the law, set forth a vague holding without practical tools to combat the larger epidemic of online piracy, while simultaneously presenting conflicting applications of the previous standard for distributor liability under *Sony*. The Court's attempt to resolve the issue of copyright infringement runs the risk of overstepping the judicial branch's power by regulating a legislative policy matter better left to the discretion of Congress.

In the fall of 2000, Metro-Goldwyn-Mayer Studios (MGM) filed suit against two distributors of free online software, Grokster and Streamcast Networks, alleging that the defendants knowingly and intentionally circulated their software in order to enable users to reproduce and distribute copyrighted works in violation of the Copyright Act.[8] Moreover, MGM maintained that the software provided by the two companies served no substantial noninfringing purpose.[9] The software allowed computer users, through the use of peer-to-peer networks,[10] to

5. Debate within the courts over the development of online files has raged for years. Although the exponential growth of online music piracy is a relatively recent development, a substantial number of cases have attempted to deal with the difficulties inherent in copyright infringement. *See, e.g.*, *In re* Aimster Copyright Litigation, 334 F.3d 643, 653 (7th Cir. 2003); A&M Records Inc. v. Napster Inc., 239 F.3d 1004, 1029 (9th Cir. 2001) (holding the distributor of online music could be held liable for the actions of users where the distributor ran a central server and was informed of users' copyright infringement).

6. 125 S. Ct. 2764 (2005).

7. *Id.* at 2780.

8. 17 U.S.C. § 101. *See Grokster*, 125 S. Ct. at 2771.

9. *Grokster*, 125 S. Ct. at 2774.

10. The use of "peer-to-peer" networks presents a critical distinction in this case from previous internet file sharing cases. Peer-to-peer networks allow users to

connect with other members and download both copyrighted and noncopyrighted music and video. In its lawsuit, MGM sought damages arising from the copyright infringement carried out by users and an injunction to prevent the continued distribution of the software.[11]

Several facts presented in the case were undisputed by either party. First, users on the network were downloading copyrighted music and video files that were made available by the software.[12] Second, users downloading the copyrighted music and video were engaged in illegal activities in violation of the Copyright Act.[13] Respondents argued that, under the doctrine adopted by the Court in *Sony*, the product was "capable of substantial noninfringing use" and that therefore they could not be held liable for the actions of their users.[14]

The case was originally brought before the federal District Court for the Central District of California, which granted summary judgment in favor of Grokster and StreamCast as to liability arising from the distribution of their software.[15] The court predicated its decision on the distributors' lack of actual

---

communicate directly with each other, as opposed to communicating through a central server. *See id.* at 2770–71. As such, there is no single point through which all the information passes, and, according to Grokster and StreamCast, such feature renders it impossible for the companies to track file exchanges. *See id.* Previous cases, including *Napster* and *Aimster*, involved the use of file-sharing programs that connected users through central data servers, through which users would find and download files from other users. *See Napster*, 239 F.3d at 1011–13.

11. *Grokster*, 125 S. Ct. at 2771.

12. Although it was undisputed that users were downloading copyrighted materials, the number of copyrighted works available on the networks was a point of contention, as was the availability of works that were not subject to any copyright protection. *Id.* at 2771–72. Indeed, respondents did not dispute that the primary use of their software by its users was the illegal downloading of copyrighted works. *See id.* at 2772.

13. *Id.* at 2772.

14. *See id.* The debate in *Sony* centered on the use of Betamax video recorders that allowed users to record programs and build video libraries: activities possibly in violation of the Copyright Act. Furthermore, there was strong evidence that users would record copyrighted works in order to possess those works free of charge. *Sony*, 464 U.S. at 423–24. The Court, however, held that a distributor of a product capable of substantial noninfringing use (there, "time-shifting" by recording a program for the limited purpose of viewing it at a different time) may not be held vicariously liable for the illegal actions of its users. *Id.* at 456.

15. MGM Studios Inc. v. Grokster, Ltd., 259 F. Supp. 2d 1029, 1046 (C.D. Cal. 2003).

knowledge about specific acts of infringement and the product's capability of substantial legitimate uses.[16]

The Court of Appeals for the Ninth Circuit affirmed the decision of the District Court, concluding that a defendant was liable for contributory infringement only when it had knowledge of the infringement committed by the users of the product and materially contributed to the infringement.[17] The Ninth Circuit determined that the software was capable of substantial non-infringing uses and found that liability under the *Sony* doctrine did not exist in this case.[18] Additionally, the companies were determined not to be responsible for the actions of their users because of the decentralized structure of the software.[19] Accordingly, the Ninth Circuit found that Grokster and StreamCast did not materially contribute to their users' infringement because they took no action in the search, retrieval, or storage of infringing files.[20] The only role taken by Grokster and StreamCast, according to the court, was providing the free software, while the users committed the illegal acts.[21]

MGM then appealed to the Supreme Court, presenting a novel legal problem to the Court: to determine "under what circumstances the distributor of a product capable of both lawful and unlawful use is liable for the acts of copyright infringement by third parties using the product."[22]

Justice Souter authored the opinion of the Court,[23] which held that "one who distributes a device with the object of promoting its use to infringe copyright, as shown by clear expression or other affirmative steps taken to foster infringement, is liable for the resulting acts of infringement by third party us-

16. *See id.* at 1045–46.

17. *See* MGM Studios Inc. v. Grokster, Ltd., 380 F.3d 1154, 1157 (9th Cir. 2004).

18. *See id.* at 1161.

19. *Id.* at 1164–66.

20. *See id.* at 1165. The Ninth Circuit's opinion regarding this point again refers to the nature of the programs as peer-to-peer networks. *Id.* Peer-to-peer networks do not require the existence of a central server and thus require little involvement on the part of the distributor. Moreover, these programs do not allow the distributor to know the contents of the traded material available on the network. *Id.* This fact was critical to the determination that Grokster and StreamCast were not liable, according to the Ninth Circuit. *See id.* at 1164–66.

21. *See id.* at 1164–66.

22. MGM Studios Inc. v. Grokster, Ltd., 125 S. Ct. 2764, 2770 (2005).

23. *Id.* Although the opinion of the Court was unanimous, there were two concurring opinions that sought to clarify the application of *Sony*. *See infra* text accompanying notes 43–56.

ers."[24] The Court thus determined that both the District Court and the Ninth Circuit had misapplied the Court's *Sony* decision.[25] In other words, the lower courts had incorrectly determined that a producer cannot be held contributorily liable for the infringing use of a product that is capable of substantial noninfringing use even when an *actual purpose to cause infringing use* is demonstrated.[26]

Instead, the Court held that a lack of specific knowledge of infringement and a failure to act upon that lack of knowledge did not bar the application of other theories of secondary liability.[27] Justice Souter further explained that, under *Sony*, distributors of products may not be held liable under theories of contributory liability where the users are acting of their own accord; however, *Sony* did not foreclose other avenues to liability in situations such as were presented by Grokster and StreamCast.[28] Justice Souter stated that nothing in *Sony* requires courts to ignore evidence of intent to promote infringement, if such evidence exists.[29] *Sony* was not meant, contrary to the opinions of the lower courts, to exclude rules of fault-based liability that arise under common law.[30] Based on this finding, Justice Souter determined that the *Sony* rule regarding third-party liability would not apply where evidence goes beyond a product's characteristics or the knowledge that a product may be put to infringing uses, and instead demonstrates statements or actions directed to promoting infringement.[31]

Indeed, the Court held that both of the parties were incorrect in their arguments applying the *Sony* test to the present facts.[32]

---

24. *Grokster*, 125 S. Ct. at 2770.

25. *Id.* at 2778–79. The Court seemed to explain that this misapplication by the lower courts did not necessarily arise out of a misunderstanding of the test in *Sony*; rather, it arose from the misperception that other grounds for liability did not pertain to the case. *See id.*

26. *Id.*

27. *Id.* at 2779.

28. *Id.*

29. *Grokster*, 125 S. Ct. at 2779.

30. *Id.*

31. *Id.*

32. *Id.* at 2778. The parties argued the resolution of the case on the basis of the Court's previously established standard under *Sony* and its third-party liability principles. *Id.* Other than noting the lack of applicability of *Sony* in this matter, Justice Souter did not engage in further analysis regarding the previous standard set forth by the Court.

Justice Souter found that the correct legal principle on which to rule is not the contributory liability of Grokster and StreamCast under *Sony*, but rather the common-law concept of "inducement" infringement.[33] The existence of evidence demonstrating a purposeful inducement on the part of the distributor will overcome the reluctance of the courts to hold a third party liable where the defendant distributes a product suitable for some lawful, noninfringing use.[34]

Justice Souter noted that three aspects of the evidence in the trial record demonstrated that the respondents induced the infringement of copyrights and acted with an unlawful objective: (1) marketing the illicit use of their programs; (2) failing to attempt to curb the illegal use of the products; and (3) making a profit as a result of the illegal activities by users of the programs.[35] With regard to each of the factors, the Court found specific qualities of the distributors particularly noteworthy. First, both companies exhibited a desire to satisfy the demand for illegal downloading of copyrighted materials by expressly attempting to appeal to previous users of Napster.[36] Second, neither party installed or attempted to create any filter in order to prevent the infringing activities undertaken by the users of their product.[37] In the opinion of the Court, this elucidated the

---

33. *Id.* at 2779. The Court explained, "'Inducement' infringement refers to situations where evidence demonstrates more than mere knowledge, to the point where there are statements or actions promoting infringement. Inducement infringement often occurs where one 'entic[es] or persuad[es] another' to engage in infringing activities." *Id.* (citing BLACK'S LAW DICTIONARY 790 (8th ed. 2004)). Common law has long recognized the existence of liability on the basis of inducement infringement. *See, e.g.*, Kalem Co. v. Harper Bros., 222 U.S. 55, 62–63 (1911) *cited in Grokster*, 125 S. Ct. at 2779.

34. *Grokster*, 125 S. Ct. at 2779 (citing Water Techs. v. Calco, Ltd., 850 F.2d 660, 668 (Fed. Cir. 1988) (holding that inducement liability occurs where one "actively and knowingly aid[s] and abet[s] another's direct infringement") (emphasis omitted)).

35. *See id.* at 2781.

36. *Id.* Because Napster was a program whose users engaged in copious amounts of known copyright infringement, the Court derived a great deal of significance from both Grokster's and StreamCast's attempts to divert old Napster customers to its own programs. It also appears from the record that the companies would respond to emails that inquired about how to download copyrighted materials that were available through use of their programs. *See id.* at 2772.

37. *Id.* at 2781. The district court and the Ninth Circuit had determined that the defendants' failure to develop filtering tools was not essential to the holding, as that would place an independent duty on the defendants to monitor users' activity. *See* MGM Studios Inc. v. Grokster, Ltd., 380 F.3d 1154, 1164–66 (9th Cir. 2004). Instead, the Court notes that StreamCast even took the step of rejecting another company's offer to monitor infringement and blocked the IP addresses of

respondents' desire to facilitate its users' infringement.[38] Third, the revenue the companies derived from selling advertising space demonstrated that the respondents had a desire to increase the infringing use of their products.[39]

The Court, therefore, examined the totality of the circumstances, viewing the record in the context of all the evidence presented and granting no single factor more weight than the others.[40] The tension of the case, according to Justice Souter, was balancing the value of copyright-protected creative pursuits and the value of innovation in communications and technology.[41] The greater the emphasis on protection of copyright holders, the more the courts will stifle the development of new technology; the greater the emphasis on permitting the free development of new products without fear of liability, the less protection the courts provide to those who hold copyrights.[42]

Justice Ginsburg filed a concurring opinion.[43] She agreed with the holding of the case regarding third-party liability on the theory of inducement infringement, but wrote separately to express her opinion on the application of *Sony*'s principles to the present case.[44] The concurring Justices maintained that, even under the standard set forth in *Sony*, there is "at least a 'genuine issue as to [a] material fact' . . . on the liability of Grokster or

---

entities that it felt were engaged in any such monitoring. *Grokster*, 125 S. Ct. at 2774.

38. *Grokster*, 125 S. Ct. at 2781. The Court also noted that, in the absence of other evidence, the lack of affirmative action by the parties would not be enough to find contributory infringement if the device otherwise was capable of substantial noninfringing uses. *Id.* at 2781 n.12.

39. *See id.* at 2781–82.

40. *See id.* Although the Court does not explicitly mention the use of a "totality of the circumstances" test, the Court's mention that certain factors alone would not serve as dispositive grounds on which to rule demonstrates that the Court is not creating a bright line rule, but rather a standard weighed through various factors.

41. *Id.* at 2775–76.

42. *See id.* While noting the tension between the two policy objectives, Justice Souter also observes that the two goals are not always mutually exclusive. *See id.* at 2776 n.8. Ironically, he mentions, even those who develop new technologies that allow for the infringement of copyrights of others may desire to protect their own intellectual property. *Id.* (citing Tim Wu, *When Code Isn't Law*, 89 VA. L. REV. 679, 750 (2003)).

43. *Grokster*, 125 S. Ct. at 2783. Chief Justice Rehnquist and Justice Kennedy joined the opinion.

44. *Id.*

StreamCast" on a theory of contributory infringement.[45] With both parties, there existed a question of whether the programs are capable of substantial noninfringing use.[46] Most evidence regarding the potential noninfringing use of the programs was anecdotal or generalized.[47] As such, it was not possible for the lower courts to tell whether, at the time of summary judgment, the programs were capable of a substantial noninfringing use.[48] To Justice Ginsburg, the test demanding that a product be "capable" of substantial noninfringing use necessarily implies that *actual* noninfringing use be shown.[49] Justice Ginsburg thus concluded that, if the case is not resolved on remand by summary judgment in favor of MGM on the basis of inducement infringement by Grokster and StreamCast, she would have the lower court overturn the summary judgment in favor of Grokster and StreamCast on the basis that it does not meet the requirement set forth by the Court in *Sony*.[50]

Justice Breyer wrote a separate concurrence,[51] agreeing with the Court that the distributor of a product that may be used for both legal and extralegal purposes may be held liable for the actions of third-party users where the distributor actively induces the infringement.[52] He further agreed that revisiting *Sony* was not necessary for the holding of the case, but noted that he would consider the question as a result of Justice Ginsburg's concurring opinion.[53] Justice Breyer argued that, considering the record at the time of summary judgment, there was an adequate basis for granting summary judgment and for the Ninth Circuit opinion upholding that decision.[54] According to Justice

---

45. *Id.*

46. *See id.* at 2785.

47. *Grokster*, 125 S. Ct. at 2785.

48. *See id.* at 2786.

49. Supporting this contention, Justice Ginsburg noted that, according to the trial record, there was little evidence that any noninfringing use was occurring. *Id.* The number of noninfringing users was dwarfed by the number of infringing users, and, at the time of summary judgment, there was no real evidence that the amount of noninfringing users would grow over time. *See id.*

50. *See id.* at 2786–87.

51. *Grokster*, 125 S. Ct. at 2787 (Breyer, J., concurring). Justices Stevens and O'Connor joined the opinion.

52. *Id.* at 2787.

53. *Id.*

54. *Id.* Again, it is important to note that Justice Breyer takes this issue under consideration in the theoretical absence of any evidence regarding active inducement on the part of Grokster and StreamCast.

Breyer, the purpose of the ruling in *Sony* was to "'strike a balance between a copyright holder's legitimate demand for effective . . . protection of the statutory monopoly, and the rights of others freely to engage in substantially unrelated areas of commerce.'"[55]

Comparing the capabilities of the Betamax recorder from *Sony* with the programs distributed by Grokster and StreamCast, Justice Breyer maintained that the programs in question are just as capable of substantial noninfringing use as was the Betamax.[56] Justice Breyer also argued that other, unknown noninfringing uses of peer-to-peer software may develop in the future, just as they developed with the Betamax and the VCR.[57] The inquiry, according to Justice Breyer, turns on analyzing the *Sony* standard itself to determine whether or not it still strikes the balance that it set out to strike.[58] As Justice Breyer interpreted the rule, *Sony* provides the balance necessary to permit the development of new technology while still protecting copyright interests.[59] According to Justice Breyer, *Sony*'s rule both provides entrepreneurs with assurance that they will not be held liable for the actions of their users as they bring new technologies to the market and protects copyright holders from devices without a redeeming legal value.[60] Justice Breyer argued

---

55. *Id.* (quoting *Sony*, 464 U.S. at 442).

56. *Id.* at 2787–89. Indeed, Justice Breyer compares the percentage of noninfringing use occurring within each of the devices. Although it was difficult to ascertain exact numbers about the noninfringing uses of Grokster and StreamCast, Justice Breyer estimated, on the basis of MGM's expert, that approximately ten percent of the files available on the networks were noninfringing. Comparatively, Justice Breyer noted that the Court found in *Sony* that approximately nine percent of the use of the Betamax recorder was noninfringing. *Id.* at 2788–89.

57. *Id.* at 2789–90. An important point discussed by Justice Breyer is the advent of new noninfringing uses of video recorders. Originally considered a threat to the studio industry, the VCR later served as the platform for a whole new industry of home video rental, creating a substantial legitimate use for the VCR. *Id.* The strength of this argument is demonstrated by similar occurrences with portable digital audio tape and portable digital audio players.

58. Justice Breyer's analysis consists of three questions: (1) has the *Sony* rule protected the development of new technology, (2) would a "modification or strict interpretation significantly weaken that protection," and (3) would strengthening the copyright benefits outweigh the interest in any such weakening. *Grokster*, 125 S. Ct. at 2791.

59. *See id.* at 2791–92.

60. *Id.* This is because the *Sony* rule is clear, strongly protects technology, is forward looking, and is mindful of the limitations on judges when dealing with matters of technology.

that adopting a modified standard, as proposed by Justice Ginsburg, would only chill the development of technology by requiring distributors to demonstrate the noninfringing use of their products with concrete evidence.[61] Finally, Justice Breyer returned to the language of *Sony* and noted that "[c]ourts are less well suited than Congress to the task of 'accomodat[ing] fully the varied permutations of competing interests that are inevitably implicated by such new technology.'"[62]

Although the Court's opinion was unanimous, the unique circumstances of the case and the conflicting applications of the *Sony* test demonstrate anything but cohesion, leaving developers and copyright holders unsure of the Court's likely disposition of future cases. Media and record companies immediately hailed the ruling of the Court as a substantial victory for copyright holders,[63] yet their jubilance may be short lived. A deeper examination of *Grokster* demonstrates that the ruling develops a situation in which similar cases may simmer in the lower appellate courts before the Supreme Court makes a decisive decision in the arena of vicarious liability for third-party users. The uncertain application of the Court's ruling on the facts of *Grokster* and the dicta contained in the Justices' concurring opinions may lead to the least-optimal scenario: simultaneously stifling development of emerging technology while accommodating large quantities of copyright infringement. In examining the Court's holding, it is necessary to consider the promise the case held for the resolution of many disputed cyberlaw questions, the effect the decision will have on technology, the effect on copyright protection, and the role of the judiciary in determining the laws of cyberspace.

For many, *Grokster* represented the hope that the Court would confront some of the most challenging legal issues of a new digital age.[64] With the rise of the Internet and new forms of

---

61. *Id.* at 2792–93.

62. *Id.* at 2796 (quoting *Sony*, 464 U.S. at 431).

63. *See, e.g.*, Tom Zeller, Jr., *The Imps of File Sharing May Lose in Court, but They are Winning in the Marketplace*, N.Y. TIMES, July 4, 2005, at C3 (quoting former head of the Recording Industry Association of America, "We won big. Unanimous."); *Court Sides with Music Industry*, FT. WORTH STAR-TELEGRAM, June 28, 2005, at C1; *Cf.* Jeef Leeds & Steve Lohr, *No Pot of Gold in Court Ruling for the Studios*, N.Y. TIMES, June 28, 2005, at C1; *Hollywood Wins Internet Piracy Battle*, CNN/MONEY, (June 27, 2005), http://money.cnn.com/2005/06/27/technology/grokster/index.htm.

64. Many legal scholars and technology experts believed that *Grokster* would resolve many of the lingering questions arising in a new digital age where there were no servers, but rather only individuals exchanging information with each

copyright infringement, the old standards set forth by the Court's decision in *Sony* were coming under intense pressure.[65] Accordingly, developers and studios alike demanded the resolution of these issues in a time of increasing copyright infringement.[66] *Grokster* did not provide that hoped-for resolution. While confronting the immediate problem presented by Grokster and StreamCast, the Court did little to resolve the question of liability in situations where no inducement on the part of the distributor is present.[67]

The impracticality of such a narrow holding is evident from the facts of the case. Both Grokster and StreamCast voluntarily sought to replace the void left by Napster's demise, despite Napster's well known reputation for allowing, maintaining, and perhaps even encouraging the illegal distribution of copyrighted material.[68] Internal documents demonstrated that the goal of the programs was to enable infringement by their users; these facts, the Court argued, constituted a clear case of inducement infringement.[69] Because of the narrow facts at issue in *Grokster*, any application to a different set of facts is tenuous at best; many other companies whose programs or creations allow such copyright infringement to occur will not make the mistake of having such clear evidence of inducement by their distributors.[70]

The difficulty for creators of new technology is not that inducement theory provides an unacceptable standard; rather, the difficulty lies in the vagaries of the Court's explanation of such a standard. When new companies seek to develop technology that may threaten the existence of the status quo, the developers may have to consider possible lawsuits stemming

---

other through a program provided by the distributor. *See, e.g.*, Michelle Kessler, *High Court Ruling on File Sharing Could Bring Clarity*, USA TODAY, June 27, 2005, at 04B.

65. *See generally* A&M Records Inc. v. Napster Inc., 239 F.3d 1004 (9th Cir. 2001); *In re* Aimster Copyright Litigation, 334 F.3d 643 (7th Cir. 2003).

66. Justice Souter stated that the amount of copyright infringement occurring on the Grokster and StreamCast networks was "staggering." *Grokster*, 125 S. Ct. at 2772.

67. *See id.* at 2782.

68. *See id.* at 2781–82.

69. *Id.* at 2781.

70. Justice Souter took careful note of internal documents that revealed the management of StreamCast and Grokster would respond to emails inquiring how to download copyrighted works. *Id.*

from the actions of third parties because of the uncertainty arising out of the Court's decision in *Grokster*.[71] Indeed, in oral arguments, using the example of a photocopier, Justices Breyer and Scalia pondered the question of how the Court's ruling would affect the inventor in his garage who does not yet know the effects of his technology.[72] The same analogy, however, could be applied to a situation of inducement infringement: the creator of a photocopier gives instructions on the operation of the machine, including instructions on how to make copies. At this point, has the creator induced users to infringe copyright by providing the mechanism, not monitoring users, and instructing them on the tools of infringement? The answer to this question remains unclear with the Court's opinion in *Grokster*. The paradox of development, of course, is that copyrights are as essential to fostering the creation of new technology and products as they are to the protection of established ideas.[73] The inherent difficulty with the Internet is that the demise of one company will often spur the creation of ten more to fill its gap.[74]

Because of the difficulty of applying the standard set forth by the Court outside of the facts of the case, lower courts will be forced to turn back to the *Sony* standard, still unsure of its true application in a digital age. With three Justices, led by Justice Ginsberg, arguing that cause for summary judgment did not exist on the basis of the facts; three Justices, led by Justice Breyer, stating that summary judgment under the *Sony* standard was correct; and three Justices agreeing with Justice Souter that a discussion of the *Sony* standard is not necessary to the holding of the case, the result is a legal quandary in which

---

71. For a discussion of the uncertainty that may arise as a result of the Court's decision, *see* Press Release, *Supreme Court Ruling Will Chill Technology Innovation, Electronic Frontier Foundation*, June 27, 2005, http://www.eff.org/news/archives/2005_06.php.

72. *See* Transcript of Oral Argument at 10–15, *Grokster*, 125 S. Ct. 2764 (2005) (No. 04-4800), http://www.supremecourtus.gov/oral_arguments/argument_transcripts/04-480.pdf.

73. *See generally* Note, *Copyright Law—Ninth Circuit Holds that Computer File Sharing Software Vendors Are Not Liable for Users' Copyright Infringement*, 118 HARV. L. REV. 1761, 1768 (2005).

74. Indeed, one need look no further than the case at hand to see the occurrence of this phenomenon. As noted earlier, the Ninth Circuit's decision in *Napster* only led to the new market dominance of Grokster and StreamCast. *See Grokster*, 125 S. Ct. at 2772–73.

the lower courts are left with little settled law and possibly even more uncertainty than previously existed.[75]

The concern expressed by the Justices over stifling technological innovation is not a new one.[76] Copyright holders are often hesitant to accept emerging technology that may or may not threaten the value of their copyrights. In the past, record and movie studios have fought similar battles with other technologies that were capable of infringing use. The most ubiquitous example is the development of the video cassette recorder and Betamax, as demonstrated by the long legal battles waged in *Sony*.[77] Although the VCR is now considered a typical household item,[78] the studios originally fought adamantly against the release of the device, as it could allow for massive amounts of copyright infringement.[79] Even with the loss suffered by the studio industry after the *Sony* battle, the war against infringement and technology was far from over. In 1986, with the introduction of Digital Audio Tapes, the studios again brought a lawsuit to prevent the distribution of these devices on the claim that they would permit copyright infringement.[80] In 1999, the record companies also brought suit against Diamond Multimedia, a company producing what, at that time,

---

75. Although the Court seemingly resolved the issue in the case, as it set out to do, by reaching a unanimous opinion as to the applicability of the inducement theories of liability, the various opinions simply demonstrate what appears to be a clear split within the Court on where the *Sony* test should stand.

76. *See generally* William Fisher & Christopher Yang, *Peer-to-Peer Copying*, Internet Law Program, The Berkman Center for Internet and Society at Harvard Law School (2003), http://cyber.law.harvard.edu/ilaw/stanford03/participants/p2p.html.

77. *See id.* As mentioned *supra*, the *Sony* case established the standards by which distributors could be held contributorily liable for the actions of third party users. *See generally* Sony Corp. v. Universal City Studios, 464 U.S. 416 (1984).

78. Although VCRs became nearly ubiquitous in the United States, they are now outsold by Digital Video Disc players and are becoming a device of the past. *See Buying Guide for Video Playback and Recording Equipment*, Consumer Guide Products, http://products.consumerguide.com/buyingguides/guide.epub?articleId=10072.

79. *See supra* note 14.

80. *See* Fisher & Yang, *supra* note 76. Digital Audio Tapes (DATs) presented a unique problem that the copyright holders had yet to face. DAT provides for the digital reproduction of copyrighted works, meaning that each copy will be exact, without any degradation in quality. *Id.* As such, this presented a more serious problem than the introduction of cassette tapes, which suffered the problem of poor reproduction quality. *Id.*

was a novel audio tool: a portable digital audio player.[81] As the studios and record companies overcame their fear of the technology, however, an entire digital music industry developed.[82]

There are no easy answers to the questions posed by the advent of online piracy, but the Court's solutions have had little effect on curbing continued copyright infringement.[83] The Court's difficulty in confronting this issue poses the question of whether the courts are adequately positioned to deal with such a question of policy. Justice Breyer noted in his concurring opinion that courts are not as well suited as Congress to taking on the policy concerns of the competing interests implicated by new technology.[84] Congress is in a better position than the courts to respond to the concerns of constituents and undertake the cost-benefit analysis required in such complex and competing matters. Congress has made attempts to tackle copyright infringement in a digital age, but these attempts have done little to deal with the problems of Internet piracy, perhaps explaining the expanding role of the courts in this area.[85] The courts, while acting with the best of intentions, have only managed to create chaos, upsetting both copyright holders and market innovators.

The Supreme Court's decision in *MGM v. Grokster* will not, unfortunately, change the status quo: confusion will reign for those who seek to obey the law, and anarchy in online digital piracy will continue unabated. The Court only delayed the inevitable resolution of this pressing issue; as new decisions percolate up from lower courts confused about the Court's

---

81. *See* Recording Indus. Ass'n of Am. v. Diamond Multimedia Sys., Inc., 180 F.3d 1072 (9th Cir. 1999).

82. The ubiquity of the iPod and the growing popularity of pay-as-you-go online music may demonstrate the growth of an industry the recording industry once feared.

83. BigChampagne, an Internet research firm that tracks usage of online file sharing, found that at any given time there are an average of 8.6 million simultaneous users swapping files on peer-to-peer networks (April 2005). This is up from 7.6 million in 2004. *See* Lisa DiCarlo, *File Size on P2P Nets Explodes*, Forbes.com, June 8, 2005, http://www.forbes.com/digitalentertainment/2005/06/08/p2p-limewire-napster-cx_ld_0608chart.html.

84. *See Grokster*, 125 S. Ct. at 2796 (Breyer, J., concurring).

85. *See generally* Audio Home Recording Act of 1992, 106 Stat. 4237 (1992); *Privacy and Piracy: The Paradox of Illegal File Sharing on Peer-to-Peer Networks and the Impact of Technology on the Entertainment Industry: Hearing Before the S. Comm. on Governmental Affairs*, 108th Cong., 2d Sess. (2003).

Copyright Act jurisprudence, the Court will undoubtedly be called on to return to its decisions in *Grokster* and *Sony*.[86]

John Lobato

86. Grokster reached a settlement with MGM and shut down its website. As part of the settlement, Grokster agreed to pay $50 million to put to rest the piracy claims. The agreement also requires the company to stop giving away its software. According to reports, Grokster will now seek to launch a legal, fee-based music service. The settlement does not affect the disposition of users already possessing the software, nor does it involve any other distributors of file-sharing software. *See* Gary Gentile, *Grokster to Shut Down Downloading Service*, ASSOCIATED PRESS, Nov. 8, 2005, http://hosted.ap.org/dynamic/stories/D/DOWNLOADING_MUSIC?SITE=AP&SECTION=HOME&TEMPLATE=DEFAULT&CTIME=2005-11-07-18-57-49.

CHASING FINALITY: FEDERAL COLLATERAL RELIEF IN THE WAKE OF *Pace v. DiGuglielmo*, 125 S. Ct. 1807 (2005)

I. INTRODUCTION

The Antiterrorism and Effective Death Penalty Act of 1996 (AEDPA) provides the first statute of limitations on the filing of federal habeas corpus petitions by state defendants.[1] Persons in custody pursuant to a state court judgment must seek federal habeas relief within one year of the latest of four dates.[2] Most commonly, the statute of limitations runs from "the date on which the judgment became final by the conclusion of direct review or the expiration of the time for seeking such review."[3] The statute of limitations tolls, however, for "[t]he time during which a *properly filed* application for State post-conviction or other collateral review with respect to the pertinent judgment or claim is pending."[4] Last Term, the Supreme Court defined the term "properly filed," holding in *Pace v. DiGuglielmo*[5] that

---

1. Pub. L. No. 104-132, 110 Stat. 1214 (1996) (codified as amended in scattered sections of 21 & 28 U.S.C.); *see* Clay v. United States, 537 U.S. 522, 528 (2003) ("Prior to [AEDPA], no statute of limitations governed requests for federal habeas corpus . . . relief.").

2. *See* 28 U.S.C. § 2244(d)(1)(A)–(D) (2000). The statute provides that:

> A 1-year period of limitation shall apply to an application for a writ of habeas corpus by a person in custody pursuant to the judgment of a State court. The limitation period shall run from the latest of—
>
> (A) the date on which the judgment became final by the conclusion of direct review or the expiration of the time for seeking such review;
>
> (B) the date on which the impediment to filing an application created by State action in violation of the Constitution or laws of the United States is removed, if the applicant was prevented from filing by such State action;
>
> (C) the date on which the constitutional right asserted was initially recognized by the Supreme Court, if the right has been newly recognized by the Supreme Court and made retroactively applicable to cases on collateral review; or
>
> (D) the date on which the factual predicate of the claim or claims presented could have been discovered though the exercise of due diligence.

Id.

3. *Id.* § 2244(d)(1)(A).

4. *Id.* § 2244(d)(2) (emphasis added).

5. 125 S. Ct. 1807 (2005).

an untimely application for state collateral relief is never properly filed and cannot toll the federal habeas statute of limitations.[6]

Pace is a vigorous attack against delays that have characterized federal habeas review of state judgments for decades. Although the Court declared a bright-line rule that can be easily applied, *Pace* is in other ways unclear. Part II of this Note gives an overview of the relevant facts and procedural history; Part III briefly outlines the issues that gave rise to a statute of limitations on the filing of federal habeas corpus petitions by state defendants; and Part IV elaborates on the Court's application of AEDPA's statute of limitations and tolling provisions, makes suggestions for further implementation of the Court's holding, and argues that the Court's draconian position against untimely applications for state collateral review was taken to advance the finality of state judgments.

II. FACTS AND PROCEDURAL HISTORY

John Pace pled guilty to second-degree murder and possession of an instrument of crime in a Pennsylvania state court and was sentenced to life imprisonment without the possibility of parole in February 1986.[7] On November 27, 1996, he filed a pro se petition for collateral review of his sentence under the Pennsylvania Post Conviction Relief Act (PCRA).[8] The PCRA became effective in 1988 and was amended in November 1995 to include, for the first time, a one-year statute of limitations on postconviction petitions.[9] The amended statute states that any postconviction petition, including a second or subsequent petition, shall be filed within one year of the date the petitioner's judgment becomes final, unless the petition falls under one of three exceptions.[10] The statute of limitations went into effect on January 16, 1996.[11] Pace filed his PCRA petition well over one year after his conviction became final and approximately ten

6. *Id.* at 1814.

7. *Id.* at 1810.

8. *Id.*; 42 PA. CONS. STAT. ANN. §§ 9541–9545 (West 1998).

9. *Pace*, 125 S. Ct. at 1810.

10. A PCRA petition otherwise barred by the new time limit could be accepted for review on the merits if (1) governmental interference prevented filing, (2) new facts arise that could not have been discovered through due diligence, or (3) a new constitutional rule is made retroactive. 42 PA. CONST. STAT. ANN. § 9545(b)(1).

11. *See* 1995 Pa. Laws 1118.

months after the PCRA statute of limitations went into effect, but at the time of his filing "no Pennsylvania court had yet applied the PCRA statute of limitations to a petitioner whose conviction had become final prior to the effective date of the Act."[12] AEDPA's one-year statute of limitations was also in effect at the time of Pace's filing, having been enacted in April 1996. Because the Third Circuit Court of Appeals had created a one-year grace period within which petitioners whose convictions had become final prior to the enactment of AEDPA could file a timely federal habeas application,[13] he had five months, without tolling, to the April 1997 federal deadline. In July 1997, the Pennsylvania Court of Common Pleas dismissed Pace's petition on its merits,[14] without ruling on the timeliness of the application.[15] Pace appealed in August 1997, approximately four months after the federal habeas grace period had run.[16] In December 1998 the Superior Court of Pennsylvania held that Pace's PCRA petition in the Court of Common Pleas and his petition in the Superior Court were both untimely.[17] The court retroactively applied the PCRA limitations period to Pace's application and held that his time for filing under the PCRA expired one year after his conviction became final.[18] The state supreme court denied review in July 1999.[19]

In December 1999, Pace filed a pro se federal habeas petition in the District Court for the Eastern District of Pennsylvania.[20] Recognizing that AEDPA's statute of limitation would bar Pace's federal petition, the district court held that Pace was entitled to statutory tolling from the time he filed his PCRA petition in November 1996 to the time he received a final judgment on his petition in July 1999.[21] The court reasoned that, because the PCRA time bar was unclear as it applied to Pace's petition,

---

12. *Pace*, 125 S. Ct. at 1817 (Stevens, J., dissenting).

13. *See* Burns v. Morton, 134 F.3d 109, 111 (3d Cir. 1998) (cited in Pace v. Vaughn, 151 F. Supp. 2d 586, 590 (E.D. Pa. 2001)).

14. *See* Brief for the Petitioner at 11, Pace v. DiGuglielmo, 125 S. Ct. 1807 (2005) (No. 03-9627), 2004 WL 2945392.

15. *Id.*

16. *Id.* at 12.

17. *Vaughn*, 151 F. Supp. 2d at 589–90.

18. *See id.* at 589.

19. *Id.* at 590.

20. *Id.*

21. *Id.* at 594.

such bar could not prevent his application from being properly filed.[22]

The Third Circuit Court of Appeals reversed. Relying on its own precedent, it concluded that when a state court deems a petition untimely, it is not properly filed.[23]

The Supreme Court affirmed. In a 5-4 vote,[24] the Court held that a postconviction petition that is rejected by the state courts as untimely is not properly filed within the meaning of AEDPA.[25] The Court did not address the specific predicament Pace faced: that the timeliness rule was unclear. Instead, the Court announced a general rule that addressed whether the existence of judicially reviewable exceptions to a time limit could prevent a late application from being improperly filed. The Court held that "[i]n common understanding, a petition filed after a time limit, and which does not fit within any exceptions to that limit, is no more properly filed than a petition filed after a time limit that permits no exception."[26] To hold otherwise would contravene AEDPA's purpose, the Court reasoned, by allowing a state prisoner to toll the statute of limitations at will, simply by filing untimely state postconviction petitions.[27] The Court also held that a petitioner's "reasonable confusion" about the timeliness of her application for state collateral relief would constitute "good cause" to file a federal habeas petition within AEDPA's deadline, and to seek a "protective" stay and abeyance of the federal proceedings thus initiated to preserve the opportunity for federal review when state remedies are exhausted.[28]

---

22. *Id.* at 590–94.

23. Pace v. Vaughn, 71 F. App'x 127, 128 (3d Cir. 2003) (citing Merritt v. Blaine, 326 F.3d 157 (3d Cir. 2003)).

24. Chief Justice Rehnquist delivered the opinion of the Court, in which Justices O'Connor, Scalia, Kennedy, and Thomas joined. Justice Stevens filed a dissenting opinion, in which Justices Souter, Ginsburg, and Breyer joined.

25. Pace v. DiGuglielmo, 125 S. Ct. 1807, 1810 (2005).

26. *Id.* at 1811–12 (internal quotation marks omitted).

27. *See id.* at 1812.

28. *See id.* at 1813–14. The Supreme Court also considered whether Pace was entitled to equitable tolling from the time he filed his PCRA petition to the time the state supreme court denied review of that petition. It held that a litigant seeking equitable tolling must establish that (1) she has been pursuing her rights diligently, and (2) some extraordinary circumstances stood in her way. *Id.* at 1814. Pace was denied equitable tolling because he failed to establish that he had been pursuing his rights diligently. The Court found that Pace waited too long after his claims were available to him to assert them in his PCRA petition, and that he also delayed after his PCRA proceedings became final before asserting his claims in

Justice Stevens dissented, arguing that the PCRA time bar operates like an ordinary procedural bar that precludes a prisoner from obtaining relief based on the merits of the claims made in a petition for collateral review, such as a bar against claims previously determined on the merits on direct appeal, or a bar against claims that the petitioner unjustifiably failed to raise on direct appeal. The Court had previously held that such rules could not prevent an application for collateral relief from being properly filed.[29] If this precedent were applied to Pace's case, the PCRA time bar could not cause his application to be improperly filed.[30] Justice Stevens reminded the majority that AEDPA was designed to streamline and simplify the federal habeas system, which had been characterized by indeterminate delays and shameful overloading, and argued that the majority's holding was unfaithful to those goals because it would inevitably subject district courts to a flood of protective filings.[31] Justice Stevens would have held that an application for state collateral relief that is ultimately held to be untimely, based on judicial review of its claims, is properly filed and tolls the federal habeas limitations period unless there is explicit evidence that the application was filed as a dilatory tactic.[32] He rebuffed the majority's concern that petitioners would intentionally postpone federal habeas proceedings by filing untimely petitions for state collateral relief as misguided, stating that a federal habeas petitioner's principal interest is in speedy federal review of her claims and in immediate relief from state incarceration.[33]

---

federal court. *Id.* at 1815. Although the Court explained why Pace was not entitled to equitable tolling, it declined to address whether equitable tolling is applicable to AEDPA's statute of limitations. *Id.* at 1814 n.8. Equitable tolling is outside the scope of this Note, and is not discussed beyond this footnote.

29. *See id.* at 1817–21 (Stevens, J., dissenting) (citing Artuz v. Bennett, 531 U.S. 4 (2000)).

30. *See id.*

31. *Id.* at 1820–21 (citing Hohn v. United States, 524 U.S. 236, 264–65 (1998) (Scalia, J., dissenting)).

32. *Id.* at 1820 (Stevens, J., dissenting).

33. *See id.*

III. THE NEED FOR A FILING DEADLINE

Federal habeas corpus review of state judgments is the most controversial and friction-producing issue in the relation between the federal courts and the states.[34] Scholars have noted that "[t]here is an affront to state sensibilities when a single federal judge can order [the] discharge of a prisoner whose conviction has been affirmed by the highest court of a state."[35] Traditionally, there was no statute of limitations on federal habeas review and the doctrine of laches did not apply.[36] A prisoner could postpone federal litigation for years and the district courts could release state prisoners at any time, thereby undercutting the finality of state judgments, and compromising a state's ability to defend its judgments against collateral attack or retry the defendant should she be granted federal habeas relief.[37]

The Supreme Court responded to this problem with Rule 9(a),[38] which became effective in 1977.[39] Rule 9(a) allowed district courts to dismiss a habeas petition if the state "ha[d] been prejudiced in its ability to respond to the petition by [the] delay in its filing."[40] The rule was insufficient to quell concerns about delayed filing because it did not protect the state against federal habeas petitions filed and granted so late that the state was

---

34. 17A CHARLES ALAN WRIGHT, ARTHUR R. MILLER & EDWARD H. COOPER, FEDERAL PRACTICE AND PROCEDURE § 4261 (2d ed. 1988).

35. *Id.* Chief Justice Rehnquist, author of *Pace*, held similar views. *See* Snead v. Stringer, 454 U.S. 988, 993–94 (1981) (Rehnquist, J., dissenting from denial of cert.) ("It is scarcely surprising that fewer and fewer capable lawyers can be found to serve on state benches when they may find their considered decisions overturned [on tenuous grounds] by the ruling of a single federal district court judge."). Justice O'Connor has also expressed reservations concerning the writ. *See* Engle v. Isaac, 456 U.S. 107, 128 (1982) (O'Connor, J.) ("Federal intrusions into state criminal trials [via the writ of habeas corpus] frustrate both the States' sovereign power to punish offenders and their good-faith attempts to honor constitutional rights.").

36. *See* WRIGHT ET AL., *supra* note 34, § 4268.2.

37. *See* Mark Tushnet & Larry Yackle, *Symbolic Statutes and Real Laws: The Pathologies of the Antiterrorism and Effective Death Penalty Act and the Prison Litigation Reform Act*, 47 DUKE L.J. 1, 6 (1997).

38. Rules Governing Section 2254 Cases in the United States District Courts, 28 U.S.C. app. (2000) (effective Feb. 1, 1977, Act of Sept. 28, 1976, Pub. L. No. 94-426, § 1, 90 Stat. 1334, and as amended thereafter).

39. WRIGHT ET AL., *supra* note 34, § 4268.2.

40. Vasquez v. Hillery, 474 U.S. 254, 265 (1986) (internal citation omitted).

prejudiced in mounting a retrial.[41] Capital defendants were of particular concern because they had real motive to postpone execution by filing federal habeas petitions. Members of Congress complained that delays in seeking federal habeas relief attenuated the deterrent effect of the death penalty and stifled confidence in the criminal justice system.[42] The problem with enacting a strict filing deadline was that it created tension with the habeas "exhaustion doctrine," which requires persons in custody pursuant to a state court judgment to exhaust state remedies with respect to all claims before seeking federal habeas relief.[43] Some prisoners would not be able to exhaust state remedies in time to meet the federal deadline. Congress responded by enacting AEDPA and its one-year filing deadline.[44]

AEDPA advances the principles of comity, finality, and federalism "to provide for an effective death penalty."[45] Section 2244(d)(1)'s one-year time limit serves the finality of state court judgments while section 2244(d)(2)'s tolling provision gives the deadline the flexibility to accommodate the requirement that petitioners exhaust state remedies before seeking federal review.[46] Like any other statute of limitations, AEDPA was designed to compel petitioners to file within a reasonable time so that the state has a fair opportunity to defend.[47] The goal was to encourage promptness and diligence in the filing of habeas petitions and to promote repose, stability, and public confidence in the justice system by bringing finality to state judgments.[48]

---

41. *See id.* (noting that Rule 9(a) did not permit dismissal of a habeas corpus petition if the state would be prejudiced in retrying the defendant).

42. *See* 141 CONG. REC. S7480 (daily ed. May 25, 1995) (letter to President Clinton from the attorneys general of several states, dated May 10, 1995, reprinted in full in the Congressional Record at the request of Senator Hatch) ("In our own states, we continue to experience endless appeals and continuous delay. We believe that such abuse of the criminal justice system produces a disrespect for the law, and serves to undermine deterrence. This is particularly true with respect to the enforcement of the death penalty.").

43. *See, e.g.*, Rose v. Lundy, 455 U.S. 509, 510 (1982).

44. AEDPA did not pass with ease. It is the fruit of decades of federal habeas reform debate. *See* Tushnet & Yackle, *supra* note 37, at 4–12.

45. Pub. L. No. 104-132, 110 Stat. 1214, 1217 (1996) (codified as amended in scattered sections of 21 & 28 U.S.C.); *see* Williams v. Taylor, 529 U.S. 420, 436 (2000).

46. *See* Duncan v. Walker, 533 U.S. 167, 179 (2001).

47. *See* 51 AM. JUR. 2D *Limitation of Actions* § 14 (2000).

48. *See id.* § 16; *see also* 141 CONG. REC. S7657 (daily ed. June 5, 1995) (remarks of Senator Dole) ("If we really want justice that is 'swift, certain, and severe,' . . . then

The main concern was delay in death penalty cases, but Congress placed the statute of limitations on all federal habeas cases.[49]

IV. TOLLING THE FEDERAL DEADLINE

AEDPA is poorly drafted and difficult to interpret in light of the federal and state processes of collateral review in place at the time of its enactment.[50] The Supreme Court has observed that "in a world of silk purses and pigs' ears, AEDPA is not a silk purse."[51] Shortly after AEDPA's enactment, Professor Larry Yackle judged that the new law would require a "practical, problem-solving brand of construction" that achieves sensible results.[52] The *Pace* Court's application of the statute and procedural setup does not seem to meet Professor Yackle's sensibleness requirement, particularly if it will result, as Justice Stevens warned, in an excess of protective filings in the federal district courts.[53] The remainder of this Note elaborates on the Court's application of AEDPA's limitation period in *Pace*, and suggests that district courts issue a stay of federal habeas proceedings when any reasonable jurist could find the application for state collateral relief timely under state law. It also attempts to provide a rationale for the procedure prescribed in *Pace*, arguing that the Court's holding was made in the interests of finalizing state judgments.

A. Proper Filing

AEDPA's tolling provision provides that "[t]he time during which a properly filed application for State post-conviction or other collateral review with respect to the pertinent judgment

---

we must stop the endless appeals and endless delays that have done so much to weaken public confidence in our system of criminal justice . . . . In America, today there is clearly a big disconnect between crime and punish [sic]. Our habeas reform proposal seeks to bridge this gap by imposing a 1-year filing deadline on all death row inmates, State or Federal.").

49. *See* Pub. L. No. 104-132, 110 Stat. 1214, 1217 (1996) (codified as amended in scattered sections of 21 & 28 U.S.C.).

50. *See* Larry W. Yackle, *A Primer on the New Habeas Corpus Statute*, 44 BUFF. L. REV. 381, 381 (1996) ("[T]hat [AEDPA] takes the preexisting habeas landscape as its baseline . . . will be of vital significance for the interpretive task that lies ahead. That task will not be easy. The new law is not well drafted.").

51. Lindh v. Murphy, 521 U.S. 320, 336 (1997).

52. *See* Yackle, *supra* note 50, at 381–82.

53. *See Pace*, 125 S. Ct. at 1820–21 (Stevens, J., dissenting).

or claim is pending shall not be counted toward any period of limitation."[54] The Supreme Court first defined the term "properly filed" in *Artuz v. Bennett*.[55] In *Artuz*, the petitioner argued that Bennett's application for state collateral relief was not properly filed because it was subject to two procedural bars: first, a bar against raising an issue that had been "previously determined on the merits upon an appeal from the judgment,"[56] and second, a bar against raising a claim that was available on direct appeal but was not raised because of the defendant's "unjustifiable failure."[57] The Supreme Court rejected the argument that Bennett's application was improperly filed, holding that "an application is *properly* filed *when its delivery and acceptance are in compliance with the applicable laws and rules governing filings*."[58] The Court distinguished between "conditions to filing," which, if not met, precluded an application from being properly filed, and "conditions to obtaining relief" such as the procedural bars at issue in that case, which could not prevent an application from being properly filed.[59] It elaborated that conditions to filing usually prescribe, for example, "the form of the document, *the time limits upon its delivery*, the court and office in which it must be lodged, and the requisite filing fee."[60]

The *Artuz* Court declared that applications that are delivered and accepted in compliance with state law governing filings are properly filed, and that procedural bars that function as conditions to obtaining relief cannot prevent an application from being properly filed. But the Court's holding was less clear on the question of whether the phrase "when . . . delivery and acceptance are in compliance with the applicable laws and rules governing filings" meant that a determination of "properly filed" was governed by state law.[61] *Artuz* can easily be read to accommodate varying requirements for invoking the tolling provision by giving states the authority to determine what

54. 28 U.S.C. § 2244(d)(2) (2000).

55. 531 U.S. 4 (2000).

56. N.Y. CRIM. PROC. LAW § 440.10(2)(a) (McKinney 2005).

57. *Id.* § 440.10(2)(c).

58. *Artuz*, 531 U.S. at 8 (second emphasis added) (internal quotation marks omitted).

59. *See id.* at 10–11.

60. *Id.* at 8 (emphasis added).

61. *Id.* at 8.

conditions to filing are required for an application for collateral relief to be properly filed. This reading of *Artuz* is supported by the Court's acknowledgment that "[i]n some jurisdictions the filing requirements also include . . . preconditions imposed on particular abusive filers or on all filers generally,"[62] but that in other jurisdictions these filing requirements may not exist. Legal commentators have understood that, under *Artuz*, "properly filed" is governed by state law.[63] If, under *Artuz*, proper filing is governed by state law, then whether an application barred by a state statute of limitations could still be properly filed would depend on whether the statute of limitations is a condition to filing or a condition to obtaining relief under state law. In jurisdictions where a statute of limitations is nothing more than an ordinary procedural bar that prevents an applicant from obtaining relief on her claims, *Artuz* would permit the state to determine that the statute of limitations is a condition to obtaining relief that cannot prevent an untimely application from being properly filed. On the other hand, in jurisdictions where, under state law, a statute of limitations is not an ordinary procedural bar, *Artuz* would give the state authority to determine whether the statute of limitations is a condition to filing that, when not met, precludes and application from being properly filed.

Whereas *Artuz* could be read to accommodate such state-to-state variation in the requirements for tolling the federal limitations period, and specifically in the ability of a time-barred application to trigger the federal tolling provision, *Pace* declined to extend flexibility where a petition for state collateral relief is ultimately held to be untimely. The *Pace* Court declared that "time limits, no matter their form, are filing conditions."[64] It stated that "[f]or purposes of determining what are 'filing'

---

62. *Id.* at 8–9 (internal citations omitted).

63. *See, e.g.*, Gary Knapp, *Supreme Court's construction and application of 28 USCS § 2244(d)(2), generally placing 1-year time limit on state prisoner's application for federal writ of habeas corpus*, 153 L. Ed. 2d 981, 985 (2004) ("[T]he Supreme Court has decided that . . . state law governs the issue whether an application for state [collateral] relief is 'properly' filed."); see also Brief for the Respondent at 29, Pace v. DiGuglielmo, 125 S. Ct. 1807 (2005) (No. 03-9627), 2005 WL 226923 ("The tolling provision of § 2244(d)(2), to be sure, is an exception to the [§ 2244](d)(1) limitation. Congress wanted to maintain some opportunity for additional exhaustion in state court before proceeding to federal court—*but only to the degree that states themselves were willing to entertain more litigation*.") (emphasis added).

64. Pace v. DiGuglielmo, 125 S. Ct. 1807, 1814 (2005) (internal quotations marks omitted).

conditions, there is an obvious distinction between time limits, which go to the very initiation of a petition and a court's ability to consider that petition, and the type of 'rule of decision' procedural bars at issue in *Artuz*, which go to the ability to obtain relief."[65] The Court's draconian position, that all timeliness rules are conditions to filing, leaves no room for state law determinations to the contrary, suggesting that "properly filed" is a matter of federal law. Though the dissenters disagreed with the Court on whether an untimely application for state collateral relief could toll the federal limitations period, they seemed to agree that the issue was governed by federal law. The dissenters argued that "[a] time bar is nothing more than a species of the larger category of procedural bars that may preclude consideration of the merits of the state petition."[66] Although Justice Stevens implicitly acknowledged that some states have timeliness rules that do not function like the procedural bars at issue in *Artuz*,[67] his proposed rule would not have allowed those states to declare that untimely petitions do not toll the limitations period. The consensus seems to be that, although proper filing will be determined with reference to state law governing filings, it is ultimately a matter of federal law. Specifically, timeliness is governed by state law, but proper filing is governed by federal law.

The Supreme Court provided the district courts with an unequivocal rule on the issue of timeliness. But it left them with little guidance for resolving criteria for determining what procedural rules can prevent applications for state collateral relief from being properly filed. In *Artuz*, the Court held that a state court application containing procedurally barred claims was properly filed so as to toll the federal habeas statute of limitations in part because to hold otherwise would "require judges to engage in verbal gymnastics when an application contains some claims that are procedurally barred and some that are not."[68] Justice Scalia wrote, for a unanimous Court, that a district court should not have to say that an application is properly filed *as to* nonbarred claims and not properly filed *as to* the

65. *Id.*

66. *Id.* at 1819 (Stevens, J., dissenting).

67. *See id.* ("most state laws respecting untimely filings of postconviction petitions function in a manner identical to the procedural bar at issue in Artuz").

68. Artuz v. Bennett, 531 U.S. 4, 10 (2000).

rest because AEDPA refers only to properly filed *applications* and does not contain the "peculiar suggestion" that a single application can be both properly filed and not properly filed.[69] In *Pace*, the Court appeared to have turned this analysis on its head, despite its rhetoric to the contrary. The *Pace* Court held that whether an application was properly filed in a state court could be determined by a claim-by-claim inquiry into the timeliness of the application.[70] It follows logically that the Court's holding in *Pace* would permit a district court to say that an application is properly filed as to claims not time-barred and not properly filed as to time-barred claims, leading to the "peculiar" result—that of a single application being at one time properly filed and not properly filed—that *Artuz* expressly intended to avoid.[71]

The inconsistency between *Artuz* and *Pace* is further demonstrated by the Court's policy analysis. The *Pace* Court expressed the concern that if a state petitioner could toll the statute of limitations at will simply by filing untimely state postconviction petitions, this would turn the tolling provision into a de facto extension mechanism contrary to AEDPA's purpose and open the door to abusive delay.[72] Yet *Artuz* permits a state petitioner to toll the federal habeas statute of limitations at will simply by filing state postconviction petitions containing procedurally barred claims.[73] The ease with which the Court turned on its reasoning in *Artuz* suggests that the law governing tolling of the federal limitations period will not develop in a consistent and predictable manner.[74]

B. Reasonable Confusion

The *Pace* majority understandably desired one rule to govern the effect of untimely petitions for state collateral review on the federal deadline. Many states provide exceptions to their post-

69. *Id.*

70. *See* Pace v. DiGuglielmo, 125 S. Ct. 1807, 1813 (2005).

71. *See id.* at 1818 (Stevens, J., dissenting).

72. *See id.* at 1812 (majority opinion).

73. Admittedly, *Pace* attenuates this policy concern with respect to *Artuz*, for a petitioner could toll the federal habeas statute of limitations at will by filing state applications containing procedurally barred claims, but eventually her applications would be barred by any applicable state filing deadline, at which point the petitioner would no longer be able to toll abusively the federal limitation period.

74. *See also infra* note 94 for another example of inconsistent reasoning regarding the application of AEDPA's tolling provision.

conviction statutes of limitation that apply to individual claims.[75] A federal habeas procedure that lumps timeliness rules having exceptions with those that do not is simple and easy to administer. The Court's holding, however, is unfair on its face. Under *Pace*, a petitioner could run the federal limitations period while the state judge is considering the timeliness of the application, and would not know that the federal deadline has passed until after the judge had declared the application untimely.[76] To remedy this concern, the *Pace* Court held that "[a] petitioner's reasonable confusion about whether an application for state collateral review is timely will ordinarily constitute 'good cause' for [her] to file in federal court" for a stay and abeyance of federal habeas proceedings until state remedies are exhausted.[77] If a petitioner has established good cause and there is no indication that the petitioner engaged in intentionally dilatory tactics, then a stay of federal habeas proceedings should be granted.[78] The stay should be conditioned upon the petitioner's pursuing state court remedies thirty days after the stay is entered and returning to federal court within thirty days after state court exhaustion is completed.[79] The stay is meant to honor the intent behind the original enactment of the tolling provision by protecting a petitioner from the expiration of the federal deadline while she is trying to exhaust state remedies.

The Court did not elaborate on what constitutes reasonable confusion or good cause for the purposes of obtaining a stay and abeyance. At a minimum, reasonable confusion and therefore good cause could be established by the mere existence of an exception to the state bar against untimely petitions. In this situation, a federal stay and abeyance would be granted if a petitioner chooses to file an untimely state petition, hoping that she will obtain review of her claims on the merits by way of an exception to the time bar. The federal courts are unlikely to take this approach to good cause. Many states have exceptions to their bars to untimely petitions.[80] Under this expansive ap-

---

75. *Pace*, 125 S. Ct. at 1818 (Stevens, J., dissenting).
76. *See id.* at 1813 (majority opinion).
77. *Id.*
78. *See id.* at 1814 (citing Rhines v. Weber, 125 S. Ct. 1528, 1535 (2005)).
79. *See Rhines*, 125 S. Ct. at 1535.
80. *See Pace*, 125 S. Ct. at 1818 (Stevens, J., dissenting).

proach, a petitioner who is nearing the end of the one-year federal limitations period and who knows to a certainty that her potentially meritorious claims cannot fall under an exception to a state time bar could nevertheless obtain, at a minimum, a sixty-day extension on the federal deadline, thereby frustrating the purpose of AEDPA's statute of limitations. Even though such an application for a stay and abeyance would likely be denied based on the petitioner's dilatory tactics,[81] the possibility that the petitioner's plan could succeed makes it likely that determinations of reasonable confusion and therefore good cause will involve some degree of judicial analysis of the claims as they apply to the state timeliness rules.

At a maximum, reasonable confusion and therefore good cause could be established upon a finding, by the district court judge applying the state law on exceptions to untimely applications, that the petitioner's application for state collateral relief is timely. In this scenario, the judge would then grant a stay and abeyance and allow the petitioner to proceed to state court. Similarly, if the judge, applying the state timeliness rules, finds that the petitioner's claims do not fit the state timeliness exceptions, then she would deny the stay and abeyance. I will term this approach the "do-it-yourself" approach.

Case law indicates that the do-it-yourself approach might be adopted by the federal courts. In *Carey v. Saffold*,[82] the Supreme Court ordered the Ninth Circuit Court of Appeals to apply a complicated California timeliness law to determine whether a petitioner's previously filed and denied application to a state court was a "pending" application for collateral relief as the term is used in AEDPA's statute of limitations provision.[83] Because the law was complicated, the Court permitted the Court of Appeals to certify a question to the California Supreme Court for a clarification of its law on timeliness.[84] In *Saffold*, therefore, the Court promulgated a do-it-yourself approach whereby the federal court would apply the state law for the purposes of determining whether a petition for state collateral relief was pending under AEDPA. In *Rhines v. Weber*,[85] the Supreme Court held that the district court could grant a stay and

81. *See id.* at 1814 (majority opinion).
82. 536 U.S. 214 (2002).
83. *Id.* at 225–27.
84. *Id.* at 226–27.
85. 125 S. Ct. 1528 (2005).

abeyance of a federal habeas petition containing both exhausted and nonexhausted claims if the petitioner had good cause for her failure to first exhaust state court claims.[86] The approach in *Rhines* is in accordance with the do-it-yourself standard. The district court judge is asked to determine, by applying her own sound judgment and, of course, precedent, whether a petitioner had good cause for her failure to exhaust in state court.

The do-it-yourself approach is an inappropriate method of determining reasonable confusion and therefore good cause. First, one could reasonably argue that, because proper filing is governed by federal law, a federal judge's determination that a state application is timely under state law means that the application tolls the federal deadline and a protective stay is unnecessary. Second, the do-it-yourself approach fails to give adequate weight to the inherently uncertain nature of judging. A federal judge could apply a state timeliness rule and find that a petitioner's application is untimely, while a state court judge, applying the same law to an identical case, could find the application timely. The more complicated the state timeliness law, the higher the likelihood that this discrepancy will occur. If a petitioner is denied a federal stay based on a district judge's incorrect determination that the application is untimely under state law, but the court does not dismiss the application on the merits and the state court ultimately determines that the petition is timely, it seems that the petitioner's application would toll the federal deadline and she would later be able to proceed to federal court as if she had never been before, making the previous encounter a waste of time and resources for all parties involved. Third, in cases where the state timeliness law is complicated or unclear, the district court may need to certify a question to the state supreme court. Federal habeas proceedings do not toll the limitations period.[87] It would be painfully unfair for the federal habeas limitations period to run while the state court responds to the district court's certification. Fourth, the do-it-yourself method frustrates principles of comity and federalism because it allows federal judges to make prelimi-

86. *Id.* at 1535.

87. *See* Duncan v. Walker, 533 U.S. 167, 181–82 (2001).

nary rulings on timeliness questions that should be first addressed by state courts.

A better approach is for a judge to find reasonable confusion and therefore good cause when the timeliness of the application for state collateral relief is less than clear and any reasonable jurist could find the petitioner's application timely under state law. This approach best serves the statute of limitations and the exhaustion requirement. It avoids the delay and waste of resources associated with sending to state courts applications that cannot possibly be accepted under state law. It also recognizes the variations and uncertainties that are associated with judging, and gives the states the first opportunity to address the timeliness of applications for collateral relief. Finally, a procedure that does not require judges to make precise applications of state timeliness rules is easy to administer. Requiring district court judges to apply the state law could become unduly burdensome if the *Pace* ruling results in copious protective filings.

C. Finality of State Judgments

The *Pace* opinion is likely to be heavily critiqued. Most federal habeas petitioners file pro se.[88] Capital federal habeas petitioners have a statutory right to counsel[89] but noncapital petitioners, who constitute the bulk of federal habeas petitioners,[90] are appointed counsel in only rare circumstances.[91] Un-

88. RICHARD H. FALLON, JR., DANIEL J. MELTZER & DAVID L. SHAPIRO, HART AND WECHSLER'S THE FEDERAL COURTS AND THE FEDERAL SYSTEM 1298 (5th ed. 2003).

89. 102 Stat. 4393 (codified as amended at 21 U.S.C. § 848(q)(4)(B)). Habeas petitioners generally do not, however, have a constitutional right to effective counsel during federal or state collateral proceedings. *See, e.g.*, Pennsylvania v. Finley, 481 U.S. 551, 554–55 (1987); Johnson v. Avery, 393 U.S. 483, 488 (1969). Habeas scholarship is replete with calls for mandatory appointment of effective counsel, especially for death penalty petitioners. *See, e.g.*, Celestine Richards McConville, *Protecting the Right to Effective Assistance of Capital Postconviction Counsel: The Scope of the Constitutional Obligation to Monitor Counsel Performance*, 66 U. PITT. L. REV. 521 (2005); Celestine Richards McConville, *The Right to Effective Assistance of Capital Postconviction Counsel: Constitutional Implications of Statutory Grants of Capital Counsel*, 2003 WIS. L. REV. 31 (2003); Andrew Hammel, *Effective Performance Guarantees for Capital State Postconviction Counsel: Cutting the Gordian Knot*, 5 J. APP. PRAC. & PROCESS 347 (2003); Burke W. Kappler, *Small Favors: Chapter 154 of the Antiterrorism and Effective Death Penalty Act, the States, and the Right to Counsel*, 90 J. CRIM. L. & CRIMINOLOGY 467 (2000).

90. In 2000, capital petitions constituted only 259 of the 21,345 federal habeas petitions by state defendants. BUREAU OF JUSTICE STATISTICS SPECIAL REPORT, U.S. DEP'T OF JUSTICE, PRISONER PETITIONS FILED IN U.S. DISTRICT COURTS, 2000, WITH TRENDS 1980–2000, at 3 tbl.2 (2002).

represented and poorly educated petitioners will make protective filings because they will not be able to predict accurately the timeliness of their applications,[92] and some will mistakenly make protective filings in federal court when their state application is plainly timely or untimely. Therefore, the Court's protective filing requirement makes the federal courts vulnerable to a "flood" of such filings,[93] particularly where state timeliness rules are complicated or unclear. Given the possibility of these burdensome results, one could reasonably argue that the Court has failed to apply AEDPA's statute of limitations and tolling provision sensibly.[94]

---

91. A federal court must appoint counsel for an indigent petitioner if an evidentiary hearing is required, *see* R.8(c), Rules Governing Section 2254 Cases in the United States District Courts, 28 U.S.C. app. (2000) (effective Feb. 1, 1977, Act of Sept. 28, 1976, Pub. L. No. 94-426, § 1, 90 Stat. 1334, and as amended thereafter), or when necessary to utilize effectively discovery authorized by the court, *see* R. 6(a).

92. *See* Pace v. DiGuglielmo, 125 S. Ct. 1807, 1821 (2005) (Stevens, J., dissenting).

93. *See id.* at 1820–21.

94. The *Pace* procedure is also awkward; whether an application filed with the court clerk has tolled the federal deadline depends on events that will occur in the future. In other words, AEDPA's tolling provision is applied retroactively. A petitioner should be able to determine that her application has tolled the federal deadline once the court clerk accepts the application for review of its claims. Interestingly, four of the Justices in the majority (Chief Justice Rehnquist and Justices Scalia, Kennedy, and Thomas) had previously expressed dissatisfaction with retroactive application of AEDPA's tolling provision. In *Carey v. Saffold*, 536 U.S. 214 (2002), all four Justices expressed disapproval of a procedure that left uncertain the tolling effect of a previously filed state application for collateral relief. *See id.* at 234–35 (Kennedy, J., dissenting). In *Saffold*, the Court held that an application for postconviction or other collateral review is "pending," as the term is used in AEDPA's tolling provision, during the interval between a state lower court's determination and the further filing of either an appeal or an original state habeas petition to a higher court, but only if the further filing is timely. *See id.* at 217, 225–27. Under *Saffold*, a petitioner will not know whether her application for state collateral relief was pending so as to toll the federal limitations period from the date of the last judgment on state collateral review until the state court ruled on the timeliness of the latest application. In other words, an application for state collateral review retroactively becomes pending after the state court has ruled on the timeliness of the latest application. Justice Kennedy dissented, joined by Chief Justice Rehnquist and Justices Scalia and Thomas, stating that

> This [was] not a sensible way of determining when an application is 'pending' under the federal tolling provision. Whether an application is pending at any given moment should be susceptible of a yes or no answer. On the Court's theory, the answer will often be 'impossible to tell,' because it depends not on whether an application is under submission in a particular court but upon events that may occur at some later time.

Id. at 234–35 (Kennedy, J., dissenting). A similar argument could easily be made in Pace's favor. He should be able to determine at any given moment whether his

A brief look at the options the Court forwent in settling with the *Pace* stay and abeyance procedure yields insight into the Court's motive for adopting such a cumbersome procedure. The *Pace* opinion outlined, among others, two issues: whether an untimely application tolls the federal deadline, and whether a petitioner should be required to make a protective filing in federal court. The Court held that an untimely petition does not toll the federal deadline and a petitioner must make a protective filing to preserve the opportunity for federal review if she is uncertain about whether her state application for collateral relief is timely.[95] It also held that the district court should dismiss a petition, without granting a stay, if the claims presented are meritless.[96] A petitioner whose application is dismissed on its merits would be able to return to federal court under very limited circumstances.[97] *Pace* permits the federal courts to "peek" at the merits of a petitioner's application before state remedies have been exhausted. It encourages early conclusion of federal review of meritless petitions, and therefore promotes the finality of state judgments. From this perspective, the inefficiencies of the *Pace* procedure become favorable; petitioners who mistakenly make protective filings in federal court will face early dismissal of their applications if the claims within them are meritless.

---

application was properly filed under the federal tolling provision. It is not sensible for events that may occur in the future to be the basis for whether an application already under submission in a court is properly filed. The four Justices' departure from their position in *Saffold* is likely to be a point of criticism; the Court will likely be chastised for failure to apply AEDPA's tolling provision sensibly. The four votes can be seen, however, as an acceptance of retroactive application as an established characteristic of AEDPA's tolling provision. Strictly on the point of retroactivity, *Pace* does no damage that *Saffold* has not already done.

95. *See Pace*, 125 S. Ct. at 1813–14.

96. *See* Rhines v. Weber, 125 S. Ct. 1528, 1535 (2005). *See also* WRIGHT ET AL., *supra* note 34, § 4261 ("[P]risoners thrive on [federal habeas corpus] as a form of occupational therapy and for a few it serves as a means of redressing constitutional violations . . . . The great bulk of [federal habeas] applications are meritless.").

97. To appeal a denial of relief by a district judge, a petitioner must obtain from either the district court where he was denied relief or the court of appeals a "certificate of appealability," which issues upon a "substantial showing of the denial of a constitutional right." 28 U.S.C. § 2253(c)(1)–(2) (2000). The certificate of appealability must indicate the specific issues showing the denial of the right. *Id.* § 2253(c)(3). Second or successive petitions for federal habeas relief by state prisoners are governed by 28 U.S.C. § 2244(b) (2000) and can be filed only under very limited circumstances.

Assuming that the process of issuing a federal protective stay is not so time-consuming as to exacerbate the problem of delay, other options, though they may be more fair or less cumbersome, might not be as effective at promoting the finality of state judgments. The dissenters took the position that an untimely petition should toll the federal deadline if the petitioner did not file for state postconviction relief "explicitly to prolong the federal statute of limitations," and that a protective filing was not necessary.[98] An alternative option would be to require a petitioner to file for a stay because an untimely petition does not toll the federal deadline, but that the stay is to be issued if the petitioner did not file for state postconviction relief explicitly to prolong the federal statute of limitations. Yet another option is one in which petitioners are not required to make a protective filing, but are granted tolling after leaving the state court if there was reasonable confusion about the timeliness of the state application. None of these three options would be as effective at promoting the finality of state judgments as the procedure adopted by the Court. A petitioner can more easily obtain tolling or a stay if evidence against her must be "explicit." And a scheme that does not require petitioners to make a protective filing does not allow district courts to peek at petitions and dismiss the meritless ones early.

Another scheme would be to require petitioners to file for a stay because an untimely petition does not toll the federal deadline, but the stay and its sixty-day grace period would be invalidated by the tolling provision if the state court ultimately determines that the petition is timely. The principle behind this scheme is that a timely petition that otherwise complies with state law governing filings activates AEDPA's tolling provision and "that is the end of the matter for the purposes of [the tolling provision]."[99] Under this option, petitioners who meet the requirements for obtaining a stay would receive one, but only those whose state petitions are held to be untimely would ultimately use it. This option would be less effective at promoting the finality of state judgments in cases where activation of the tolling provision gave the petitioner longer than thirty days to

98. *Pace*, 125 S. Ct. at 1820–21 (Stevens, J., dissenting).

99. *See Pace*, 125 S. Ct. at 1812 (holding that when a postconviction petition is untimely under state law, that is the end of the matter for the purposes of AEDPA's tolling provision).

return to federal court. The thirty-day return was specifically designed to prevent capital petitioners from using the stay and abeyance procedure to prolong incarceration and avoid execution.[100] The procedure adopted by the Court allows the district court to rid the federal system of meritless petitions at an earlier time and to usher more quickly federal habeas petitioners through the system.

V. CONCLUSION

AEDPA's statute of limitations is designed to curb the delay that has plagued the federal habeas system. Because only capital prisoners have real incentive to delay, *Pace* is likely to have an effect on delay only in collateral attacks on death penalty judgments. *Pace*, however, furthers the interests of finalizing state judgments, one of the established goals of AEDPA's statute of limitations,[101] regardless of the nature of the case. Restricting access to federal habeas review of state judgments in the interests of the finality is an attack on the procedure at its core. Federal habeas by its very nature unsettles state judgments. Proponents of limited federal review of state judgments argue that state judges are as equally bound and competent to uphold the constitutional rights of defendants as federal judges. As Professor Bator so aptly put it, "if a job can be well done once, it should not be done twice."[102] Others see AEDPA as a means of eliminating a vital avenue for the vindication of constitutional rights. In *Sanders v. United States*,[103] Justice Brennan declared that "[c]onventional notions of finality of litigation have no place where life or liberty is at stake and infringement of constitutional rights is alleged."[104] The latter camp values federal habeas particularly in capital cases where correction of constitutional errors made in the state courts is literally a matter of life and death and where federal habeas has proven to be a vital means of correcting those errors.[105] Al-

100. *See Rhines*, 125 S. Ct. at 1535.

101. *See supra* Part IV.C.

102. Paul M. Bator, *Finality in Criminal Law and the Federal Habeas Corpus for State Prisoners*, 76 HARV. L. REV. 441, 451 (1963). See Professor Bator's article for a discussion of the role of finality in the criminal justice system.

103. 373 U.S. 1 (1963).

104. *Id.* at 8.

105. Studies reveal that between 1973 and 1995, federal habeas courts overturned forty percent of the state death penalty sentences they reviewed

though the *Pace* system promises to promote the finality of state judgments, it will exacerbate the problem it attempts to alleviate if the stay and abeyance procedure is too time-consuming and burdensome. A burdensome stay and abeyance procedure also compromises AEDPA's goal to streamline federal habeas proceedings.[106] In any case, the message to all petitioners who seek federal habeas review of a state judgment is clear: File early!

N. Noelle Francis

---

because of prejudicial or reversible error. *See* JAMES S. LIEBMAN, JEFFREY FAGAN & VALERIE WEST, A BROKEN SYSTEM: ERROR RATES IN CAPITAL CASES, 1973–1995 (2000).

106. *See* Pace v. DiGuglielmo, 125 S. Ct. 1807, 1820–21 (Stevens, J., dissenting) (expressing the concern that the Court's holding in *Pace* was unfaithful to AEDPA's goal of streamlining the federal habeas system); *cf.* Duncan v. Walker, 533 U.S. 167, 180 (2001) (stating that the exhaustion requirement was designed to reduce piecemeal litigation).