ENGAGE

The Journal of the Federalist Society Practice Groups

Project of the E.L. Wiegand Practice Groups

NOTA BENE

Incorporating Gun Rights: A Second Round in the Chamber for the Second Amendment
by Kenneth A. Klukowski

Getting Beyond Guns: Context for the Coming Debate over Privileges or Immunities
by Clark M. Neily III & Robert J. McNamara

Miranda *with an English Accent*
by Lauren J. Altdoerffer

A Cold Breeze in California: ProtectMarriage *Reveals the Chilling Effect of Campaign Finance Disclosure on Ballot Measure Issue Advocacy—Comment and Rebuttal*
Stephen R. Klein & Steve Simpson

Nashville in Africa: Intellectual Property Law, Creative Industries, and Development
by Mark Schultz & Alec van Gelder

The Fourth Amendment Goes to War
by Robert J. Delahunty

BOOK REVIEWS

Federalism: Political Identity and Tragic Compromise
by Malcolm M. Feeley & Edward Rubin

Regulation by Litigation
by Andrew P. Morriss, Bruce Yandle & Andrew Dorchak

Volume 10, Issue 3 November 2009

ENGAGE: THE JOURNAL OF THE FEDERALIST SOCIETY PRACTICE GROUPS provides original scholarship on current, important legal and policy issues. It is a collaborative effort, involving the hard work and voluntary dedication of each of the organization's fifteen Practice Groups. Through its publication, these Groups aim to contribute to the marketplace of ideas in a way that is collegial, measured, and insightful—to spark a higher level of debate and discussion than we often see in today's legal community.

Likewise, we hope that members find the work in the pages to be well-crafted and informative. Articles are typically chosen by our Practice Group chairmen, but we strongly encourage members and general readers to send us their commentary and suggestions at info@fed-soc.org.

Volume 10, Issue 3 ENGAGE November 2009

Civil Rights

Two Civil Rights Decisions Close Out Supreme Court's 2008 Term
by Christian J. Ward & Edward C. Dawson 4

Incorporating Gun Rights: A Second Round in the Chamber for the Second Amendment
by Kenneth A. Klukowski 10

Getting Beyond Guns: Context for the Coming Debate over Privileges or Immunities
by Clark M. Neily III & Robert J. McNamara 18

Corporations, Securities & Antitrust

Looks Can Be Deceiving: Holdout Litigation Under the Foreign Sovereign Immunities Act
by Saloni Kantaria 26

Judicial Review of Mutual Fund Advisory Fees: Reliance on Markets or Statutory Language?
by Joanne T. Medero 32

Criminal Law & Procedure

Miranda with an English Accent
by Lauren J. Altdoerffer 35

Whither the Rule of Lenity
by Dan Levin & Nathaniel Stewart 42

Environmental Law & Property Rights

Ripening Federal Property Rights Claims
by J. David Breemer 50

The Supreme Court and the Judicial Takings Doctrine
by Steven Geoffrey Gieseler & Nicholas M. Gieseler 54

Financial Services & E-Commerce

The Stop Trading on Congressional Knowledge Act
by Stephen M. Bainbridge 59

The Financial Reform Plan: What It Means for Insurance Companies
by Laura Kotelman 66

Free Speech & Election Law

A Cold Breeze in California: *ProtectMarriage* Reveals the Chilling Effect of Campaign Finance Disclosure on Ballot Measure Issue Advocacy
by Stephen R. Klein 68

Attack Ballot Issue Disclosure Root and Branch:
Comment on *A Cold Breeze in California:* ProtectMarriage *Reveals the Chilling Effect of Campaign Finance Disclosure on Ballot Measure Issue Advocacy*
by Steve Simpson 74

Rebuttal to Steve Simpson's Response to *A Cold Breeze in California:* ProtectMarriage *Reveals the Chilling Effect of Campaign Finance Disclosure on Ballot Measure Issue Advocacy*
by Stephen R. Klein 78

Politics for Professionals Only: Ballot Measures, Campaign Finance "Reform," and the First Amendment
by Dick M. Carpenter, Ph.D., Jeffrey Milyo, Ph.D. & John K. Ross 80

Intellectual Property

Qui-tam-osaurus, the Statutory Dinosaur: Evolution or Extinction for the Qui Tam Patent False Marking Statute?
by Trevor K. Copeland & Laura A. Lydigsen 86

Nashville in Africa: Intellectual Property Law, Creative Industries, and Development
by Mark Schultz & Alec van Gelder 94

International & National Security Law

The Fourth Amendment Goes to War *by Robert J. Delahunty* 107

Litigation

Constitutional Limits on Punitive Damages: An Evaluation of the Role of Economic Theory in Prescribing Constitutional Constraints on Punitive Damage Awards
by Dorothy Henderson Shapiro 115

A Wave of ADA Public Accommodation Lawsuits Moves from Florida and California to the Rest of the United States
by Gregory P. McGuire 123

Professional Responsibility & Legal Education

Government Ethics in President Obama's First Year: A Preliminary Assessment
by Richard W. Painter 128

Religious Liberties

Why the Supreme Court Has Fashioned Rules of Standing Unique to the Establishment Clause
by Carl H. Esbeck 134

Smith, *Stormans*, and the Future of Free Exercise: Applying the Free Exercise Clause to Targeted Laws of General Applicability
by Mark L. Rienzi 143

Book Reviews

Federalism: Political Identity and Tragic Compromise by Malcolm M. Feeley & Edward Rubin
Reviewed by George D. Brown 149

Regulation by Litigation by Andrew P. Morriss, Bruce Yandle & Andrew Dorchak
Reviewed by Margaret A. Little 151

Civil Rights

Two Civil Rights Decisions Close Out Supreme Court's 2008 Term

*By Christian J. Ward & Edward C. Dawson**

The Supreme Court's 2008 Term concluded with opinions in two closely followed civil rights cases, *Northwest Austin Municipal Utility District No. 1 v. Holder* (*Northwest Austin MUD)*, and *Ricci v. DeStefano* (*Ricci*). Both cases were anticipated as presenting possibilities for sweeping constitutional holdings—in *Northwest Austin MUD*, the invalidation of the Voting Rights Act, and in *Ricci*, the application of Equal Protection analysis to workplace claims of "reverse" discrimination under Title VII. In fact, neither case produced a constitutional seachange, but instead both were decided on grounds of statutory interpretation, consistent perhaps with Chief Justice Roberts's articulated preference for a "minimalist" jurisprudential approach. Nonetheless, both cases achieved significant, incremental change—in recognizing the Nation's significant advances in guaranteeing equal voting rights to all, and in advancing the vision of antidiscrimination employment law as a vehicle for ensuring equal, race-neutral employment opportunities. This article summarizes and analyzes each of the two decisions, and offers some thoughts about the respective implications of each for future developments in voting-rights and employment-discrimination law.

I. Northwest Austin Municipal Utility District No. 1 v. Holder: A Time to Move Forward

In *Northwest Austin MUD*,[1] the U.S. Supreme Court signaled that it is time to reevaluate four-decade-old presumptions underlying enforcement of the Voting Rights Act, unanimously requiring the Department of Justice and the U.S. District Court for the District of Columbia to significantly broaden the availability of "bailout"—i.e., exemption from Section 5 of the Act–and expressing skepticism regarding whether Section 5 remains a constitutional exercise of Congress's enforcement power.

Congress enacted the landmark Voting Rights Act of 1965 "to banish the blight of racial discrimination in voting, which ha[d] infected the electoral process in parts of our country for nearly a century."[2] Previous statutory attempts to enforce the guarantees of voting rights enshrined in the Constitution had met with little success in overcoming deep-rooted intransigence in certain parts of the country, notably–and unsurprisingly–in the Civil Rights Era South.

At a time when a Southern governor might stand in a schoolhouse doorway attempting to stave off court-ordered integration, case-by-case litigation of voting rights abuses proved largely ineffective at substantially increasing voter registration and participation by African Americans. The Justice Department might have received an order from a court declaring, say, a literacy test for voter registration unconstitutional only to have a recalcitrant state change its registration requirements enough to evade the court order while retaining their discriminatory purpose and effect. That type of gamesmanship—which has been compared to a game of Whac-A-Mole[3]—is what prompted inclusion of Section 5 in the 1965 Act. As the Supreme Court later explained, "Section 5 was a response to a common practice in some jurisdictions of staying one step ahead of the federal courts by passing new discriminatory voting laws as soon as the old ones had been struck down."[4]

The Section 5 response was an unprecedented, and still unparalleled, feature in American law—a literal federal veto power over certain laws and policy choices made by state and local governmental entities. Basically, a state or political subdivision that is covered by Section 5 must get federal preapproval—known as preclearance—for any change affecting voting, either by seeking a declaratory judgment from the U.S. District Court for the District of Columbia or by submitting the change for vetting by the Attorney General.[5] The district court or the Justice Department must reject the change if it finds that the change has the purpose or effect of abridging the right to vote.[6] It almost goes without saying that the administrative route of submitting changes to the Justice Department is used far more often than the more cumbersome option of litigating in the district court. Because the Supreme Court has made clear that laws subject to Section 5 preclearance "are not now and will not be effective as laws until and unless cleared pursuant to § 5,"[7] the Justice Department's role essentially places the federal Executive Branch in the position of a sort of super-governor, with the power to overrule a state's legislature and its governor. Section 5 was originally set to expire in five years, viewed as a temporary, emergency measure, but has been extended repeatedly, most recently in 2006 until 2031.[8]

States and certain localities (counties, parishes, and other entities if they register voters) were subjected to Section 5 coverage by a formula that takes into account the existence of a "test or device" and voter registration and turnout rates for specified presidential election years.[9] The original 1965 formula was reverse-engineered to capture well-known offenders against the voting rights of African Americans, including the states of Alabama, Georgia, Louisiana, Mississippi, South Carolina, and Virginia. It relied on registration and turnout rates from the 1964 presidential election.[10] In 1975, Section 5 was extended for the second time and the coverage formula amended to capture jurisdictions believed to have discriminated against language minorities, bringing into the fold the states of Alaska, Arizona, and Texas and counties in states including California and New York.[11] The 1975 formula uses data through the 1972 presidential election, and subsequent reenactments of Section 5 have involved no further changes to the coverage formula,

** Christian J. Ward and Edward C. Dawson are partners in the Austin, Texas office of Yetter, Warden & Coleman, L.L.P. Ward represented the plaintiff utility district in Northwest Austin Municipal Utility District No. 1, and is president of the Austin Lawyers Chapter of the Federalist Society. Dawson represented the petitioning firefighters in Ricci v. DeStefano. Both cases were argued by their partner, Gregory S. Coleman.*

meaning that all jurisdictions covered today are covered based on registration and election data from no later than 1972.[12]

The Supreme Court also made clear early on that the VRA takes a very broad view of what constitutes a change affecting voting, meaning Section 5's preclearance requirement extends to tiny alterations or those that have even the remotest effect on voting,[13] which might include personnel policies adopted by a school board that did not conduct elections,[14] annexation of unpopulated land, or moving a utility district's polling place from a residential garage to a public school a short distance away. Additionally, and importantly to the outcome of *Northwest Austin MUD*, the Court, in the 1978 case *United States v. Board of Commissioners of Sheffield*,[15] interpreted Section 5 to require that *any* political subunit within the territory of a covered state submit its changes for preclearance, notwithstanding a more restrictive definition of "political subdivision" elsewhere in the Act that appeared to limit the term's application to counties, parishes, and other entities only if they registered voters.[16]

The Supreme Court upheld the constitutionality of Section 5 in 1966 but cautioned even then that the 1965 enactment had been supported specifically by evidence of "exceptional conditions" that could "justify legislative measures not otherwise appropriate."[17] As early as the 1982 extension of Section 5, Congress had begun to recognize that the original emergency may have passed and Section 5 outlived its usefulness. Its extension included a newly expansive "bailout" provision for "political subdivisions," intended to allow localities that could demonstrate a decade of compliance with the Constitution and the VRA an exemption from Section 5.[18]

Northwest Austin Utility District Number One, the utility district for an Austin, Texas neighborhood of about 3,500 residents known as Canyon Creek, sought to take advantage of the bailout provision. In 2004, members of the district's board learned that they had to get federal preclearance before moving the district's polling place from a private garage to the nearby public elementary school where the other local elections were held on the same day. They regarded this as a ridiculous federal intrusion into local affairs, especially given that Canyon Creek did not exist until the late 1980s, long after the turbulent civil rights struggles of the 1960s, and that it has absolutely no history of voting discrimination. The district filed a bailout suit, as required, in the D.C. federal district court. Recognizing, however, that the Justice Department and others had long interpreted the bailout provision restrictively—specifically, applying the statutory definition of "political subdivision" to conclude that governmental units smaller than counties were ineligible to seek bailout—the district's suit included an alternative claim that the 2006 reenactment of Section 5 exceeded Congress's constitutional authority.

The district argued, based on the statutory language and the Supreme Court's holding in *Sheffield*, that it must be regarded as a "political subdivision" eligible to bail out. Its alternative argument was that, under cases including *South Carolina v. Katzenbach* and *City of Boerne v. Flores*,[19] Congress could not reenact Section 5 as a prophylactic measure in 2006 because, so long after the original emergency ended, it was far too broad to be regarded as simply enforcing guarantees in the Fourteenth and Fifteenth Amendments. The fact that less than a tenth of a percent of proposed voting changes now draw objections from the Justice Department demonstrates that Section 5 is an outmoded federal intrusion into local government.

The district court rejected both of the district's arguments, but the Supreme Court reversed, holding that the district correctly interpreted the bailout statute to make the district eligible to pursue a bailout.[20] The Court agreed with the district's reasoning that the prior holdings in *Sheffield* and *Dougherty County* compelled the conclusion that "political subdivision" must be given its ordinary meaning, which obviously includes entities like utility districts.[21]

Tellingly, the entire Court signed onto language suggesting that they find the 2006 extension of Section 5 at least constitutionally suspect. The majority noted, for example, that "Section 5 goes beyond the prohibition of the Fifteenth Amendment by suspending *all* changes to state election law—however innocuous—until they have been precleared by federal authorities in Washington, D.C.," that "the Act also differentiates between the States, despite our historic tradition that all the States enjoy 'equal sovereignty,'" and that "[t]he statute's coverage formula is based on data that is now more than 35 years old, and there is considerable evidence that it fails to account for current political conditions."[22] Chief Justice Roberts wrote the majority opinion, with which seven other justices (six of whom remain on the Court today) concurred. Justice Thomas concurred in the judgment in part and dissented in part only to express his view that the Court should have gone further, reaching the constitutional challenge and striking Section 5 down.[23]

The Court has indicated that it wants to see bailout become a frequently-used, effective mechanism for reducing the scope of Section 5 coverage. The Court as a whole is at least skeptical whether Section 5 remains a constitutional remedy at all. Through the expansion of bailout and, perhaps soon, the eventual ending of Section 5, the nation is ready to move forward with voting rights enforcement that is no longer based on the presumption that race relations remain mired in the 1960s.

II. Ricci v. DeStefano–Walking the Line Between Discrimination and "Reverse" Discrimination in Employment Testing

In *Ricci v. DeStefano*,[24] decided June 29, 2009, the Supreme Court clarified the law governing the interaction of disparate-impact and disparate-treatment discrimination claims under Title VII. The decision will have wide-ranging implications for employment practices and litigation under Title VII in both the public and private sectors and may also signal a deeper tension between the commands of the Constitution and the dictates of disparate-impact law under Title VII. In addition, the case is notable for the prominent role it played in the confirmation hearings of the newest Associate Justice, Sonia Sotomayor.

Title VII prohibits intentional acts of discrimination based on race, color, religion, sex, and national origin.[25] It also prohibits policies that do not intentionally discriminate but have a disproportionate adverse effect on minorities.[26] However,

disparate-impact discrimination only violates Title VII when an employer is unable to show that a challenged practice is job-related, or the employee-plaintiff can show that there are less discriminatory alternative practices that equally serve the same legitimate business need. In the public-employment context, moreover, the constitutional guarantee of equal protection requires that any race-based action by a government actor must be subjected to strict scrutiny and invalidated in all but the rarest of circumstances.[27] The thorny question the Court faced in *Ricci* was how to resolve these competing commands when a governmental employer administers a promotion test that produces racially disparate results and has to decide whether to use or reject the test based only on the skewed racial distribution of the scorers.

Ricci arose out of a dispute over firefighter promotions in New Haven, Connecticut. The New Haven Fire Department uses objective oral and written examinations to decide who should be considered to fill vacant lieutenant and captain positions, which are meant to determine the most qualified individuals for command positions. This examination system for promotions within the classified civil-service industries is governed by the city's charter, in addition to federal and state law. The promotion process also has separate requirements through a contract between the city and its firefighter union, which specifies that a promotion candidate's composite examination score must be determined through an examination process that is sixty percent written and forty percent oral. Normally, the city administers the test, and then, once it receives the results, the New Haven Civil Service Board (CSB) is asked to certify the ranked list of applicants who passed by achieving a composite score of seventy or higher. After the list is certified, the city charter requires that a "rule of three" is used by the hiring authority to fill the vacancy. This rule allows the hiring authority (here, the NHFD) to promote any one candidate from among the top three scorers on the list.

New Haven had previously experienced racial disparities in the number of eligible candidates for promotion selected through its tests. Therefore, before administering its 2003 tests, it undertook extensive efforts to ensure that the tests were fair and free of any non-job-related tendency to produce racially disparate results. After reviewing various consultants, the City hired Industrial/Organizational Solutions, Inc. (IOS) to create and administer the promotional examinations at a cost of $100,000. To begin the test-design process, IOS performed job analyses for the captain and lieutenant positions. IOS also went through an extensive interview process with incumbent captains and lieutenants and their supervisors in order to determine the knowledge, skills, and abilities that are essential for the positions. Throughout the research for the test design, IOS intentionally oversampled minority firefighters to prevent the tests from favoring white applicants. IOS compiled a list of reading materials approved by the fire chief and assistant fire chief and disseminated that list to the candidates, including the specific chapters that were used in the development of the examination. In addition, IOS took painstaking measures to make sure that the scoring of the oral portion did not favor any race.

The examinations were given in November and December of 2003. Seventy-seven candidates completed the lieutenant examination—forty-three whites, nineteen blacks, and fifteen Hispanics. Of those, thirty-four candidates scored high enough to qualify for the eligibility list—twenty-five whites, six blacks, and three Hispanics. Based on these results and the "rule of three," only the top ten scorers could be considered for the eight lieutenant positions open at that time. All ten were white. Forty-one candidates completed the captain examination—twenty-five whites, eight blacks, and eight Hispanics. Of those, twenty-two candidates passed—sixteen whites, three blacks, and three Hispanics. Seven of the top scorers were white and two were Hispanic. There were seven open captain positions, which, together with the rule of three, meant that the top nine scorers could be considered for the immediately-available promotions. In addition, further vacancies at both the captain and lieutenant level were anticipated to arise during the two years the eligibility list would remain in effect.

After receiving these results, city officials became worried that the examinations unintentionally discriminated against minorities. Some firefighters were upset by the results and threatened to sue the city if it promoted from eligibility lists based on the tests, claiming that because the test results had racial disparities, the tests violated the disparate-impact provision of Title VII. In addition, city officials came under political pressure from local activists not to certify the results. There was evidence in the record that, once the racial distributions of the test scores were known, city officials orchestrated a campaign designed, for a mixture of political and racial reasons, to result in the rejection of the test results. The CSB then held several hearings at which it heard testimony from persons interested in the certification issue. During these hearings, some witnesses raised questions about the tests that had been given, but at no point were the tests shown to have been impermissibly biased or unrelated to the jobs for which the applicants were applying. The CSB also heard testimony from an expert employed by a competitor of IOS who speculated that it might be possible to design an equally job-related test that would have less racial numerical disparity; however, he did not identify any actual alternatives and moreover advised the CSB that the best thing for it to do would be to certify the test results.

Ultimately, the CSB deadlocked by tie vote, which meant that the eligibility list was not certified, and no promotions were made. A group of white and Hispanic firefighters who believed they had done well on the tests and had been wrongly denied their chance at promotion sued the city and its officials, claiming the city had intentionally discriminated against them because of their race in violation of Title VII and the U.S. Constitution's Equal Protection Clause. Discovery later confirmed that most of the plaintiffs indeed were on the rejected promotion-eligibility lists. The plaintiff firefighters argued that the defendants' decision to throw out the test results because of the racial distribution of the successful candidates was intentional, impermissible racial discrimination that was not and could not be justified by the defendants' claimed concerns that certifying the test results would result in impermissible unintentional disparate-impact discrimination.[28] The defendants countered that the decision not to certify the test results was justified because they had a good-faith belief

that certifying the test results could have exposed the City to litigation and potential liability under Title VII's disparate-impact provisions. The defendants did not argue that the test as given was actually flawed or that they had concrete evidence of superior alternatives. Nor did they argue that their actions had been justified on the basis of achieving diversity—they limited themselves to the Title-VII-compliance rationale.

After discovery and briefing, the district court granted summary judgment to the defendants. On the equal protection point, the district court reasoned that there had been no racial classification at all because the test results were thrown out for all test takers, without regard to race. On the Title VII issue, the district court concluded that the defendants were immune from liability as a matter of law because they had a subjective good-faith belief that certifying the test results could result in exposure to disparate-impact litigation or liability. The Second Circuit affirmed in a per curiam, one-paragraph opinion that merely adopted the district court's opinion. Judge Sotomayor was one member of the Second Circuit panel. Subsequently, the appellate court sua sponte considered whether to rehear the case en banc, and voted 7-6 against rehearing over a strenuous and compelling dissent by Judge Jose Cabranes.

After the firefighters petitioned for review, the Supreme Court granted certiorari. The central questions before the Court were whether and when under Title VII an employer may engage in intentional racially disparate treatment in order to avoid or forestall potential, unintentional racial disparities, and how in the public-employer context this analysis is informed by the constitutional guarantee of equal protection.

The petitioning firefighters argued that the defendants' refusal to certify the test results was a race-based action subject to strict scrutiny under the Equal Protection Clause, and that it could not survive that scrutiny because it was neither justified by any compelling state interest nor narrowly tailored to achieve any such interest. They noted, in particular, the absurdity of the district court's conclusion that refusing to certify test results based on the racial distribution of the successful candidates was "race-neutral" because the officials had canceled all the candidates' scores. On the statutory question, the petitioners maintained as their lead position that it is *never* permissible to engage in race-based disparate treatment in order to avoid a potential disparate-impact violation, because, for public employers, such disparate treatment violates the Constitution. As their fallback position, petitioners argued that if it is ever permissible for a public employer to engage in disparate treatment in order to avoid disparate impact, it can only be when the employer has a "strong basis in evidence" to believe that a disparate-impact violation will otherwise result. The petitioners' suggested "strong basis in evidence" standard was drawn from the Court's equal protection cases, such as *Richmond v. J.A. Croson Co.*,[29] which have held that a governmental actor wishing to take race-based action in order to remedy past racial wrongs must have a strong basis in evidence to support its belief that remedial action is required. Underpinning the petitioners' arguments, and that of many of their amici, was the compelling theme that it is an insult to individual dignity and the fundamental principle of equality for an employer to allow candidates to sacrifice mightily to perform well on a test and then throw out the test merely because of the raw racial numbers produced.[30]

The respondents continued to maintain that the Equal Protection Clause was not implicated at all because the cancellation of the results had been race-neutral, and also argued that even if the cancellation were race-based, Title VII compliance was necessarily a compelling state interest that could justify race-based action. On the statutory question, the respondents changed tack, abandoning their position that a mere good-faith fear of possible disparate-impact liability was sufficient to justify scuttling the promotions and arguing instead that the evidentiary record objectively demonstrated that the tests were flawed and that there were better, available alternatives. This was a difficult position to maintain, however, since the city had expressly conceded in the lower courts that it did not have an objective case either that the tests were flawed or that there were known, demonstrably better alternatives, and moreover because on summary judgment the petitioners were entitled to have the evidentiary record read in the light most to their favor. Underpinning all of the city's arguments, and also prominently figuring in the briefs of several of its amici, was the persuasive theme that adopting the petitioners' position would put employers into an impossible position where, having given a test that produced racially disparate results, they would be sued and exposed to liability no matter what action they took.

The Supreme Court reversed the Second Circuit in a 5-4 opinion written by Justice Kennedy, holding that New Haven had violated Title VII by discarding the test results and denying lieutenant and captain promotions to the highest-scoring candidates based on the test results' racial distributions. The majority opinion adopted the petitioners' proposed "strong basis in evidence" standard to resolve the conflict between Title VII's provisions. That standard, according to the majority, "limits... discretion to cases in which there is a strong basis in evidence of disparate-impact liability, but it is not so restrictive that it allows employers to act only when there is a provable, actual violation."[31] The majority reasoned that this standard best reconciled the various provisions of Title VII with one another, as well as with the background concerns of constitutional equal protection. The Court did not, however, reach the constitutional question, finding the statutory ground sufficient to resolve the case.

In light of this statutory standard, the Court held that the respondents' actions had, in fact, violated Title VII because the record conclusively failed to demonstrate a strong basis in evidence to believe that certifying the test results would have led to a disparate-impact violation. The city could have been liable for a disparate-impact violation only if the tests were not job-related and consistent with business necessity, or if plaintiffs had shown an equally valid, less discriminatory alternative. Here, however, the record showed that the city had hired an expert employment test consultant, IOS, which took extensive steps to develop and administer race-neutral examinations. Vincent Lewis, a witness at the CSB hearings who examined the tests and had firefighting experience (and who was himself African-American), testified that the questions were relevant

for both exams. Even the expert witness from IOS's competitor recommended that the CSB certify the examination results. Moreover, there was no record evidence of an equally valid and less-discriminatory testing alternative; the vague statements in the CSB hearings about possible alternatives were insufficient, and proposed alternatives suggested by the respondents (for the first time) in their Supreme Court briefing would have themselves violated Title VII and were thus not equally valid.

Justice Alito supplemented the majority opinion with a concurrence, in which he walked through the record evidence to tell the story of how race and politics impermissibly influenced and determined the respondents' decision not to certify the eligibility list, and rebutted the dissent's selective presentation of the evidentiary record.[32] Justice Scalia wrote a short concurrence "to observe that [the Court's] resolution of this dispute merely postpones the evil day on which the Court will have to confront the question: Whether, or to what extent, are the disparate-impact provisions of Title VII of the Civil Rights Act of 1964 consistent with the Constitution's guarantee of equal protection?"[33] The dissenters, in an opinion written by Justice Ginsburg, criticized the Court for "leaving out important parts of the story," such as the history of discrimination in firefighting.[34] The dissent also argued that the purported conflict between disparate-impact and disparate-treatment liability was illusory and criticized the majority for not remanding the case to the lower courts for application of the strong-basis-in-evidence standard.

In the short term, the *Ricci* decision became notable for the attention it received during the confirmation hearings of Judge Sotomayor. Several Senators asked Judge Sotomayor pointed questions about the Second Circuit's decision in the case, as well as the panel's handling of the case in an unpublished, per curiam, one-paragraph opinion. Moreover, two of the *Ricci* petitioners, Frank Ricci and Ben Vargas, testified in the confirmation hearings held by the Senate Judiciary Committee. It will be interesting to see whether this augurs a trend towards litigants who had a case before a Supreme Court nominee being called to testify during that nominee's confirmation hearings. In any event, however, Ricci's and Vargas's testimony was largely uncontroversial, and Justice Sotomayor was confirmed by a comfortable margin.

In the longer term, and more importantly, *Ricci* will have significant effects on employment-discrimination law, and it also leaves open important statutory and constitutional questions to potentially be resolved in some future case. Importantly, because the decision is grounded entirely in statutory construction, its effect extends to both public and private employers. In terms of employment practices and litigation, employers will have a lower liability risk when using appropriate pre-employment and promotional examinations. Raw racial-disparity statistics will not be sufficient to allow employers to act to avoid disparate impact, and the Court's opinion at least implies that they similarly will not be sufficient to prove disparate impact in any lawsuit brought by complaining minorities. As long as the employer can show that the employment examinations were a business necessity and job-related in their content and design, the employer will be able to effectively fight a lawsuit even if there is a racial disparity in the test results or business policies. The employers' lower liability risk acts as a security for the applicants who take a hiring or promotional examination. Applicants will not have to fear that the employers will discard the tests, studying for which requires considerable financial and personal expenses, whenever the results fail to satisfy a racial quota or provide the desired diverse outcome.

The ruling will also strongly encourage employers to take the necessary steps to ensure its examinations for hiring or promoting decisions are racially neutral *before* administering them. The Court specifically stated that "Title VII does not prohibit an employer from considering, before administering a test or practice, how to design that test or practice in order to provide a fair opportunity for all individuals, regardless of their race."[35] This preparation and thoughtfulness before the examination is administered could eventually become the employer's defense in a disparate-impact suit. Indeed, the extensive precautions taken by the City of New Haven, the NHFD, and IOS in the test-making process were weighed heavily in the Court's decision.

Additionally, certain practices that are currently prevalent in employment may come into serious question after *Ricci*. While employers may face less risk of liability for disparate impact, many affirmative action practices that have been accepted to increase diversity will face a higher risk of disparate-treatment liability. One such practice, identified by Roger Clegg, is colleges' rejection of finalist pools for hiring decisions when the pool lacks the racial diversity that the employers were seeking.[36]

A contrary consequence of *Ricci* may be that employers who simply wish to achieve raw racial balance in their employment and promotion numbers, whether to avoid being sued or to promote diversity, will have a significant incentive to avoid giving objective examinations altogether. Once a test is developed and given, under *Ricci*, an employer will need to have a very solid evidentiary record that the test is discriminatory before it can decide to throw it out. "But once that process has been established and employers have made clear their selection criteria, they may not then invalidate the test results, thus upsetting an employee's legitimate expectation not to be judged on the basis of race."[37]

Finally, and more broadly, the Court's decision in *Ricci* endorses the principle that intentional employment discrimination is a greater injustice than unintentional racial employment inequality, and that the law must incline towards preventing the former instead of the latter. This principle is certainly consistent with the Court's pronouncement in equal protection cases such as *Croson*, *Adarand Constructors, Inc. v. Pena*,[38] and *Parents Involved In Community Schools v. Seattle Sch. Dist. No.1*,[39] but *Ricci* takes a step forward in carrying this principle into the context of statutory, employment law.

Beyond these implications, *Ricci* raises several intriguing, unanswered questions. One is whether, and when, the Court will have to confront the lurking conflict between the disparate-impact provision of Title VII and the Constitution's promise of equal protection. Justice Scalia's concurrence was devoted solely to this point. As he notes, when an employer can ascertain with

certainty that certifying a test (or, more broadly, taking any given employment action) will have impermissible, unintentional disparate impact, then Title VII allows and indeed requires the employer to engage in intentional race-based action to avoid that disparate impact. Yet that action is, by definition, disparate racial treatment mandated by the government, which seemingly violates the Equal Protection Clause. As Justice Scalia aptly put it, "the war between disparate impact and equal protection will be waged sooner or later, and it behooves us to begin thinking about how—and on what terms—to make peace between them."[40]

Another important, unanswered question is whether diversity can be a compelling state interest in public employment. In *Grutter v. Bollinger*,[41] the Court held that diversity can be a compelling state interest in public higher education, and in *Parents Involved*, five justices (including Justice Kennedy) indicated their belief that diversity can be a compelling interest in elementary and secondary education. But the Court has never held that diversity is a compelling interest in public employment. The *Ricci* petitioners argued, albeit briefly, that public employment is materially different from education since the primary consideration should be effectiveness at doing the required job. Since the *Ricci* defendants did not assert diversity as a compelling interest to justify their throwing out the test results, and since the Court did not reach the equal-protection question, it remains to be settled whether diversity in employment is a compelling state interest that can justify race-based action by a government employer. The *Ricci* decision, however, may lead some to hope that the five justices in the *Ricci* majority would decline to so hold.

Ricci has shifted the field in employment law away from raw racial numbers, and toward a system that focuses on merit and qualifications to do the job. How far this trend will go, and whether the Court may in the future explicitly ground it not just in Title VII, but in the Constitution itself, remains to be seen, as the Court works through these questions in future cases and also as the composition of the Court changes in years to come.

Endnotes

1 129 S. Ct. 2504 (2009).

2 *South Carolina v. Katzenbach*, 383 U.S. 301, 308 (1966).

3 *E.g.*, Riley v. Kennedy, 128 S.Ct. 1970 (2008) (No. 07-77) (argument by Pamela S. Karlan).

4 Beer v. United States, 425 U.S. 130, 140 (1976).

5 42 U.S.C. § 1973c(a).

6 *Id.*

7 Connor v. Waller, 421 U.S. 656, 656 (1975) (per curiam).

8 *See* 42 U.S.C. §1973 b(a)(8).

9 42 U.S.C. §1973 b(b).

10 *See id.*

11 *See id.* The 1975 changes brought residents of Manhattan, Brooklyn, and the Bronx under Section 5. A complete list of covered jurisdictions is available at http://www.usdoj.gov/crt/voting/sec_5/covered.php.

12 *See* 42 U.S.C. §1973 b(b).

13 *See* Allen v. State Bd. of Elections, 393 U.S. 544, 566 (1969) ("The Voting Rights Act was aimed at the subtle, as well as the obvious, state regulations which have the effect of denying citizens their right to vote because of their race. Moreover, compatible with the decisions of this Court, the Act gives a broad interpretation to the right to vote, recognizing that voting includes 'all action necessary to make a vote effective.'").

14 Dougherty County, Ga., Board of Educ. v. White, 439 U.S. 32, 44 (1978).

15 435 U.S. 110 (1978).

16 *Id.* at 128-29 & n.15, 130-31 & n.18; *see also* Dougherty County, 439 U.S. at 43-44.

17 South Carolina v. Katzenbach, 383 U.S. 301, 334 (1966).

18 42 U.S.C. §1973 b(b).

19 521 U.S. 507 (1997).

20 Nw. Austin Mun. Util. Dist. No. 1 v. Holder, 129 S. Ct. 2504, 2508 (2009).

21 *Id.* at 2513-15.

22 *Id.* at 2511-12.

23 *Id.* at 2517 (Thomas, J., concurring in the judgment in part and dissenting in part).

24 129 S.Ct. 2658 (2009).

25 42 U.S.C. §2000e-2(a)(1).

26 42 U.S.C. §2000e-2(k)(1)(A)(i).

27 *See, e.g.*, *Wygant v. Jackson Bd. of Educ.*, 476 U.S. 267, 273 (1986) (plurality op.).

28 The group called themselves "the New Haven 20," and set up a website, www.NewHaven20.com, to gain support and explain their cause.

29 488 U.S. 469, 500 (1989).

30 Several of the petitioners—including lead petitioner Frank Ricci, who is dyslexic—had gone to extraordinary lengths and expense to prepare for the tests.

31 129 S.Ct. at 2676.

32 *Id.* at 2683 (Alito, J, concurring).

33 *Id.* at 2681 (Scalia, J, concurring).

34 *Id.* at 2689 (Ginsburg, J, dissenting).

35 *Id.* at 2676.

36 *See* Roger Clegg, "Dousing the Fires of Racial Discrimination," July 28, 2009, available at http://www.popecenter.org/clarion_call/article.html?id=2209.

37 129 S.Ct. at 2676.

38 515 U.S. 200 (1995).

39 127 S.Ct. 2738 (2006).

40 129 S.Ct. at 2683 (Scalia, J, concurring).

41 539 U.S. 306 (2003).

Note from the Editor:

The following two articles discuss *McDonald v. City of Chicago*, the highly publicized Supreme Court case from the October 2009 Term addressing the interaction of the Second Amendment of the U.S. Constitution with state regulation of gun ownership, and the Privileges or Immunities Clause of the Fourteenth Amendment. We chose to include both in this edition of Engage because of the extensive public interest in *McDonald* and because of the conflicting stances the authors take on the meaning and legacy of the *Slaughter-House Cases*.

Incorporating Gun Rights: A Second Round in the Chamber for the Second Amendment

*By Kenneth A. Klukowski**

The Supreme Court's October Term 2009 will see another major gun-rights case, the second in three years. Although the first case was undisputedly a watershed, from both a constitutional law perspective and from a societal-impact perspective, this second case will likely prove more consequential than the first.

The question presented in this case, *McDonald v. City of Chicago*, is whether the Second Amendment right to keep and bear arms is applicable to the states either through the Privileges or Immunities Clause or the Due Process Clause of the Fourteenth Amendment.[1] The lawyers bringing this case wisely took the opportunity to ask the Court to consider two alternative routes for "incorporation,"[2] creating the possibility for the Court to use this case as a vehicle to remediate aspects of incorporation doctrine and Fourteenth Amendment jurisprudence, as well as to extend to the states an important right enshrined in the Constitution's Bill of Rights.

This issue also benefits from the extraordinary caliber of circuit court judges who have developed the case law undergirding this case. Chief Judges David Sentelle, Alex Kozinski, and Frank Easterbrook, and Judges Diarmuid O'Scannlain, Laurence Silberman, Richard Posner, William Garwood, and Janice Rogers Brown, are among the jurists who have developed this issue in an exceptionally brief period of time, either since the *Heller* decision or in the years immediately preceding it. Their opinions—both regarding the nature of the right to bear arms and also on both sides of the incorporation question—have made this issue ripe for Supreme Court review, barely more than one year after the groundbreaking case that set the Second Amendment in motion in the federal judiciary.

I. *Heller* Redux

In 2008 the Supreme Court decided the landmark case of *District of Columbia v. Heller*,[3] presenting the question of whether the Second Amendment secured a right to private individuals to keep and bear firearms, versus merely some form of aggregate "right" of the people acting collectively, such as in organized National Guard units or some other form of state-controlled public service. The Court held that the right secured by the Second Amendment is indeed an individual right, consistent with the other enumerated rights declared in the Bill of Rights.[4] Accordingly, the Court struck down the law at issue, a D.C. statute that categorically banned handguns and other readily-usable firearms, even in the home,[5] affirming Judge Silberman's opinion for the D.C. Circuit.[6]

The majority opinion, written by Justice Scalia, was for the most part a sterling example of originalism.[7] That assessment is qualified with "for the most part" because there were some statements in the opinion that are problematic from an originalist viewpoint.[8] Most of these are ably explored by Professor Nelson Lund of George Mason—likely the foremost scholar today on the Second Amendment and whose argument may have helped win the *Heller* case[9]—in a recent law review article.[10] I also discuss what I regard as several problematic statements in my own law review article,[11] where I note that the impact of these problematic statements in *Heller* could be minor in that at least some of them are *obiter dicta*.[12]

As significant as the *Heller* decision was, its holding was nonetheless narrow. The facts in *Heller* were extreme, concerning essentially an absolute ban on firearm ownership, and the Court's opinion was appropriately tailored to resolve a case involving such extreme facts.[13] The Court held the Second Amendment secures an individual right, reasoning that the Amendment's prefatory clause (referencing a militia) must only be read in a fashion that does not restrict the scope of the operative clause (referencing the right to arms).[14] Thus *Heller* merely resolved the threshold issue on the right to bear arms;[15] had it held that there was no individual right in the Second Amendment, no further questions or cases on gun rights would be forthcoming.

Heller had been a long time coming. For many years, the only Supreme Court precedent clearly on point was *United States v. Miller*,[16] where in 1939 the Court remanded a gun-rights case for evidentiary development,[17] accompanied by a brief opinion containing various opaque statements about the nature of the Second Amendment so difficult to navigate that all sides of the gun-rights debate claimed that *Miller* supported their view. Finally, in 2001 the Fifth Circuit became the first circuit court to adopt the view that the Second Amendment

* *Fellow and Senior Legal Analyst, American Civil Rights Union. The author is also writing an amicus brief on behalf of Attorney General Edwin M. Meese III and the American Civil Rights Union in the upcoming Supreme Court case* McDonald v. Chicago. *Most of the material in this article is found in the author's recent law review article,* Citizen Gun Rights: Incorporating the Second Amendment Through the Privileges or Immunities Clause, 39 New Mexco Law Review 195 (forthcoming Nov. 2009).

secures an individual right in *United States v. Emerson*,[18] where Judge Garwood's opinion for the court engaged in a very long and thorough examination of the issue.[19]

The Ninth Circuit then wrote a lengthy opinion adopting the view that the Second Amendment confers no private right whatsoever in the 2002 case *Silveira v. Lockyer*,[20] which was essentially a rebuttal to *Emerson*.[21] When the full Ninth Circuit declined to rehear the case en banc, several judges wrote opinions dissenting from the denial,[22] including one by Chief Judge Alex Kozinski.[23]

Neither of these cases went before the High Court. But the case law they created set the stage, with a thorough examination of the literature and research that had been assembled over three decades. And while this case law reached its culmination in *Heller*, it marks only the beginning of the struggle over gun rights in America. Many questions remain, such as what level of scrutiny will attend Second Amendment claims,[24] a question that did not need to be decided in *Heller*,[25] and likewise is not at issue in *McDonald*. Indeed, a case almost identical to *Heller* was dismissed by a divided panel of the D.C. Circuit for lack of standing[26] (a fate that *Heller* avoided by a single vote),[27] showing that the question of standing is a critical issue going forward that cannot be taken for granted in any gun-rights case.[28]

This may in fact have been the reason that the Court granted certiorari in *McDonald v. Chicago*, but not *NRA v. Chicago*. These two cases had been consolidated at both the district and circuit level, and the Seventh Circuit decided both cases with one opinion bearing the NRA's name. Furthermore, the National Rifle Association was the first to petition for certiorari.[29] The only clear distinction between the two cases is that the *McDonald* plaintiffs applied for permits to possess firearms within Chicago city limits, and cited the denial of that permit as their particularized injury to satisfy the Article III case-or-controversy requirement,[30] while the *NRA* plaintiffs all claimed that their right to own a firearm within Chicago was being abridged, but did not assert any injury other than that which was suffered by the public at large, which the Supreme Court has held is insufficient to present a justiciable case.[31]

But of all the questions remaining after *Heller*, perhaps none is as consequential for the right to bear arms as whether the Second Amendment is applicable to the states through the Fourteenth Amendment.[32]

II. The Circuit Split Without a Split: When *Heller* Met *Agostini*

The Supreme Court in *Heller* had no occasion to consider whether the Second Amendment is incorporated against the states because the Bill of Rights directly applies to the District of Columbia as a federal enclave.[33] *Heller* was a test case deliberately brought in D.C. to avoid the incorporation question, so the Court expressly disclaimed the question, thus obviating the need to reconsider its precedent on the issue.[34]

In the 1876 case *United States v. Cruikshank*, the High Court held that the Second Amendment does not apply to the states through the Fourteenth Amendment.[35] This holding was then reaffirmed in *Presser v. Illinois*,[36] the 1886 precedent cited in Justice Sotomayor's per curiam decision on Second Amendment rights that was so often discussed in her confirmation hearings.[37] This proposition was reaffirmed yet again shortly thereafter in *Miller v. Texas*.[38]

The key question therefore becomes the breadth of the Court's holding in the *Cruikshank* line of cases. There are two possibilities, discussed in more detail below in Parts IV & V.

One is that *Cruikshank* and its progeny only considered whether the Second Amendment applies to the states through the Fourteenth Amendment Privileges or Immunities Clause. Even though the Court did not limit its holding to this clause, referring instead to the Fourteenth Amendment *in toto*, the 1876 decision was before the advent of substantive due process.[39] The argument therefore arises that the Court never contemplated the question of incorporating through the Due Process Clause, which would then leave that route open for applying the right to bear arms to the states. Professor Lund is the foremost advocate of this position,[40] which was also briefly adopted by the Ninth Circuit in an excellent opinion written by Judge O'Scannlain.[41]

The alternative position is that the *Cruikshank* line of cases precludes incorporation through any clause of the Fourteenth Amendment. Although it is accurate to argue that the Court's holdings in these cases were based solely on arguments involving the Privileges or Immunities Clause, such an argument is beside the point. The plain text of the Court's opinion encompassed all of the Fourteenth Amendment: the Court simply held that this amendment does not apply the right to bear arms to the states.[42] This is the position most consistent with the Supreme Court's most recent pronouncement on how inferior courts are to regard Supreme Court precedents that speak directly to an issue, but seem anachronistic.[43] Lower courts are therefore constrained to regard these precedents as foreclosing application of the Second Amendment to the states until the Supreme Court revisits the issue.[44] The Seventh Circuit recently adopted this very position, in an opinion by Chief Judge Easterbrook, written with his characteristic style.[45]

In terms of the Supreme Court, this question of the breadth of the *Cruikshank* holding determines whether stare decisis is an impediment to incorporating the Second Amendment. If *Cruikshank* is read narrowly as only concerning Privileges or Immunities, then the Court could freely incorporate the Second Amendment through the Due Process Clause while leaving those nineteenth-century precedents intact. If the Court takes the plain text of these antiquated opinions at face value and finds that they cover the entirety of the Fourteenth Amendment, then it would have to overrule all three cases to incorporate the Second Amendment.

Should the Court adopt the latter position, then it is worth noting that *Cruikshank* and its progeny are textbook examples of cases that are fit to be overruled.[46] The *Heller* Court expressly noted that *Cruikshank* also stated that the First Amendment does not apply to states, and further noted that these cases did not engage in any part of the analysis that the Court's subsequent case law requires for Fourteenth Amendment inquiries.[47] Thus the Court clearly signaled that *Cruikshank* and its progeny are deficient from a modern jurisprudential perspective.

Stare decisis requires courts to adhere to precedent absent some special justification for overruling it.[48] This principle is

predicated on the concept that, in the American common-law system, it is usually better for a rule of law to be decided, than to be decided correctly.[49] However, the Court has repeatedly held that the hurdle of stare decisis is not as difficult to clear when constitutional issues are at bar.[50] Further, the Court has held that a significant development in constitutional law can provide such a special justification.[51] Given that every Supreme Court case incorporating provisions of the Bill of Rights was decided after the *Cruikshank* trio of cases,[52] and that the entire framework of Fourteenth Amendment jurisprudence currently employed by the Court was developed in the twentieth century,[53] it seems clear that stare decisis should not impede overturning *Cruikshank*, *Presser*, and *Miller*.

But all this is beyond the purview of the circuit and district courts. The Supreme Court's precedents on this point are still controlling,[54] and the Court has recently reemphasized that inferior courts must consider themselves bound by Supreme Court precedent even when those precedents seem irreconcilable with current law, as only the High Court can overrule its own precedent.[55] Although the existence of a court split is usually a prime factor counseling in favor of granting certiorari, in the instant issue the Court should regard the lack of a split as still more compelling, as the circuits that have denied incorporation expressly claim that Supreme Court precedent forecloses the opportunity for the intermediate courts to even consider the question.

III. Obviating the Problem of Strict Scrutiny

The reality is that perhaps the single greatest impediment to incorporating the Second Amendment is strict scrutiny. The Supreme Court has only incorporated fundamental rights into the Due Process Clause, and in so doing has only applied rights that are fundamental to the states.[56] The general test for burdens on fundamental rights is strict scrutiny.[57]

Some may argue, therefore, that whether the Court is willing to incorporate will depend on uncoupling fundamentality from strict scrutiny. This is understandable, given that strict scrutiny is a daunting standard for laws to overcome. Laws subject to strict scrutiny are presumptively invalid,[58] and are only upheld if the government can carry the burden of showing that the challenged state action is narrowly tailored to advance a compelling state interest.[59]

Opponents of gun rights can argue that the Court's promulgation of a rule that gun control laws trigger strict scrutiny would lead to deranged individuals walking through public parks with assault rifles, violent individuals taking submachine guns onto school playgrounds, and criminal defendants carrying concealed weapons into courthouses. Given that there are many thousands of firearm laws in the United States between the federal, state, and local levels, it is not difficult to conclude that some Justices might be reluctant to subject every gun control law to a test that few laws survive.

This is especially significant in light of the presumptive invalidity of actions triggering strict scrutiny. There are over 200 million firearms in the United States, possessed by perhaps 90 million individuals throughout the fifty states and U.S. territories, under a patchwork legal framework of the thousands of laws referenced above.[60] The number of permutations for possible case fact patterns is essentially infinite. Shifting the burden from the challenger to the government sets a Herculean task before the government, requiring it to satisfy that burden in the multitudinous lawsuits that could arise. Strict scrutiny is called strict for a reason—it is usually fatal to the law at issue.

But this is a false choice. Strict scrutiny is not the uniform test for burdens on fundamental rights. It is simply the general test, and this provides a route to incorporate the Second Amendment's right to bear arms while circumscribing the parade of horribles mentioned above that opponents to gun rights will trot out in an effort to persuade the Court to refuse incorporation. For example, voting is a fundamental right for which only severe burdens are subject to strict scrutiny, with the remainder being held to a standard of reasonableness.[61] Burdens on Fourth, Fifth, and Sixth Amendment rights are generally not subject to strict scrutiny. Finding the right to bear arms to be fundamental does not necessitate applying strict scrutiny to every gun law.

The Court can instead begin establishing a multi-tiered framework of review, analogous to the one employed for free speech issues.[62] Content-based speech controls are subject to strict scrutiny.[63] Viewpoint-based discrimination is even more demanding, in that the rebuttable presumption of invalidity is elevated to an irrebuttable presumption, creating a per se rule that viewpoint discrimination is always unconstitutional.[64] In the opposite direction, content-neutral regulations on speech, such as those concerning the time, place, or manner of speech, are subject to intermediate scrutiny,[65] under which the law must be narrowly tailored to achieve a significant government interest. (This test is different, and more demanding, than intermediate scrutiny under the Equal Protection Clause.)[66] Speech on government land that is a limited public forum can be further restricted to force speech to conform to the public purposes of the forum.[67] And laws governing speech in a nonpublic forum, such as an airport, are subject to a test of mere reasonableness.[68] As the Court is presented with different forms and degrees of gun control laws, developing this multi-tier framework should satisfy public needs while upholding the right to keep and bear arms consistent with the design of both the Founding Fathers that adopted the Second Amendment and also the Framers of the Fourteenth Amendment.

Adapting the multi-tier framework employed under the Free Speech Clause is the optimal solution that satisfies all of these concerns.[69] The Court can then hold the right to bear arms to be a fundamental right and therefore applicable to the states, but without the concomitant issue of strict scrutiny mowing down every vestige of gun laws in the United States.

IV. Incorporation through the Due Process Clause

Every right applied to the states thus far has been applied by being incorporated into the Due Process Clause as a substantive right. Most seem to overlook the fact that incorporating rights though the Due Process Clause is a form of substantive due process.

The problem with substantive due process is that it is perhaps the most pernicious doctrine ever promulgated by judicial activism. It first became ascendant in the infamous *Lochner v. New York*.[70] But although *Lochner* has long since

been overruled,[71] the substantive due process spawned by that discredited precedent refuses to die along with its creator. Instead, it resurfaced in *Griswold v. Connecticut* (although *Griswold* was predicated on "penumbras from emanations" in the Bill of Rights, not the Due Process Clause).[72] But when this doctrine became resurgent with a vengeance in *Roe v. Wade* it cast aside any pretence of these spectral penumbras or ethereal emanations.[73] Instead, the Court simply proclaimed that the Due Process Clause, a provision that by its own diction is purely procedural, somehow implicitly contains substantive rights in its utterly non-substantive verbiage. This principle was then reaffirmed in *Planned Parenthood v. Casey*[74] and stated in completely explicit terms in *Lawrence v. Texas*.[75] All this precedent can be laid at the feet of substantive due process,[76] making it a dubious vehicle at best for applying the Bill of Rights to the states.

There is an additional complication with applying the right to bear arms through the Due Process Clause. The Due Process Clause, like the Equal Protection Clause, applies to every person in the United States. Aliens can avail themselves of these protections.[77] Even illegal aliens can claim many—if not all—of these rights.[78] To incorporate the Second Amendment right into the Due Process Clause would extend gun rights to some, if not all, of these noncitizens. While most aliens are law-abiding people who have the same self-defense concerns as any other person, the inherent deadliness of firearms and the reasons explained below should give pause to empowering aliens, possibly including those in the United States illegally, with the right to demand a gun.

V. The Privileges or Immunities Clause and the Rights of Federal Citizenship

The solution to the problems attending incorporation through the Due Process Clause is to return the application of substantive federal rights to the states to the clause of the Constitution designed for that purpose: the Privileges or Immunities Clause of the Fourteenth Amendment.

After lying dormant for many decades, the Privileges or Immunities Clause has recently reemerged in constitutional jurisprudence.[79] The Supreme Court resuscitated this provision from the Fourteenth Amendment in the 1999 case *Saenz v. Roe*.[80] Although dissenting in that case, Justice Thomas wrote for himself and the late Chief Justice Rehnquist that he would be willing to revisit Privileges or Immunities in a case that appropriately presented an opportunity to do so.[81] *McDonald v. Chicago* is such a case.

Whereas the Due Process Clause extends procedural protections against the states for all persons, by its very terms the Privileges or Immunities Clause extends such protections only to American citizens. This provision alone was designed to convey those substantive rights against the states.

Not all constitutional rights extend to every person. The most obvious example of this is voting. The Constitution specifies that voting, which is a fundamental right, is guaranteed to American citizens that are age eighteen, regardless of gender or race. Noncitizens have no right to vote, because it is a right of federal citizenship.

The Second Amendment contains two rights, one of self-defense and the other the right to keep the government in check.[82] The *Heller* Court expressly recognized the self-defense right, as this was the basis for the Court's holding.[83] The Court also suggested that it recognized the political right, referencing the right to protect oneself against public violence (in contradistinction to private violence),[84] and referencing the concern of the Framers that the government would disarm the citizenry to avoid being held to account,[85] making the Second Amendment a "safeguard against tyranny."[86]

The right to self-defense having been firmly established by *Heller*, little more need be written on that point; only the right to hold government in check by force of arms requires explication here. While *Heller* only references this right obliquely, case law from other courts delves into this issue in more depth.

The most powerful articulation of this principle comes from now-Chief Judge Alex Kozinski, dissenting from the denial of en banc in *Silveira v. Lockyer*. Chief Judge Kozinski's opinion declares that "the simple truth—born of experience—is that tyranny thrives best where government need not fear the wrath of an armed people."[87] He continues:

> The Second Amendment is a doomsday provision, one designed for those exceptionally rare circumstances where all other rights have failed—where the government refuses to stand for reelection and silences those who protest; where courts have lost the courage to oppose, or can find no one to enforce their decrees. However improbable these contingencies seem today, facing them unprepared is a mistake a free people get to make only once.[88]

This couplet of rights—one personal and the other political—is perhaps best explained by Judge Janice Rogers Brown. As a member of the California Supreme Court, then-Justice Brown noted that "[e]xtant political writings of the [founding] period repeatedly expressed a dual concern: facilitating the natural right of self-defense and assuring an armed citizenry capable of repelling foreign invaders and quelling tyrannical leaders."[89] Such writings, taken with the voluminous works on self-defense discussed throughout the *Heller* opinion, draw a picture of the Framers guaranteeing two rights in one constitutional provision.

The Second Amendment's use of the words "the people" (as in "We the People") strongly supports the proposition that this amendment reserves a political right only to citizens, and would have been understood as such in the 1790s when the Bill of Rights was adopted.[90] The term "the people" was also often used by Congress to refer to the American citizenry in the 1860s, when the Fourteenth Amendment was adopted.[91] Although this point can be confusing, given that "the people" is used in several places in the Bill of Rights for rights that all persons within the United States enjoy regardless of citizenship,[92] the Declaration of Independence and the Constitution also use the term in a way that clearly concerns only citizens.[93] The Second Amendment employs this term in the latter sense.[94]

VI. Incorporating through Privileges or Immunities Without Overruling the *Slaughter-House Cases*

Many would say that incorporating the Second Amendment through the Privileges or Immunities Clause is precluded by the *Slaughter-House Cases*.[95] Therefore, this argument goes, *Slaughter-House* would have to be overruled.[96]

This argument is incorrect.[97] It is difficult to think of any case where what the Court did is so profoundly different from what scholars say the Court did as the *Slaughter-House Cases*.[98] But the Court is not bound by any post-hoc gloss imposed on its precedent by the legal academy, and thus can freely remedy the situation.

In *Slaughter-House*, a group of Louisiana butchers challenged a state law granting a monopoly for the slaughtering of animals within city limits, alleging that this statute violated the "privileges or immunities" guaranteed by the Fourteenth Amendment.[99] The Court rejected this argument,[100] noting that states exercise police power to regulate public health.[101] The Court held that the Privileges or Immunities Clause secures rights derived from the U.S. Constitution,[102] and held that no such federal right was implicated in this case.[103]

For the Supreme Court to have held to the contrary would have been gross judicial activism. The Constitution is silent over butchering animals; there is no constitutional provision that invalidates state laws regulating the butchering trade within dense population centers. The plaintiffs in *Slaughter-House* were asking the Court to judicially invent a right out of the ether, and to use it to strike down an important public health law adopted by the people's elected legislators. The Court simply declined this invitation to open Pandora's Box; it did not eviscerate the Privileges or Immunities Clause.[104]

To the contrary, the Court in *Slaughter-House* articulated the test that the Privileges or Immunities Clause only applies to the states rights that inhere in federal citizenship.[105] The Court went on to comment in dicta that it was not defining those rights in that case,[106] but cited First Amendment rights of free assembly and seeking redress, as well as habeas corpus, as possibly among such rights.[107]

The Court's holding was thus quite narrow,[108] and unless the Court is again faced with a case where the petitioner seeks to have the Court find an implied fundamental right to be free of economic monopolies and strike down a law granting such a monopoly, *Slaughter-House* need not be overruled.

We saw this same phenomenon in connection with the Second Amendment. As noted above, for almost seven decades the only Supreme Court precedent on point was *United States v. Miller*. Both those supporting the proposition that the Second Amendment secured an individual right and those opposing it cited *Miller* as their authority. Many wondered what the Court could do with the Second Amendment without overruling *Miller*. Yet, in *Heller*, the Court relegated *Miller* to a legal footnote without discarding it (though Justice Kennedy derogated it with the comment that as a precedent it was "deficient"[109]), simply holding that *Miller* stands merely for the proposition that the Second Amendment extends to almost all firearms that can be carried by a person.[110]

Slaughter-House can in that regard become the new *Miller*. Although there is no need to call it deficient, its holding can instead be clarified to articulate this test of federal citizenship. Then, for the reasons discussed above and discussed in my law review article in much greater detail, the Court can simply hold that the right to bear arms is a right inhering in federal citizenship, and apply that right to the states through the Privileges or Immunities Clause while preserving *Slaughter-House*.[111]

There are three precedents that would have to be overruled to incorporate the Second Amendment: *Cruikshank*, *Presser*, and *Miller v. Texas* (not to be confused with *United States v. Miller*). While space constraints preclude discussing those cases in detail, for the reasons cited in Part II, it is sufficient to note that the Court has jettisoned the entire rationale underlying those cases, and that these precedents should be overruled.

But the talismanic *Slaughter-House* is not among them. The Second Amendment right to bear arms can be extended to the states without overruling the *Slaughter-House Cases*.

VII. The Risk in Overruling the *Slaughter-House Cases*

Some libertarians believe that the *Slaughter-House Cases* should be overruled. They argue that people should be able to challenge state and local employment and business laws in federal court. They welcome the opportunity to employ the federal judiciary to recognize economic rights devoid of textual support in the Constitution, urging courts to employ such rights to strike down onerous laws.

Conservatives disagree, and with good reason. There are reasons to preserve the *Slaughter-House Cases*. The law must rest on the application of neutral principles.[112] Judicial restraint requires not only that the courts not impose a liberal agenda or allow federal intrusions beyond the Constitution's enumerated powers, but, beyond that, to not impose any agenda, nor declare rights that are not enumerated. Conservative giants have propounded this principle, delineating the limited role of unelected judges in our democratic republic.[113] The *United States Reports* are likewise replete with these warnings from the Supreme Court, often quoting the Federalist Society's iconic figure, James Madison.[114] The states are the laboratories of democracy.[115] It is an unfortunate fact that the Constitution does not forbid states from passing stupid legislation. Should the Court invalidate state and local economic laws through the constitutionalizing of unenumerated economic rights, one of the critical remaining aspects of federalism will be forever abolished. Those calling for the overruling of *Slaughter-House* evidently fail to see this titanic downside risk, or ill-advisedly believe that they can control this genie to only arrive at "correct" results, once it is released from its bottle.

There is a reason that many calling for *Slaughter-House* to be overturned are squarely in the liberal camp. They see it as a cornucopia to accomplish through the courts everything they fail to achieve through the ballot box. It would facilitate a future Supreme Court declaring as among the "privileges or immunities" of U.S. citizenship the rights to healthcare, college education, "decent" housing, and a clean environment, attended by orders to enact everything from government-run healthcare to cap-and-trade. Many on the left are already laying

the predicate for such holdings.[116] As Justice Scalia cautioned when writing of rights that may be implicit in the Constitution, the lack of constitutional codification means that they should be debated and decided by the people's elected leaders; courts are not entitled to cite such rights to override the policy judgments of officials answerable to the electorate.[117]

Finally, some advance the argument that the Second Amendment right to bear arms should be incorporated through Privileges or Immunities, but extend to every person in America. They argue that self-defense is a human right, and that although it finds its locus in the Privileges or Immunities Clause, it can be extended to noncitizens.

This argument is a classic *non sequitur*, and a patently absurd one at that. The Privileges or Immunities Clause, by its own terms, applies only to citizens. The Court has recently reaffirmed as much.[118] The record is explicit. Only rights incorporated through the Due Process Clause extend to (almost) everyone. If it is a "privilege or immunity" of citizenship, then it only extends to citizens. Although individual states should grant generous gun rights to resident aliens to enable them to enjoy a means of self-defense, any such right must be a statutory right, not a constitutional right. An originalist interpretation of the Privileges or Immunities Clause will not allow extending the rights of citizenship to noncitizens as a constitutional entitlement. To argue otherwise is to sacrifice the neutral principle of historically-defensible originalism on the altar of results-oriented expediency, in derogation of the rule of law.

VIII. Conclusion

In some respects, whether to incorporate through the Privileges or Immunities Clause versus the Due Process Clause is a choice between stare decisis versus first principles. The Court has always applied federal rights to the states through the Due Process Clause, and it may opt to cleave to that approach.

But there are significant complications that arise from pursuing this substantive due process route when it comes to firearms, and the Court's entire Fourteenth Amendment jurisprudence has suffered as a result of its redirecting matters intended for the Privileges or Immunities Clause to the Due Process Clause instead.[119]

The Supreme Court should therefore take this historic opportunity to apply the Second Amendment right to bear arms to the states through the Privileges or Immunities Clause. Such a holding would be a bold step forward for originalism, as it would fulfill the designs both of the Founding Fathers who adopted the Second Amendment, and also the Framers of the Fourteenth Amendment who endeavored to extend the right to bear arms to the states.[120]

Endnotes

1 Petition for Writ of Certiorari at i, McDonald v. City of Chi. (No. 08–1521).

2 Strictly speaking, the term "incorporation" means to incorporate a provision of the Bill of Rights as a substantive right into the Due Process Clause of the Fourteenth Amendment. However, it has become the legal term of art for applying any federal right to the states through any provision of the Fourteenth Amendment, and it is used in that fashion in this article.

3 128 S. Ct. 2783 (2008).

4 *Id.* at 2799.

5 *Id.* at 2821–22.

6 *See* Parker v. District of Columbia, 478 F.3d 370 (D.C. Cir. 2007) (opinion of Silberman, J.).

7 *See* Kenneth A. Klukowski, *Citizen Gun Rights: Incorporating the Second Amendment Through the Privileges or Immunities Clause*, 39 N.M. L. Rev. 195, 201 (2009).

8 It is possible that it was Justice Kennedy who insisted that these statements, largely caveats that the Court was not calling into question laws that restricted certain firearms or prohibited them in certain places or for certain people, be inserted into the opinion. Another possibility is that Chief Justice Roberts wanted a clear majority opinion, rather than a fractured Court, so that it would be clear that the rule being promulgated was a holding of the Court. If so, then Justice Scalia would have had to qualify the broad statements in the opinion to keep everyone onboard and avoid separate partial concurrences. But no one can say with certainty that it was not Chief Justice Roberts who insisted on these provisos in the opinion, or perhaps one of the other Justices. Time may tell, if there are several more Second Amendment cases while the five Justices in the *Heller* majority are all still on the Court.

9 *See* Transcript of Oral Argument 5–6 (comments of Kennedy, J.), District of Columbia v. Heller, 128 S. Ct. 2783 (2008) (No. 07–290). Justice Kennedy refers to an argument for decoupling the two clauses of the Second Amendment, and specifies that this argument is not found in the "red brief" (the merits brief for respondents). Professor Lund's amicus brief was the only brief offering such an argument. *See generally* Brief for the Second Amendment Foundation as Amicus Curiae Supporting Respondent, *Heller*, 128 S. Ct. 2783 (No. 07–290). Although Justice Scalia's final opinion in *Heller* does not adopt this argument, instead going further with a broader argument that Justice Kennedy joined in full, Lund's argument is apparently what Justice Kennedy indicated he found persuasive during oral argument. This is also the conclusion of David Kopel, a fellow at the Cato Institute who is another leading scholar on the Second Amendment and who attended the oral argument. *See* David Kopel, *Oral Argument in D.C. v. Heller: The View from the Counsel Table*, The Volokh Conspiracy, http://volokh.com/2008/03/31/oral-argument-in-dc-v-heller-the-view-from-the-counsel-table/. This assertion in no way detracts from the excellent work of Respondent Dick Heller's legal team, especially its architect Robert Levy and its lead counsel Alan Gura. It is simply an acknowledgement that the method of constitutional interpretation suggested in Lund's amicus brief appears to have persuaded the fifth Justice in the case.

10 *See* Nelson Lund, *The Second Amendment,* Heller*, and Originalist Jurisprudence*, 56 UCLA L. Rev. 1343, 1352–56 (2009).

11 *See* Klukowski, *supra* note 7, at 201.

12 *Id.* at 201 n.59 (citing *Heller*, 128 S. Ct. at 2821 (dictum) (stating that the Second Amendment extends to law-abiding citizens that are "responsible")).

13 *Id.* at 201–02.

14 *Heller*, 128 S. Ct. at 2789–90.

15 Klukowski, *supra* note 7, at 201.

16 307 U.S. 174 (1939).

17 *Id.* at 183.

18 270 F.3d 203, 264–65 (5th Cir. 2001).

19 *See id.* at 218–60. The extraordinary detail in Judge Garwood's opinion was doubtless partially due to the fact that the comprehensive briefs submitted in the case were largely written by former Justice Department OLC chief Charles Cooper and the aforementioned Professor Lund.

20 312 F.3d 1052, 1092 (9th Cir. 2002).

21 Klukowski, *supra* note 7, at 200.

22 *See, e.g.*, *Silveira*, 328 F.3d at 583–85 (Kleinfeld, J., joined by Kozinski, O'Scannlain, and T.G. Nelson, JJ., dissenting from the denial of rehearing en banc).

23 *See infra* notes 87–88 and accompanying text.

24 Klukowski, *supra* note 7, at 202–03.

25 *See* Kenneth A. Klukowski, *Armed By Right: The Emerging Jurisprudence Of The Second Amendment*, 18 Geo. Mason U. Civ. Rts. L.J. 167, 185–86 (2008).

26 Seegars v. Gonzales, 396 F.3d 1248, 1255–56 (D.C. Cir. 2005). Now-Chief Judge Sentelle dissented in that case. *See id.* at 1256–57 (Sentelle, J., dissenting).

27 Of the six plaintiffs in the case, the appellate panel unanimously held that five lacked standing, and the sixth was found to have standing by a 2–1 vote. *See* Parker v. District of Columbia, 478 F.3d 370, 378 (D.C. Cir. 2007); *id.* at 402 n.2 (Henderson, J., dissenting). Indeed, that is why the case is called *Heller*, instead of *Parker*.

28 Klukowski, *supra* note 7, at 202.

29 The NRA's petition was filed June 3, 2009 and carries docket no. 08–1497, while McDonald's petition was filed on June 9, 2009 with the docket no. 08–1521.

30 *See* Davis v. FEC, 128 S. Ct. 2759, 2768 (2008); Lujan v. Defenders of Wildlife, 504 U.S. 555, 561 (1992).

31 *See* Frothingham v. Mellon, 262 U.S. 447, 488 (1923).

32 Klukowski, *supra* note 7, at 203.

33 *See* Pernell v. Southall Realty, 416 U.S. 363, 369–80 (1974) (applying the Seventh Amendment directly to D.C.).

34 District of Columbia v. Heller, 128 S. Ct. 2783, 2813 n.23 (2008).

35 92 U.S. 524, 553 (1876).

36 116 U.S. 252 (1886).

37 *See* Maloney v. Cuomo, 554 F.3d 56, 58–59 (2d Cir. 2009).

38 153 U.S. 535 (1894).

39 *Cf.* Klukowski, *supra* note 7, at 226.

40 *See* Nelson Lund, *Anticipating Second Amendment Incorporation: The Role of the Inferior Courts*, 59 Syracuse L. Rev. 185, 186–87 (2008).

41 *See* Nordyke v. King, 563 F.3d 439, 448, 457 n.16 (9th Cir. 2009), *vacated* Sept. 24, 2009 (No. 07–15763) (order vacating en banc submission pending Supreme Court decision on incorporation question).

42 *See Miller*, 153 U.S. at 538; Presser v. Illinois, 116 U.S. 252, 264–66 (1886); United States v. Cruikshank, 92 U.S. 524, 551 (1876).

43 Klukowski, *supra* note 7, at 254 (citing Agostini v. Felton, 521 U.S. 203, 237 (1997) (quoting Rodriguez de Quijas v. Shearson/Am. Express, Inc., 490 U.S. 477, 484 (1989))).

44 *Id.* at 254–55 & 254 n.524.

45 *See* Nat'l Rifle Ass'n of Am., Inc. v. City of Chi., 567 F.3d 856, 857–58, 860 (7th Cir. 2009), *petition for cert. filed*, 77 U.S.L.W. 3679 (U.S. June 3, 2009) (No. 08–1497).

46 Klukowski, *supra* note 7, at 253–54.

47 District of Columbia v. Heller, 128 S. Ct. 2783, 2813 n.23 (2008).

48 Payne v. Tennessee, 501 U.S. 808, 827 (1991).

49 Agostini v. Felton, 521 U.S. 203, 235 (1997) (quoting Burnet v. Coronado Oil & Gas Co., 285 U.S. 393, 406 (1932) (Brandeis, J., dissenting)).

50 Seminole Tribe v. Florida, 517 U.S. 44, 63 (1996) (citations omitted).

51 *Agostini*, 521 U.S. at 235–36.

52 *See* Klukowski, *supra* note 7, at 203 n.79 (listing all of the incorporation cases).

53 *See id.* at 207–12.

54 NRA v. City of Chi., 567 F.3d 856, 857–58 (7th Cir. 2009). *Contra* Nordyke v. King, 563 F.3d 439, 448, 457 n.16 (9th Cir. 2009), *vacated* Sept. 24, 2009 (No. 07–15763); Lund, *supra* note 40, at 186–87.

55 *See supra* note 44.

56 Klukowski, *supra* note 7, at 211.

57 *Id.* at 205–06.

58 *Id.* at 206 & n.113 (citing, inter alia, United States v. Am. Library Ass'n, 539 U.S. 194, 235 (2003)).

59 Gratz v. Bollinger, 539 U.S. 244, 270 (2003).

60 Klukowski, *supra* note 7, at 195 n.5 (citing various sources).

61 Burdick v. Takushi, 504 U.S. 428, 433–34 (1992).

62 *See* Klukowski, *supra* note 7, at 206 n.117.

63 *Id.* (citing Perry v. Perry Local Educators' Ass'n, 460 U.S. 37, 45 (1983)).

64 *See* Members of the City Council of L.A. v. Taxpayers for Vincent, 466 U.S. 789, 804 (1984).

65 Klukowski, *supra* note 7, at 206 n.117 (citing United States v. O'Brien, 391 U.S. 367, 377 (1968)).

66 *Id.* at 206 n.116 and accompanying text.

67 Rosenberger v. Rector & Visitors of Univ. of Va., 515 U.S. 819, 829–30 (1995).

68 Int'l Soc'y for Krishna Consciousness v. Lee, 505 U.S. 672, 679 (1992).

69 Klukowski, *supra* note 7, at 235.

70 198 U.S. 45 (1905).

71 *See* Ferguson v. Scrupa, 372 U.S. 726 (1963).

72 381 U.S. 479, 484 (1965).

73 *See* 410 U.S. 113, 129 (1973).

74 505 U.S. 833 (1992).

75 539 U.S. 558, 578 (2003).

76 Klukowski, *supra* note 7, at 204–05.

77 United States v. Verdugo-Urquidez, 494 U.S. 259, 271 (1990) (listing various cases).

78 *See* Klukowski, *supra* note 7, at 238 & nn.385–87 (citing various sources).

79 *Id.* at 197, 226.

80 526 U.S. 489, 499–500, 503 (1999).

81 *Id.* at 528–29 (Thomas, J., joined by Rehnquist, C.J., dissenting).

82 Klukowski, *supra* note 7, at 195.

83 District of Columbia v. Heller, 128 S. Ct. 2783, 2817 (2008).

84 *Id.* at 2799.

85 *Id.* at 2801.

86 *Id.* at 2802.

87 Silveira v. Lockyer, 328 F.3d 567, 569 (9th Cir. 2003) (Kozinski, J., dissenting from denial of rehearing en banc).

88 *Id.* at 570.

89 Kasler v. Lockyer, 2 P.3d 581, 602 (Cal. 2000) (Brown, J., concurring).

90 Klukowski, *supra* note 7, at 246 (citing, inter alia, Akhil Reed Amar, *The Bill of Rights as a Constitution*, 100 Yale L.J. 1131, 1176 (1991)).

91 Akhil Reed Amar, *The Bill of Rights and the Fourteenth Amendment*, 101 Yale L.J. 1193, 1226 & n.151 (1992).

92 Klukowski, *supra* note 7, at 246.

93 *Id.* at 245 (citing The Declaration of Independence, para. 2 (U.S. 1776)).

94 *Id.* at 244–48.

95 83 U.S. (16 Wall.) 36 (1873).

96 1 Laurence H. Tribe, American Constitutional Law 1321 & n.4, 1322 (3d ed. 2000); Laurence H. Tribe, *Taking Text and Structure Seriously: Reflections on Free-Form Method in Constitutional Interpretation*, 108 Harv. L. Rev. 1221, 1298 n.247 (1995).

97 Klukowski, *supra* note 7, at 228.

98 *Id.* at 230.

99 *Slaughter-House*, 83 U.S. (16 Wall.) at 60.

100 *Id.* at 66.

101 *Id.* at 82.

102 *Id.* at 74.

103 *Id.* at 81.

104 Klukowski, *supra* note 7, at 226.

105 *Id.* at 230; *accord* William E. Nelson, The Fourteenth Amendment: From Political Principle to Judicial Doctrine 162–63 (1988); Rebecca E. Zietlow, *Congressional Enforcement of Civil Rights and John Bingham's Theory of Citizenship*, 36 Akron L. Rev. 717, 746–49 (2003).

106 *Slaughter-House*, 83 U.S. (16 Wall.) at 74.

107 *Id.* at 79.

108 Klukowski, *supra* note 7, at 228, 230.

109 Transcript of Oral Argument at 62 (comments of Kennedy, J.), District of Columbia v. Heller, 128 S. Ct. 2783 (2008) (No. 07–290).

110 *Id.* at 2814.

111 Klukowski, *supra* note 7, at 232.

112 *See* Robert H. Bork, *Neutral Principles and Some First Amendment Problems*, 47 Ind. L.J. 1 (1971).

113 *See, e.g.*, Originalism: A Quarter Century of Debate (Steven G. Calabresi ed., 2007) (reprinting speeches of Edwin M. Meese III, Robert H. Bork, and Ronald W. Reagan).

114 *See, e.g.*, Lujan v. Defenders of Wildlife, 504 U.S. 555, 560 (1992) (quoting The Federalist No. 48 (James Madison)).

115 *See* New State Ice Co. v. Liebmann, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting).

116 *See, e.g.*, Jesse Jackson Jr., *Fundamental Right to Healthcare*, RealClearPolitics, Nov. 26, 2008.

117 *See* Troxel v. Granville, 530 U.S. 57, 91–92 (2000) (Scalia, J., dissenting).

118 *See* Saenz v. Roe, 526 U.S. 489, 502–04 (1999).

119 Klukowski, *supra* note 7, at 204 n.89 (citing John Hart Ely, Democracy and Distrust 20 (1980)).

120 *Id.* at 249–52.

Getting Beyond Guns: Context for the Coming Debate over Privileges or Immunities*

*By Clark M. Neily III** and Robert J. McNamara****

As it struggled to cope with the aftermath of the Civil War and to dismantle the system of human slavery that had both dominated and disgraced its early history, the United States adopted a trio of amendments designed to fulfill the promise of America as originally expressed in our founding documents, the Constitution and the Declaration of Independence. The Reconstruction amendments were specifically intended to reshape the relationship between government—federal, state, and local—and the people. And while an immediate goal of those amendments was to confer full and equal citizenship on newly freed African-Americans, they had a deeper, more profound purpose: to stamp out a culture of lawlessness and oppression that had grown up around the issue of slavery and attempts to abolish it, but that had grown like a cancer until it menaced the freedom of all citizens and the very notion of liberty upon which this country was founded.

While a tremendous victory, the Reconstruction amendments were not a *final* victory. The same debates over the scope of state power and states' relationships to the federal government that had raged before Reconstruction continued after the Amendments' ratification. While the Amendments represented the intellectual and legal triumph of Republican antislavery ideology, that triumph was in many ways short-lived—and, in the case of the Privileges or Immunities Clause, barely more than momentary.

Notwithstanding the imprecision with which it is frequently used, the term "judicial activism" does have a fixed meaning, namely, the substitution by a judge of his or her personal preferences for law. And that is precisely what happened in the *Slaughter-House Cases*, where a bare majority essentially announced that it considered unwise the Nation's decision to empower the federal government to enforce basic civil rights and would refuse to apply the Amendment insofar as it did so. That display of raw judicial power has deprived Americans of a properly engaged federal judiciary for more than a century.

This paper tells the story of the Privileges or Immunities Clause—its original purpose, its redaction by the Supreme Court, and its prospects for revival. The Supreme Court would do well to prepare for the challenges of the 21st century by correcting a particularly glaring mistake from the 19th. Properly understood, the Privileges or Immunities Clause speaks to a wide range of modern concerns—from gun control to property rights to occupational freedom—and provides a coherent framework for engaging those issues that is based on the text and history of the Constitution.

I. Slavery, Abolition, and the Shifting Balance of Power Between the Federal Government, the States, and the People

The Fourteenth Amendment represented a capstone—not just of the Civil War, but of a decades-long political struggle that sought to redeem the spirit of liberty from the crucible of slavery and its incidents. The Amendment can be neither understood nor interpreted without a proper appreciation of the historical dynamics that produced it, including particularly the specific evils the Amendment was designed to cure.

A proper understanding of the meaning of the Privileges or Immunities Clause has three basic components. First is the context in which the debates over the Fourteenth Amendment took place—the continuing struggle, dating back to the framing of the Constitution, over the relationship between the federal government, the states, and the people, who understood themselves to be *sovereign*. Second, one must understand what abolitionists and congressional Republicans were trying to accomplish, that is, the specific issues that gave rise to the Fourteenth Amendment. And, finally, one must look at what they actually produced, the Amendment's text and how it was crafted.

A. Pre-Civil War Debates

The U.S. Constitution was adopted as a significant change in its own right—a change meant to centralize more power in the federal government after the failure of the feeble authority created by the Articles of Confederation.[1]

In striking a new balance between federal power and state power, one question loomed large: slavery. In the original Constitution, the Framers largely punted on this question—while some implicit references to slavery (such as the notorious "three-fifths compromise" of Article I, sec. 2) were necessary, the terms "slave," "slavery," "human bondage" and the like do not appear anywhere in the document.[2]

The Framers' failure to address slavery and delineate the balance of power between the federal and state governments on that issue created a void in the Constitution with far-reaching implications. While everyone recognized that the new Constitution had created a stronger central government, there was much uncertainty about just how strong that government was and the precise bounds of its power vis-à-vis the states and the people.[3] One school of thought held that state governments retained the power to "nullify" federal laws they did not like.[4] Another, in part motivated by the Constitution's failure to grapple with the slavery problem, held that the Constitution itself was illegitimate.[5]

** This essay is a condensed version of a longer piece forthcoming in the Texas Review of Law and Politics, Volume 14, Issue 1. This piece is reprinted with their permission.*

*** Senior Attorney, Institute for Justice. In his private capacity, Mr. Neily was co-counsel for the Plaintiffs in District of Columbia v. Heller.*

**** Staff Attorney, Institute for Justice*

A third school of thought—of particular importance because it became the dominant view among many of the Reconstruction Republicans who would control Congress and propose the Fourteenth Amendment—held that the Constitution as drafted imposed substantive limitations on the states.[6] While a surprising and certainly difficult argument to accept through modern eyes, there can be no doubt that it was sincerely held at the time.[7] Though mistaken, the view that the Bill of Rights applied directly to the states was apparently fairly common,[8] while a more sophisticated view held that the Article IV Privileges *and* Immunities Clause protected substantive rights from state incursion.[9]

However sincerely held, those views had already been rejected by the Supreme Court. In *Barron v. Baltimore*, the Court held that the Constitution posed no barrier to a city's appropriating private property because the Fifth Amendment's takings provision (along with the rest of the Bill of Rights) had no application to the states.[10] And in *Dred Scott*, the Court adopted a narrow reading of the Article IV Privileges and Immunities Clause, finding that it only restrained states' ability to treat temporary visitors differently from residents, but imposed no requirements on what rights the states *denied* to different classes of citizens.[11]

Notwithstanding judicial setbacks, however, antislavery legal theorists continued to insist that the Constitution provided a meaningful check on state actions. For example, Joel Tiffany in his 1849 *Treatise on the Unconstitutionality of Slavery*, made an impassioned defense of a vision of the Constitution under which being a citizen of the United States was to be "invested with a title to life, liberty, and the pursuit of happiness," with United States citizenship providing a "a panoply of defense equal, at least, to the ancient cry 'I am a Roman citizen[]'" standing as a barrier to oppression by *any* government, including that of a state.[12]

It is worth noting that while Tiffany's theory of the scope of constitutional protection was a minority view, his use of the term "privileges" to describe substantive rights like freedom of speech was hardly unusual.[13] As Michael Kent Curtis notes, this usage "had a long and distinguished history," appearing in Blackstone's landmark *Commentaries on the Laws of England* before the American Revolution.[14] Even the reviled *Dred Scott* decision referred to the Bill of Rights as the "rights and *privileges* of the citizen."[15]

The view of many antislavery advocates that the Bill of Rights should be understood as binding state governments may have been wrong—that is, the *Barron* court may have been entirely correct in its interpretation of the Constitution—but it profoundly influenced later debates over the scope and significance of the Fourteenth Amendment nevertheless. As Yale professor Akhil Amar notes, the very phrase "bill of rights" became commensurate with the view that the first ten amendments to the Constitution were binding on the states—because, as declarations of *rights* (meaning *natural rights*), they could necessarily be asserted against any government.[16]

The Republican understanding of Article IV's Privileges and Immunities Clause—that it protected substantive rights against state infringement, not simply discrimination against nonresidents—was also shared by Ohio Republican Representative John Bingham. In 1859, speaking out against provisions in the proposed Oregon state constitution that would forbid free blacks from entering the new state, Bingham disputed the validity (or perhaps legitimacy) of both *Dred Scott* and *Barron*, arguing that free blacks were citizens of the United States and therefore held substantive rights protected by Article IV. His explanation of the Clause gives tremendous insight into the language that eventually became part of the Fourteenth Amendment, arguing that while there was "an ellipsis in the language employed in the Constitution," it was "self-evident" that it was meant to guarantee the natural rights of "citizens of the United States in the several States...."[17]

These are not simply the views of an ordinary Republican Congressman. While Bingham was active in the pre-Civil-War debates over the constitutional relationship between the states and the federal government, Bingham truly found fame several years later as the chief architect of the Fourteenth Amendment, taking the opportunity to correct the perceived "ellipsis" in Article IV's Privileges and Immunities Clause by filling in the missing text in the Fourteenth Amendment's Privileges or Immunities Clause.[18]

B. The Abuse, Redemption, and Surrender of Civil Rights in the Reconstruction South

As with any constitutional provision, the interpretation of the Fourteenth Amendment should be guided by a clear understanding of the specific evils the provision was meant to address.[19] In the case of the Fourteenth Amendment, the "mischief" that concerned Congress is easy to identify: state and local authorities throughout the South were systematically violating individual rights in open defiance of federal demands for full and equal citizenship for all. In 1866, Reconstruction Republicans undertook to set things straight.

The Fourteenth Amendment struck at three distinct "evils." First, it was meant to prevent states from locking newly freed slaves out of *political society*—an end accomplished by incorporating the Republican view that all people born within the United States were citizens thereof, effectively overruling the *Dred Scott* decision.[20] Second, the Fourteenth Amendment was meant to prevent states from discriminating against newly freed slaves by, for example, refusing to provide black citizens with police protection—a problem addressed by the requirement that no state deny any person within its jurisdiction equal protection of the laws.[21] Third, it was meant to prevent states from locking freedmen and others out of *civil* society by stripping them of certain rights—like the right to speak freely, to defend themselves, and to earn a livelihood in the occupation and on the terms of their choosing—that Reconstruction Republicans (and presumably most Americans) viewed as inherent in the definition of what it meant to be a free man.[22]

Republican concern for violations of civil liberties and natural rights did not start with the Reconstruction Congress. Indeed, the heated atmosphere of pre-Civil War debates over slavery and abolition effectively fused opposition to slavery with staunch support for civil liberties, as Southern states made clear that no individual right was sacred when it came to propping up the "peculiar institution," as states routinely prosecuted Republicans for the crime of circulating antislavery materials.[23]

And, of course the abuse of individual rights did not stop with the end of the Civil War or the adoption of the Thirteenth Amendment—to the contrary. Legislative testimony and newspaper accounts provide compelling evidence concerning the scope and intensity of the assault on civil liberties during Reconstruction.

The stories are legion. Discharged Union soldiers were forcibly stripped of their weapons; South Carolina law prescribed flogging for any black man who broke a labor contract; other laws prevented blacks from practicing trades or even leaving their employer's land without permission; minors in Mississippi were "taken from their parents and bound out to the planters"; white Union sympathizers often had their property seized or found themselves banished from a state outright.[24] In one Kentucky town, it was reported that the "marshall [took] all arms from returned colored soldiers and [was] very prompt in shooting the blacks whenever an opportunity occur[red]," while outlaws made "brutal attacks and raids upon freedmen, who [were] defenseless, for the civil law-officers disarm the colored man and hand him over to armed maurauders."[25] These acts were widely reported, fostering outrage not just in Congress, but throughout the popular press.[26] For many, if not most freedmen, life as a "free" man cannot have seemed much better than life as a slave.[27]

While it may be tempting to see these outrages as an ugly but isolated moment in our Nation's history, they are not. To the contrary, in America as everywhere else, those with power have always abused it, and the simple freedom to go about one's business unmolested and enjoy the fruits of one's labor is perpetually insecure. The Fourteenth Amendment, referred to by Justice Swayne in his *Slaughter-House* dissent as part of America's "new Magna Carta,"[28] was a very deliberate attempt secure that freedom.

C. Framing the Fourteenth Amendment

Congress in 1866 was considering several concurrent measures to address the twin problems of Reconstruction and the re-admittance of Southern states to the Union. Those measures included the bill that became the Civil Rights Act of 1866 and the various drafts of what would eventually become the Fourteenth Amendment.[29] Given the overlapping character of and motivations behind these measures, the debates surrounding them may be treated as a single coherent conversation over the central question of how to secure individual rights from predation by state and local governments.

The Fourteenth Amendment was largely drafted and guided by John Bingham, an Ohio congressman and moderate Republican whom "the *New York Times* described as 'one of the most learned and talented members of the House.'"[30] Bingham's leadership is important for several reasons, not least of which because his views explain why the debates over the Civil Rights Act are every bit as relevant to the proper interpretation of the Fourteenth Amendment as the debates over the Amendment itself. Many Congressional Republicans, given their unorthodox theory of the Constitution, believed (mistakenly) that the federal government already had all the power it needed to protect rights in the states.[31] But Bingham understood that that was not so, and he also recognized that without some sort of enabling amendment to the Constitution, the Supreme Court might well invalidate the Civil Rights Act as well.[32]

While many members of Congress appeared unaware (or unwilling to acknowledge) that the Supreme Court had long ago rejected their theory of constitutional supremacy over the states, Bingham was all too aware of those decisions, and he deliberately framed the Fourteenth Amendment as a response to *Barron*, specifically responding to that opinion's admonishment that provisions that were to limit the powers of state governments should clearly read that "no State shall...."[33]

Debates over what became the Fourteenth Amendment are replete with the natural-rights language that Republicans had used for decades in arguing against slavery.[34] Having been unable to respond effectively to state predations against natural rights before the Civil War, Reconstruction Republicans were intent on remedying what they considered a flawed constitutional rule that rendered the federal government powerless to stop those abuses as they continued after the war.

Throughout the 1866 debates, congressmen drew clear distinctions between their concern about *equality*—a concern that state laws be even-handed—and their concern about protections of *substantive rights*. Representative Thayer, for example, praised the Fourteenth Amendment as "so necessary for the equal administration of the law" *and* as "so necessary for the protection of the fundamental rights of citizenship."[35]

That distinction is essential to a proper understanding of the Privileges or Immunities Clause.[36] After all, as Michael Kent Curtis has observed, "in the South, the ideal solution to the problem of speech about slavery was compelled silence"—fully applicable to blacks and whites equally.[37] Thus, far from being concerned only with equality, congressional Republicans wanted to prevent states from violating "guaranteed privileges" like the right to speak out against slavery or cruel or unusual punishment,[38] and to reaffirm and protect certain "inalienable rights, pertaining to every citizen, which cannot be abolished or abridged by State constitutions or laws"—even ones that operated even-handedly.[39]

It was also very much the Framers' intent to ensure that *federal courts* would actively restrain state action. Representative Bingham discussed at length the Supreme Court's decision in *Barron*, citing it as evidence that "the power of the Federal Government to enforce *in the United States courts* the bill of rights under the articles of amendment to the Constitution had been denied."[40] Bingham's position was hotly disputed by Robert Hale, who insisted that the Bill of Rights already restrained state legislation but who acknowledged, in response to Bingham's challenge to name any court decision protecting liberty from state encroachment under the Bill of Rights, that he had "somehow or other" gotten that idea but could not identify any cases supporting it.[41]

The understanding of the Amendment expressed in the House of Representatives was typical of the understanding nationwide. These sentiments were echoed in the Senate, where Senator Jacob Howard relied extensively on Justice Bushrod Washington's opinion in *Corfield v. Coryell*,[42] to illustrate the natural rights or "fundamental guarantees" that were encompassed in the term "privileges and immunities."[43] The same understanding can be found in the state-level debates

over the Amendment's ratification[44] and those expressed in newspaper articles and editorials.[45] And legal scholars took the same view. Three significant legal treatises were published between the proposal of the Fourteenth Amendment and its ratification, each of which took the position that the Privileges or Immunities Clause would protect substantive rights of American citizens.[46]

In short, the congressional leadership intended to bring the Constitution in line with longstanding Republican ideology about national citizenship and natural rights, and to protect those rights from further violation at the hands of state and local officials. And the public appears by all accounts to have understood the proposed Fourteenth Amendment that way as well.[47] (If there is a credible historical counter-narrative, it has yet to be offered.) Thus, the notion that we lack the means to properly understand the Privileges or Immunities Clause of the Fourteenth Amendment is a fiction, and a rather shabby one at that.

II. The Supreme Court's Breathtaking Judicial Activism in *Slaughter-House*

The initial battles over the Privileges or Immunities Clause—during its drafting and ratification—were clear victories for proponents of federal protection for natural rights. But just seven years later, that vision was dealt a shocking blow by a narrow majority of the Supreme Court determined to substitute its preference for what we today would call minimalism over the expressed will of the people.

That blow, of course, was delivered by the Court in the infamous *Slaughter-House Cases*.[48] At issue in *Slaughter-House* was the constitutionality of a Louisiana law granting an exclusive monopoly on the right to sell and slaughter animals in New Orleans to a single politically connected company. Local butchers could continue to practice their trade under the law, but they could do so only in facilities operated by, and upon payment to, the government-favored monopolist.[49]

To the butchers, the creation of a state-sanctioned monopoly seemed an obvious violation of the Privileges or Immunities Clause, which they understood as protecting their right to earn a living free from unreasonable (including, obviously, corrupt) government interference. Just as the Black Codes had bound freedmen to an employer's land, imposed onerous contractual terms on their labor, and even barred them from participating in particular trades, the butchers viewed the challenged law as a direct affront to their livelihoods. The Supreme Court disagreed with that premise as a factual matter; as Justice Miller explained, "a critical examination of the act hardly justifies [the butcher's] assertions."[50] But instead of stopping there, as the doctrine of constitutional avoidance would have counseled, the majority went on to construe the Fourteenth Amendment, in what is arguably dicta, as an essentially meaningless provision that "most unnecessarily excited Congress and the people on its passage."[51]

In Justice Miller's opinion for the 5-4 majority, the Court posits a dichotomy of rights—those that are held by virtue of one's state citizenship on the one hand, and those that are held by virtue of one's national citizenship on the other—the rather obvious purpose of which is to disclaim any responsibility (or even authority) on the part of the federal government for protecting precisely those rights whose wanton violation by state governments was the driving force behind the enactment of the Fourteenth Amendment generally and the Privileges or Immunities Clause in particular.

The tenor of the opinion is striking, as it makes clear that its crabbed interpretation rests on a basic disapproval of the Amendment's purpose; that is, the Court effectively read the Privileges or Immunities Clause out of the Constitution because the "consequences" of reading the Clause properly would be so "radical."[52] The opinion's hostility to the Reconstruction Congress and its aims is barely masked, as Justice Miller only briefly notes the exploitative economic restrictions imposed on freedmen before suggesting that the congressional hearings were tainted with "falsehood or misconception... [in] their presentation."[53]

Rather than read the Privileges or Immunities Clause as working a significant change in the constitutional order—which it was explicitly intended and understood to have done by those who drafted and ratified it—the Court viewed the Clause as protecting only a narrow set of rights of "national citizenship," including "the right to use the navigable waters of the United States" and "the right of free access to... the subtreasuries, land offices, and courts of justice in the several States."[54] While some modern advocates have attempted to rehabilitate *Slaughter-House*, arguing that Justice Miller's opinion does not foreclose reading the Privileges or Immunities Clause to protect certain rights,[55] the opinion itself is clear on this point: it draws a distinction between rights whose very existence depends on the federal government (like access to its subtreasuries) and rights that had hitherto been the responsibility (no irony intended, at least by Justice Miller) of the states, making clear that the latter were "not intended to have any additional protection by this paragraph of the [A]mendment."[56] In short, the Privileges or Immunities Clause was rendered an essentially dead letter (though of course the possibility remained that it might one day be pressed into service by someone who is seeking access to a seaport or navigable waterway[57]).

Justice Stephen Field wrote a powerful dissent in which he chided the majority for rendering the Privileges or Immunities Clause "a vain and idle enactment, which accomplished nothing."[58] Field acknowledged the state's interest in public health, but, unlike the majority, recognized that there was a difference between the proper exercise of the state's police power to control where and how animals are slaughtered and the grant of an exclusive monopoly to one corporation. Noting that the law contained provisions prohibiting slaughtering animals in certain areas and requiring inspection of all animals to be slaughtered, Justice Field correctly observed that there was no *additional* public-health concern that would justify the creation of the slaughter-house monopoly.[59]

Having dispensed with the portions of the law that were unquestionably legitimate, Justice Field turned to the Privileges or Immunities Clause itself. In a thorough study of the context in which the Clause was adopted and the history upon which it drew (a context and history that the majority simply ignored), Justice Field noted the obvious linguistic similarity to the Privileges *and* Immunities Clause of Article

IV, and, relying (as did Congress in framing the Amendment) on Justice Bushrod Washington's explanation of privileges and immunities in *Corfield v. Coryell*, concluded that the new Privileges *or* Immunities Clause prevented states from violating the same basic rights identified in *Corfield*.[60] This, of course, included the traditional common-law abhorrence of monopolies as a violation of the right of all citizens to the "pursuit of the ordinary avocations of life."[61]

Despite compelling dissents by Justices Field, Bradley, and Swayne that utterly demolished the majority's reasoning (if it may be called that), *Slaughter-House* was universally understood as having effectively eliminated the Privileges or Immunities Clause as a source of meaningful protection for individual rights.[62] Of course, this was warmly received by opponents of the Fourteenth Amendment, many of whom applauded their fellow travelers on the Court for undoing what they viewed as a national mistake in empowering the federal courts to strike down state laws that interfered with citizens' basic civil rights.

What is striking, given the breadth and ideological diversity of modern scholarship, however, is the consensus of opinion that has emerged: simply put, nearly "everyone" now agrees that *Slaughter-House* misinterpreted the Privileges or Immunities Clause.[63] As described by historian Eric Foner, the *Slaughter-House* majority's conclusions "should have been seriously doubted by anyone who read the Congressional debates of the 1860s."[64] And Professor Thomas McAfee has observed that "this is one of the few important constitutional issues about which virtually every modern commentator is in agreement."[65] Moreover, even the few scholars who defend *Slaughter-House* do so not on the merits, but rather on overtly pragmatic grounds—i.e., that reinvigorating the Privileges or Immunities Clause would have undesirable consequences such as requiring judicial protection for currently disfavored rights like private property and occupational freedom—the very same grounds upon which the majority based its decision in the *Slaughter-House Cases*.[66]

But *Slaughter-House* did more than just misinterpret the Privileges or Immunities Clause. It fundamentally warped the Supreme Court's jurisprudence of rights in a manner that persists to this day. Having defied the will of the people by draining the Fourteenth Amendment of any real force, the Court left itself in the untenable position of either standing by while state and local officials continued to trample basic civil rights, or figuring out some way to sidestep its original mistake. And that was how substantive due process was pressed into service to protect an increasing number of rights deemed sufficiently "fundamental" by the Supreme Court.

Unfortunately, the Court's unnecessary reliance on substantive due process has had a number of negative consequences for individual-rights jurisprudence. First is the obvious tension between "substantive" and "process," which prompted John Hart Ely's comparison of "substantive due process" to "green pastel redness."[67] By contrast, the term "privileges or immunities"—which 19th-century Americans appear to have used interchangeably with "rights"[68]—needs no gloss or embellishment to do its job.[69]

Strengthening the ties between the Court's jurisprudence and the Constitution's actual text and history would not only increase the perceived legitimacy of the Court's individual-rights jurisprudence, it would give *content* to that jurisprudence. Because the debates and contemporaneous public documents surrounding the Fourteenth Amendment are replete with references to specific doctrines and even court cases the Framers meant to overturn, along with the specific evils they meant to prevent, the rights protected by the Privileges or Immunities Clause can be rooted solidly in both text and history, as can their limits.[70] The Clause is neither a meaningless nullity nor a freewheeling source of rights pulled from thin air. Relying on the Privileges or Immunities Clause would both help the Court outline the contours of its role in protecting individuals from rights violations by state governments and make that role more stable and difficult to assail.

In short, the Supreme Court read the Privileges or Immunities Clause out of the Constitution, not because of any genuine lack of clarity about what the Clause was meant to do, but simply because the Court found unsettling the change in federal-state relations that the Clause enacted. But that is obviously not a solid basis for principled jurisprudence.

III. Prospects for the Future

Why does any of this matter? The debates over the Fourteenth Amendment and the Supreme Court's evisceration of the Privileges or Immunities Clause came more than a century ago. The butchers who brought the *Slaughter-House Cases* are long-dead. But the issue remains alive today—in large part because the Supreme Court's misreading of the Privileges or Immunities Clause continues to have a direct impact on people's lives.

The abandonment of any meaningful judicial protection for economic liberty has yielded predictable, and tragic, results. In the 1950s, only 4.5% of the workforce needed a government license in order to do their job—these were largely doctors, lawyers, architects, and similar professionals. Today, nearly 30% of the workforce needs the government's permission in order to earn a living.[71] Rather than protecting public health or safety, these new licensing requirements often serve as nothing more than a means to lock a politically disfavored group out of a portion of the economy in order to allow politically favored groups to earn higher profits.[72]

The Supreme Court's incentive to reconsider *Slaughter-House* has been diminished by the fact that it has already "incorporated" most of the substantive protections of the Bill of Rights against the states using the doctrine of substantive due process.[73] The Court has also protected a number of unenumerated rights through that doctrine,[74] though many—including the right to earn a living—have been relegated to "nonfundamental" status, meaning they are recognized but not meaningfully protected. The ideal test case, then, is one presenting an indisputably fundamental, preferably enumerated right that has never been incorporated against the states.[75] The right to keep and bear arms fits that bill perfectly.

In its 2008 landmark decision *District of Columbia v. Heller*, the Supreme Court held for the first time that the Second Amendment protects an individual right to keep

and bear arms.[76] That decision resolved a long-standing and contentious constitutional debate,[77] but it left open a pressing question—given that the Second Amendment protects a right to keep and bear arms against infringement by the *federal* government,[78] does the Constitution prevent state and local governments from infringing the right to keep and bear arms, and if so how? The Court has now taken up that question[79] in a case challenging Chicago's handgun ban where the question presented asks "[w]hether the Second Amendment right to keep and bear arms is incorporated as against the States by the Fourteenth Amendment's Privileges or Immunities or Due Process Clauses."[80]

A candid, originalist reexamination of the Supreme Court's Fourteenth Amendment jurisprudence indicates that the Court has been protecting some rights (like free speech) incorrectly, by "incorporating" those rights through the Due Process Clause rather than simply recognizing them as among the inherent rights shared by all Americans and protected by the Privileges or Immunities Clause. Other rights, like economic liberty, have been all but ignored, despite their centrality to the Fourteenth Amendment's entire purpose, namely, the practical (not merely formal) elimination of servitude through the protection of those very rights necessary to overcome it.

This means that what the Supreme Court does with the gun-control question has consequences that run far deeper than gun regulations. As demonstrated above, the record is abundantly clear that the Privileges or Immunities Clause was meant to protect a right to armed self-defense by preventing the sort of forcible disarmament that became all too common in the Reconstruction South. But it is equally clear that the Clause is meant to protect other rights, like the right to earn an honest living in the occupation of one's choice, a right that most Americans—but unfortunately not most Supreme Court Justices—recognize as being among the most fundamental rights we possess.

CONCLUSION

The Fourteenth Amendment marked a watershed moment in American history, when the people of this country made a conscious decision to reject the fiction that state and local governments could, by virtue of their relative proximity to the polity, be entrusted with the protection of basic civil rights. The Fourteenth Amendment was the product of that decision, and it included a very conscious decision by the people of this country to charge the federal government not only with the power but the duty to protect a wide variety of individual rights from state governments. Unfortunately, in a breathtaking display of activism—and on the basis of a decision so profoundly flawed that it has been rejected by all serious constitutional scholars—the Supreme Court chose to defy the will of the people and forestall the constitutional revolution that culminated in the ratification of the Fourteenth Amendment. The Court has yet to honestly confront that mistake or fully acknowledge its initial refusal to implement the will of the people as expressed in their founding document. It is high time to do both by restoring the Privileges or Immunities Clause to its proper constitutional role. And while we cannot know exactly where that path might lead, there has never been any reason in this country to fear fidelity to the Constitution.

Endnotes

1 Laurence H. Tribe, American Constitutional Law, 6.3 at 404; *see also, e.g.,* The Federalist No. 11 ("A unity of commercial, as well as political, interests, can only result from a unity of government.").

2 *See generally* Lysander Spooner, *The Unconstitutionality of Slavery* (1845), *available at* http://www.lysanderspooner.org/UnconstitutionalityOfSlaveryContents.htm.

3 *But see* U.S. Const. amend. X.

4 *E.g.,* John C. Calhoun, *Speech on the Force Bill* (1833), in *Union and Liberty: The Political Philosophy of John C. Calhoun* 401, 428-29 (Liberty Fund 1992).

5 *See* Lysander Spooner, *No Treason, No. VI: The Constitution of No Authority* (1870), in 1 The Collected Works of Lysander Spooner (Charles Shively ed., 1971).

6 Michael Kent Curtis, *No State Shall Abridge* 43-56 (1986).

7 *See infra*.

8 Indeed, at one point a seemingly frustrated Representative John Bingham literally read aloud from *Barron v. Baltimore* in an attempt to persuade some of his doubtful colleagues that "the power of the Federal Government to enforce in the United States courts the bill of rights... had been denied." Cong. Globe, 39th Cong., 1st sess. 1089-90 (1866).

9 Curtis, *supra* note 6, at 47-48.

10 Barron v. Baltimore, 32 U.S. (7 Pet.) 243, 248 (1833).

11 Dred Scott v. Sanford, 60 U.S. (10 How.) at 422-23.

12 Joel Tiffany, *A Treatise on the Unconstitutionality of Slavery* 55 (1849) (cited in Michael Kent Curtis, *supra* note 6, at 42-43.

13 Akhil Reed Amar, *The Bill of Rights* 166-69 (1998).

14 Curtis, *supra* note 11, at 64.

15 60 U.S. (10 How.) at 403 (emphasis added).

16 Amar, *supra* note 13, at 286-87 (noting that the phrase "Bill of Rights" hardly ever appeared in antebellum congressional debates, and was the exclusive domain of Republicans in the debates over the Fourteenth Amendment).

17 Cong. Globe, 35th Cong., 2d Sess., 984 (1859).

18 *See infra*.

19 *Cf.* Rhode Island v. Massachusetts, 37 U.S. (12 Pet.) 657, 723 (1838) ("In the construction of the constitution, we must... examine the state of things existing when it was framed and adopted to ascertain the old law, the mischief, and the remedy.") (citation omitted).

20 *See* Don E. Fehrenbacher, *The Dred Scott Case* 579-81 (1978). The integration of freed slaves into *political* society was, of course, not complete until the introduction of the Fifteenth Amendment two years after the introduction of the Fourteenth Amendment.

21 *See, e.g.,* David Currie, *The Constitution in the Supreme Court: The First Hundred Years* 349 (1985).

22 *Cf.* Kimberly C. Shankman & Roger Pillon, *Reviving the Privileges or Immunities Clause to Redress the Balance Among States, Individuals, and the Federal Government*, 3 Tex. Rev. L. & Pol. 1 , 25-26 (1998).

23 State v. Worth, 52 N.C. 488, 492 (1860); *see also* Clement Eaton, *The Freedom of Thought in the Old South* 245 (1940) (detailing an indictment in Virginia for distribution of the same book).

24 *See* David T. Hardy, *Original Popular Understanding of the 14th Amendment as Reflected in the Print Media of 1866-68*, at 8-12 (2009), available at http://ssrn.com/abstract=1322323.

25 House Ex. Doc. No. 80, 39th Cong., 1st Sess. at 236-239 (1866).

26 *See* Stephen Halbrook, *Freedmen, the Fourteenth Amendment, and the Right to Bear Arms, 1866-1876*, at 7, 18, 31, 37 (1998) (discussing contemporaneous press coverage).

27 *Cf.* Michael Kent Curtis, *Resurrecting the Privileges or Immunities Clause and Reviving the Slaughter-House Cases without Exhuming Lochner: Individual Rights and the Fourteenth Amendment*, 38 B.C. L. Rev. 1, 72 (1996) (noting the Reconstruction South's additional abuse of the right to "free speech, the right to hold religious meetings[,] and the right to bear arms.")

28 83 U.S. (16 Wall.) at 125 (Swayne, J., dissenting).

29 Curtis, *supra* note 6, at 57-58.

30 *Id.* at 58.

31 For example, Senator Richard Yates from Illinois presumably spoke for many of his colleagues when he expressed surprise (perhaps feigned) that the question of federal power to protect individuals from state governments was even being debated: "I had," he said, "in the simplicity of my heart, supposed that 'State rights' being the issue of the war, had been decided." Cong. Globe, 39th Cong., 1st Session at 99 app. (1866).

32 *See* Curtis, *supra* note 6, at 80-81.

33 Amar, *supra* note 13, at 164-65 (1998) (quoting Cong. Globe, 42d Cong., 1st Sess. 84 app. (1871) (emphasis altered)).

34 *E.g.*, Cong. Globe, 39th Cong., 1st Session at 1064 (1866) (statement of Rep. Frederick Woodbridge) (stating that the proposal would give the federal government the power to "give to a citizen of the United States the natural rights which necessarily pertain to citizenship").

35 Cong. Globe, 39th Cong., 1st Session at 2465 (1866).

36 Some academics have argued that the Clause was meant only to require *equality* of rights, rather than to protect individual rights from infringement. *See, e.g.,* John Harrison, *Reconstructing the Privileges or Immunities Clause*, 101 Yale L. J. 1385 (1992).

37 Curtis, *supra* note 27, at 47.

38 Cong. Globe, 39th Cong., 1st Sess. 2542 (1866) (statement of Rep. Bingham).

39 210 Cong. Globe, 39th Cong., 1st Sess. 1832 (1866) (statement of Rep. Lawrence); *see also* Curtis, *supra* note 39, at 44-65 (describing and debunking what Curtis calls the "equality only" view of Privileges or Immunities).

40 Cong. Globe, 39th Cong., 1st Session 1089-90 (1866) (emphasis added).

41 Cong. Globe, 39th Cong., 1st Sess. at 1066 (1866).

42 6 F. Cas. 546 (No. 3,230) (C.C.E.D. Pa. 1823).

43 Cong. Globe, 39th Cong., 1st Sess. at 2766 (1866). Other Senators took similar positions. *See, e.g.*, Bernard Siegan, *The Supreme Court's Constitution* 55-65 (1987).

44 *See, e.g.*, Michael Kent Curtis, *supra* note 6, at 138-39.

45 *See* Hardy, *supra* note 24, at 15-23; *see also id* at 21 (quoting "Madison," *The National Question: The Constitutional Amendments—National Citizenship*, The New York Times, Nov. 10, 1866 at 2, col. 2-3).

46 Richard Aynes, *On Misreading John Bingham and the Fourteenth Amendment*, 103 Yale L.J. 57, 83-94 (1993).

47 *See, e.g.*, Hardy, *supra* note 36.

48 83 U.S. (16 Wall.) 36 (1873).

49 *Id.* at 59-60.

50 83 U.S. at 60.

51 *Id.* at 96 (Field, J., dissenting).

52 *Id.* at 77-78.

53 *Id.*at 70.

54 *Id.* at 79.

55 *E.g.*, Bryan H. Wildenthal, The *Lost Compromise: Reassessing the Early Understanding in Court and Congress on Incorporation of the Bill of Rights in the Fourteenth Amendment,* 61 Ohio St. L.J. 1051 (2000).

56 83 U.S. at 74.

57 In fact, the Privileges or Immunities Clause has been invoked for precisely that purpose by Institute for Justice client Erroll Tyler, a Boston entrepreneur seeking to launch a nautical tour company in Cambridge, Massachusetts. *See* http://www.ij.org/index.php?option=com_content&task=view&id=676&Itemid=165.

58 83 U.S.at 96 (Field, J., dissenting).

59 *Id.* at 87-88.

60 *Id.* at 98.

61 *Id.*at 110.

62 *See, e.g.,* United States v. Cruikshank, 92 U.S. 542, 549 (1875).

63 Richard L. Aynes, *Constricting the Law of Freedom: Justice Miller, the Fourteenth Amendment, and the Slaughter-House Cases*, 70 Chi.-Kent L. Rev. 627, 627 (1994); *see also* Laurence H. Tribe, *American Constitutional Law* 1320-21 (3d ed. 2000) ("The textual and historical case for treating the Privileges or Immunities Clause as the primary source of federal protection and state rights-infringement is very powerful indeed."); Akhil Reed Amar, *The Bill of Rights* 213 (1998) (explaining "[t]he obvious inadequacy—on virtually any reading of the Fourteenth Amendment—of Miller's opinion" in *Slaughter-House*).

64 Eric Foner, *Reconstruction: America's Unfinished Revolution, 1863-1877*, at 503 (1988).

65 Thomas B. McAffee, *Constitutional Interpretation—the Uses and Limitations of Original Intent*, 12 U. Dayton L. Rev. 275, 282 (1986).

66 *E.g.*, Jeffrey Rosen, *Textualism and the Civil War Amendment*, 66 Geo. Wash. L. Rev. 1241, 1268 (1998) ("[W]e can make a conscientious effort to resurrect the Privileges or Immunities Clause in its original context, but only if we are willing to look into the abyss and to acknowledge the fact that the practical consequences of a privileges or immunities revival would be, for nearly all of us, unacceptable.").

67 John Hart Ely, *Democracy and Distrust* 18 (1978). Of course, the fact that substantive due process has been subjected to criticism does not make that criticism correct or the doctrine wholly illegitimate. *See, e.g.*, James W. Ely, Jr., *The Oxymoron Reconsidered: Myth and Reality in the Origins of Substantive Due Process*, 16 Const. Comment. 315 (1999).

68 *See, e.g.*, Curtis, *supra* note 6, at 64-65.

69 C*f.* Saenz v. Roe, 526 U.S. 489, 527-28 (1999) (Thomas, J., dissenting); Brennan v. Stewart, 834 F.2d 1248, 1256 (5th Cir. 1988) ("[I]t would be more conceptually elegant to think of [protected] substantive rights as 'privileges or immunities of citizens of the United States' ....").

70 *Cf.* Rhode Island v. Massachusetts, 37 U.S. (12 Pet.) 657, 723 (1838) ("In the construction of the constitution, we must... examine the state of things existing when it was framed and adopted... to ascertain the old law, the mischief and the remedy….") (internal citation omitted).

71 *See* Morris M. Kleiner, *Licensing Occupations* 1 (Upjohn Institute 2006); *see also* Morris M. Kleiner & Alan B. Krueger, *Analyzing the Extent and Influence of Occupational Licensing on the Labor Market*, National Bureau of Economic Research Working Paper 14979, at 10-13 (2009), *available at* http://www.nber.org/papers/w14979 (providing most recent data).

72 *See* E. Frank Stephenson and Erin Wendt, *Occupational Licensing: Scant Treatment in Labor Texts*, 6 Econ. Journal Watch 181, 185-86 (2009) (summarizing research documenting how occupational licensing is used to erect barriers to entry or mobility); Walter Gelhorn, *The Abuse of Occupational Licensing*, 44 Chi. L. Rev. 6, 25 (1976) (arguing that "[o]nly the credulous can conclude that licensure is in the main intended to protect the public rather than those who have been licensed or, perhaps in some instances, those who do the licensing"); *see also* Dick M. Carpenter, *Designing Cartels* (documenting the efforts of a professional lobbying group to persuade state legislators to adopt interior-design licensing laws that have no plausible health or safety justification).

73 *E.g.,* Gitlow v. New York, 268 U.S. 652, 666 (1925).

74 *See, e.g.,* Loving v. Virginia, 388 U.S. 1 (1967) (striking down laws against racial intermarriage); Pierce v. Society of Sisters, 268 U.S. 510 (1925) (striking down an Oregon law against sending children to private schools); Meyer v.

Nebraska, 262 U.S. 390 (1923) (striking down Nebraska law forbidding teaching of German language in schools).

75 *But see* Saenz v. Roe, 526 U.S. 489, 527-28 (1999) (Thomas, J., concurring) (suggesting a willingness to reconsider the Privileges or Immunities Clause in an appropriate case).

76 District of Columbia v. Heller, 128 S. Ct. 2783, 2799 (2008).

77 *See generally* Robert J. Spitzer, *Lost and Found: Researching the Second Amendment*, 76 Chi.-Kent L. Rev. 349 (2000).

78 Provisions of the federal Bill of Rights apply directly to the District of Columbia, which is a creature of the federal government. *E.g.* Pernell v. Southall Realty, 416 U.S. 363, 370 (1974).

79 McDonald v City of Chicago, No. 08-1521, 2009 WL 1631802 (Sept. 30 2009).

80 *See* http://origin.www.supremecourtus.gov/qp/08-01521qp.pdf.

Corporations, Securities & Antitrust

Looks Can Be Deceiving: Holdout Litigation Under the Foreign Sovereign Immunities Act

*By Saloni Kantaria**

Historically, sovereign lending has been dominated by a small group of large banks and financial institutions.[1] The group of investors holding sovereign debt has become more diverse and includes commercial banks of all sizes, investment banks, pension funds, mutual funds, hedge funds, nonfinance companies, and retail investors.[2] In the late 1980's and early 1990's, a secondary market developed for distressed sovereign debt because banks sought to remove rescheduled sovereign debt from their books and did so by selling this debt at significantly discounted prices to the secondary market.[3] However, the secondary market for sovereign debt began to attract investors having no intention of making equity investments in the debtor countries. These investors, known as "vulture creditors,"[4] specialize in strategic purchase of debt on the secondary market and typically purchase sovereign debt that is trading at a deep discount as a result of the sovereign's financial distress.[5] The objective of the vulture creditors is to seek short-term gains, either through the restructuring process or by holding out of the restricting process until the debtors and majority creditors negotiate an offer of additional payment. In such a situation, the vulture creditor typically "free rides" by holding out for better terms already agreed to by other creditors in a restructuring process. If this is unsuccessful, the vulture creditor will seek to collect the full face value of its claim from the sovereign by means of litigation.[6] The term "holdout litigation" typically characterizes this situation, where a majority of creditors accept debt restructuring but a minority chooses to sue for full repayment.[7]

In order for a vulture creditor to sue a sovereign state in the United States, it must do so pursuant to the provisions of the Foreign Sovereign Immunities Act of 1976 (the "FSIA").[8] The FSIA was introduced for the purposes of curbing states from invoking sovereign immunity to prevent suits against them on commercial grounds.[9] A second purpose of the FSIA was to protect injured parties in commercial dealings with states. Although the FSIA gives the initial impression that it is suitable repayment vehicle for creditors, this paper will demonstrate that the FSIA standards for parties to bring suit and enforce a judgment against a sovereign state are rigorous and burdensome due to the dualism of immunity in the FSIA.[10] For this reason, many vulture creditors have not been successful against foreign states in U.S. courts.[11] In the context of holdout litigation by vulture creditors, this paper will explore the following issues: the evolution of the FSIA, the challenges of overcoming the jurisdictional immunity hurdle, and the practical challenges of seeking an order of attachment and execution against a sovereign state.

** LL.M. University of Chicago, LL.B. (Honors) University of Sydney, BSc. Cornell University, currently a practicing lawyer at Clayton Utz (a top tier firm in Australia).*

I. The Evolution of the FSIA

The principle of sovereign immunity derived from English law, which assumed that "the King can do no wrong."[12] In *Schooner Exchange v. McFaddon*, Justice Marshall, writing for the Supreme Court, for the first time held that sovereign states have absolute immunity.[13] In the 1926 case *Berizzi Bros. Co. v. S.S. Pesaro*,[14] the Court broadened the scope of sovereign immunity by extending the immunity traditionally accorded to military vessels to foreign commercial ships because "all ships held and used by a government... for the purpose of advancing the trade of its people or providing reserve for its treasury... are public ships in the same sense that warships are."

Following World War II, sovereigns increasingly became engaged in commercial activity and utilized the sovereign immunity defense for tort and breach of contract actions.[15] This paved way for concern that granting absolute immunity to sovereigns gave foreign nations a commercial advantage in the market over private firms not so privileged, especially because governments began to rely increasingly upon sovereign immunity to avoid commercial commitments.[16] In 1952, the U.S. State Department announced in the Tate letter that it was shifting away from the absolute sovereign immunity doctrine and adopting a restrictive theory of sovereign immunity. The letter stated:

> A study of the law of sovereign immunity reveals the existence of two conflicting concepts of sovereign immunity… According to the classical or absolute theory of sovereign immunity, a sovereign cannot, without his consent, be made a respondent in the courts of another sovereign. According to the newer or restrictive theory of sovereign immunity, the immunity of the sovereign is recognized with regard to the sovereign or public acts (juri imperii) of a state, but not with respect to private acts (jure gestionis)… [T]he Department feels that the widespread and increasing practice on the part of governments of engaging in commercial activities makes necessary a practice which will enable persons doing business with them to have their rights determined in the courts. For these reasons it will hereafter be the Department's policy to follow the restrictive theory of sovereign immunity in the consideration of requests of foreign governments for a grant of sovereign immunity.[17]

However, the Tate letter did not improve potential plaintiffs' ability to sue sovereign states because it failed to define "commercial activity," thus making it unclear under what circumstances these plaintiffs could sue. Secondly, the Tate letter was only applicable for the purposes of obtaining jurisdiction to sue a sovereign state and did not provide any guidance on executing judgments.[18] This created "considerable

uncertainty" and a "troublesome inconsistency" in immunity decisions.[19] Accordingly, to remedy the situation, Congress codified the requirements to sue a foreign state by passing the FSIA in 1976.[20] The issue of sovereign immunity thus became a question of statutory subject matter jurisdiction, the determination of which was removed from the executive branch and solely vested in the judiciary.[21] The FSIA imposes a dual form of immunity for sovereigns which presents two significant hurdles for vulture funds: (i) jurisdictional immunity,[22] and (ii) immunity from having their property attached in satisfaction of judgment.[23]

II. The Challenges of Overcoming the Jurisdictional Immunity Hurdle

Section 1604 starts from the premise that a foreign sovereign[24] is presumed to be immune from a suit in a U.S. court. Accordingly, the vulture fund has the burden of establishing that the U.S. court has jurisdiction over the foreign state. To rebut this presumption, the plaintiff must demonstrate that one of the exceptions to immunity under sections 1605-1607 of the FSIA is applicable.[25] In the context of vulture fund litigation, the most commonly utilized exceptions are (i) the commercial activity exception,[26] and (ii) the waiver of immunity exception.[27]

A. Commercial Activity Exception

The commercial activity exception in section 1605(a)(2) is the most litigated exception in the FSIA.[28] To demonstrate that a court has jurisdiction over a foreign state, a vulture fund must establish two elements: (i) the foreign state's act is a "commercial activity" within the definition of the FSIA, and (ii) there is some connection between the commercial activity and the United States.

Whether a foreign state's action is a commercial activity is determined by reference to the nature of the course of conduct or particular transaction or act, rather than by reference to its purpose.[29] In the 1992 case *Republic of Argentina v. Weltover, Inc.*,[30] the issue before the Supreme Court was whether the issuance of debt by the Republic of Argentina constituted a "commercial activity" under the FSIA. In its discussion of the definition of "commercial activity" under section 1603(d), the Court stated:

> When a foreign government acts, not as regulator of a market, but in the manner of a private player within it, the foreign sovereign's actions are "commercial" within the meaning of the FSIA. Moreover, because the Act provides that the commercial character of an act is to be determined by reference to its "nature" rather than its "purpose," the question is not whether the foreign government is acting with a profit motive or instead with the aim of fulfilling uniquely sovereign objectives. Rather, the issue is whether the particular actions that the foreign state performs (whatever the motive behind them) are the type of actions by which a private party engages in "trade and traffic or commerce." Because the FSIA has now clearly established that the "nature" governs, we perceive no basis for concluding that the issuance of debt should be treated as categorically different from other activities of foreign states.[31]

In relation to the second element, the FSIA provides three situations in which there is a connection between the commercial activity and the United States: (i) the activity was carried on in the United States by the foreign state; (ii) an act performed in the United States was connected to a commercial activity of the foreign state elsewhere; or (iii) an act performed outside the territory of the United States caused a direct effect in the United States and was connected to a commercial activity of the foreign state.[32]

In *Croesus EMTR Master Fund L.P. v. Brazil*,[33] three hedge funds sued the Federative Republic of Brazil ("Brazil") for failure to pay the principal and interest on Brazilian bonds they held. Brazil filed a motion to dismiss the complaint on the basis that the court lacked subject matter jurisdiction under the FSIA. The court first considered whether a commercial activity was carried on in the United States by Brazil. The hedge funds argued that this exception to immunity applied because Brazil engaged in the commercial activity of issuing securities and promoting secondary markets for its securities to and by persons in the United States.[34] The court found that there was no exception to immunity under the first clause of §1605(a)(2) because even if Brazil knew there was a secondary market for the bonds in the United States, and even if Brazil fostered that market, its failure to pay the principal and interest on the bonds was not "based upon" its purported promotion of the secondary market.[35]

The court also considered whether there was an act performed outside the territory of the United States which was connected to a commercial activity of Brazil elsewhere and which caused a direct effect in the United States.[36] The hedge funds identified two acts by Brazil that fit within this section: (i) Brazil's alleged failure to "keep market participants informed of actions affecting the value of the [b]onds"; and (ii) Brazil's failure to repay the principal and interest on the bonds.[37] The court promptly dismissed the first alleged act by Brazil because it considered that Brazil's failure to pay the principal and interest on the bonds was not "based upon" an alleged failure to inform market participants.[38] The hedge funds submitted that Brazil's failure to repay the principal and interest on the bonds had a "direct effect" in the United States because they had intended to direct the payments to their U.S. bank accounts.

The court acknowledged the Supreme Court's decision in *Weltover*, where the Court had held that a foreign state's unilateral rescheduling of bond maturity dates had a "direct effect" in the United States because the plaintiffs had indicated that their New York accounts were the place of payment and the foreign state had already made some interest payments into these accounts.[39]

However, the court found the facts distinguishable from *Weltover* since the hedge funds failed to designate any U.S. location as the "place of performance" where money was "supposed"[40] to have been paid.[41] Ultimately, the court held that Brazil was entitled to jurisdictional immunity because the hedge funds had not established an exception under §1605(a)(2) of the FSIA.

In *Global Index, Inc. v. The Honourable H.E. Benjamin W.Mkapa*,[42] the court was concerned with whether the failure by a sovereign to honor promissory notes constituted an exception to jurisdictional immunity pursuant to §1605(a)(2) of the FSIA. The court had little difficulty determining that the issuance of promissory notes by the sovereign constituted a commercial activity under §1605(a)(2) of the Act and primarily focused on whether the sovereign's failure to honor the promissory notes had a direct effect in the United States, thereby satisfying §1605(a)(2) of the FSIA. The court held that although the express text of the promissory note required payment in U.S. currency to a U.S. company in the United States, this was not sufficient to establish that the sovereign's commercial activity had a direct effect in the United States. In arriving at this finding, the court also focused on the fact that the plaintiffs had not designated any place of payment, let alone a bank or city in the United States.[43]

The decisions in *Croesus*, *Weltover*, and *Global Index* suggest that for vulture funds to prove that an exception to immunity exists under §1605(a)(2), they must first ensure that a clause exists within the contract that designates a location in the United States as the place of payment of principal and interest. If the vulture fund then proceeds to sue the foreign state on the basis that it has failed to repay the principal and interest on the bonds, it should have little difficulty convincing a court that the issuance of sovereign bonds constitutes a "commercial activity" which caused a direct effect in the United States pursuant to §1605(a)(2) of the FSIA.

B. Waiver of Immunity Exception

The second frequently-used exception to the general grant of immunity is §1605(a)(1) of the FSIA provides that a foreign sovereign is not immune if it has implicitly or expressly waived immunity from suit.[44] Explicit waivers, which are construed narrowly, could only be obtained if the foreign sovereign clearly and unambiguously granted a waiver in its contract with the vulture fund.

The FSIA does not define an implied waiver, but the courts narrowly construe this provision as well.[45] Determining whether a foreign sovereign has impliedly waived its immunity will depend on the facts of each case. For example, an implied waiver of immunity will be found if the foreign state demonstrates a conscious decision to take part in litigation in the United States by failing to raise the sovereign immunity defense despite an opportunity to do so.[46] However, in general, the courts are likely to find an implied waiver of sovereign immunity under the FSIA in only three circumstances: (i) a foreign state has agreed to arbitration in another country;[47] (ii) a foreign state has agreed that a contract is governed by the laws of the United States;[48] or (iii) a foreign state has filed a responsive pleading in a case without raising the defense of sovereign immunity.[49]

III. The Practical Challenges of Enforcing a Judgment Against a Foreign State

Merely obtaining jurisdiction over a foreign state is futile if the sovereign's property is immune from execution. A court must be capable of enforcing its judgment by allowing the vulture creditor to execute on the assets of the sovereign debtor. Although the FSIA applies the restrictive principle to immunity from execution on a sovereign's property, the FSIA requires a plaintiff to clear several hurdles before it allows a court to grant an execution order.[50]

Section 1609 of the FSIA starts with a general presumption that a foreign state's property located in the United States is immune from attachment unless it qualifies under an exception.[51] If a sovereign asserts this immunity, then the vulture fund must prove that one of the exceptions in §1610 of the FSIA is applicable. Section 1610 is applicable only to property that is located in the United States and creates no exception for property located outside of the United States,[52] since such property is absolutely immune from execution.[53] §1611 of the FSIA further extends unwaivable immunity to any property used in connection with a military activity and of either a military character or under the control of a defense agency. The FSIA also provides immunity to immovable property used for a diplomatic or consular mission.

Assuming that the sovereign's property can be located, it is a relatively simple task to demonstrate that the foreign state's property is located within the United States when a party seeks to attach tangible property.[54] However, complexities arise when dealing with intangible property. In *Af-Cap Inc. v. Republic of Congo* ("*Af-Cap II*"),[55] following the entry of a default judgment against the Republic of Congo in London, the Connecticut Bank of Commerce filed suit in New York state court to turn the foreign judgment into a U.S. judgment. The New York court entered judgment in the bank's favor and entered an order of attachment authorizing the bank to execute judgment against the Congo's property. On appeal to the Fifth Circuit, the issue before the court was whether the property the bank sought to garnish was in the United States. The court held that the relevant property was the garnishee's obligation to pay taxes and royalties to the Congo. It found that since the garnishees were business entities formed and headquartered in the United States, the property was in the United States for the purpose of the FSIA.[56]

Even if the vulture fund identifies property situated in the United States, §1610(a)(2) of the FSIA further provides that execution may be sought only against property in the United States which "is or was used for the commercial activity upon which the claim is based."[57] When dealing with sovereign states, one of the difficulties is distinguishing commercial from "official" property. This difficulty is apparent from the Fifth Circuit's previous decision in *Connecticut Bank of Commerce v. Republic of Congo*,[58] where it grappled with the words "used for" in §1610(a)(2):

> What matters under the statute is what the property is "used for," not how it was generated or produced. If property in the United States is used for a commercial purpose here, that property is subject to attachment and execution even if it was purchased with tax revenues or some other noncommercial source of government income. Conversely, even if a foreign state's property has been generated by commercial activity in the United States, that property is not thereby subject to execution or attachment if it is not "used for" a commercial activity within our borders.[59]

In *Birch Shipping Co. v United Republic of Tanzania*,[60] where the Tanzanian embassy in Washington, D.C. had used funds in an account for both commercial and non-commercial purposes, the D.C. district court took a practical approach in interpreting §1605(a)(2), holding that the funds could be attached to satisfy judgment against Tanzania on the basis that any other finding would mean that all sovereign accounts with funds used for multiple purposes would be beyond the reach of creditors. In *Eastern Timber Corp. v. Government of Republic of Liberia*,[61] the same court took a narrower approach, concluding that "funds used for commercial activities which are 'incidental' or 'auxiliary,' not denoting the essential character of the use of the funds in question, would not cause the entire bank account to lose its mantle of sovereign immunity."[62]

In *Af-Cap, Inc v. Chevron Overseas (Congo)*,[63] the judgment creditor sought to appeal the district court's judgment dissolving and vacating garnishments and liens filed against any property of the Republic of Congo, held by third party ChevronTexaco Corporation. The property that the creditor sought to garnish included intangible obligations of ChevronTexaco owed to Congo for various bonuses, taxes, and royalties related to the extraction of hydrocarbons, oil, and other of Congo's natural resources. The issue before the Ninth Circuit was whether Congo's property was used for a commercial activity in the United States. The Ninth Circuit adopted the Fifth Circuit's principle from *Connecticut Bank of Commerce v. Republic of Congo*[64] and held that under §1610(a) "property is 'used for a commercial activity in the United States' when the property is put into action, put into service, availed or employed for a commercial activity, not in connection with a commercial activity, or in relation to a commercial activity." Ultimately the court concluded that Af-Cap could not garnish the obligation to pay bonuses or the bonus payments up to the prepayment amount because they did not belong to the Congo. The court also found that since the obligation or bonus payment merely had a nexus or connection with a commercial activity in the United States, this was not sufficient to satisfy §1605(a) of the FSIA. For these two reasons, it held that the obligations of ChevronTexaco owed to the Congo were immune from execution or collection.

These decisions suggest that if a vulture creditor seeks an order of attachment against a foreign state's intangible property, such as a bank account, there is an onerous evidentiary burden to prove that foreign state's funds were being used for a commercial activity in the United States. In practice, the decisive question is whether the foreign state's moneys have been specifically designated for a particular purpose. Even if a creditor is able to overcome the evidentiary hurdle of proving that a foreign state's property is located in the U.S. and is used for a commercial activity, that creditor must still establish entitlement to a §1610 exception before a court will grant an order of attachment execution.[65] The most common exceptions in the context of vulture fund litigation are the waiver of immunity exception and use of the property for the commercial activity upon which the claim is based.

A. Waiver of Immunity Exception

The first exception is the foreign state's waiver of its immunity from attachment "either explicitly or by implication."[66] The Ninth Circuit has held that this provision must be construed narrowly.[67] With respect to contractual waiver of immunity, the court has held that the language must be explicit, although it is not necessary to recite the words "prejudgment attachment" to effect a waiver.[68]

An example of language within a contract that satisfies the "explicit" waiver test is in the case of *Karaha Bodas Co. v. Perusahaan Pertambangan Minyak Dan Gas Bumi Negara* ("*Pertamina*").[69] The court held that pursuant to the FSIA, a company waived its sovereign immunity from attachment in United States courts when, in its geothermal energy contracts with the judgment creditor, it waived any "right of immunity (sovereign or otherwise) which it or its assets now has or may have in the future... and consent[ed] in respect of the enforcement of any judgment against it."[70]

A commonly-used clause in sovereign bond and loan agreements is "the borrower waives its immunity from attachment prior to entry of judgment and from attachment in aid of execution against any of its property and assets irrespective of their use or intended use."[71] However, such a clause does not entitle a vulture creditor to attach any of the sovereign's assets because the FSIA permits courts to execute only against a foreign state's property that is located in the United States and is used for a commercial activity in the United States.[72] The reason for this strict approach is that "confiscating funds that are being put immediately to some sovereign use interrupts a sovereign's public acts."

B. Property Used for Commercial Activity Upon Which the Claim is Based

The second exception states that property in the United States of a foreign state is not immune from attachment if "the property is or was used for the commercial activity upon which the claim is based."[73] This provision serves to ensure that an antecedent basis for the adjudicative jurisdiction exists, and limits the property at issue to satisfy the judgment to resources that had already been allocated to a commercial transaction.[74] In the context of sovereign debt litigation, this provision is understood to mean that the judgment creditor may only execute on the sovereign state's property located in the United States used for the issuance of bonds. This provision places an onerous evidentiary burden on the creditor to prove the sovereign's intended use of the funds. These funds are unlikely to exist, or, if they do exist, are unlikely to be separated from other funds for the sole purpose of servicing a loan.

IV. The Final Steps

Even if the creditor is able to cross the jurisdictional immunity and immunity-from-enforcement hurdles, the court must not make an order of attachment or execution until it has determined that "a reasonable period of time has elapsed following the entry of judgment and the giving of any notice required under §1608(c) of this chapter."[75] This provision poses particular difficulty for creditors since the foreign state's property

is required to be located in the United States at the time the court authorizes execution.[76] This loophole invites foreign sovereigns to move their property outside the United States after a suit is brought and before execution is rendered.[77]

Fortunately, §1610(d) of the FSIA provides some recourse to vulture creditors who are in danger of such a situation:

> The property of a foreign state, as defined in section 1603(a) of this chapter, used for a commercial activity in the United States, shall not be immune from attachment prior to the entry of Judgment in any action brought in a court of the United States or of a State, or prior to the elapse of the period of time provided in subsection (c) of this section, if:
>
> (1) The foreign state has explicitly waived its immunity, from attachment prior to Judgment, notwithstanding any withdrawal of the waiver the foreign state may purport to effect except in accordance with the terms of the waivers; and
>
> (2) The purpose of the attachment is to secure satisfaction of a Judgment that has been or may ultimately be entered against the foreign state, and not to obtain Jurisdiction.

Despite the usefulness of this provision, a vulture creditor cannot utilize its benefit unless it satisfies both conditions. The first condition suggests, due to the "explicit waiver"[78] language, implicit waiver of immunity is not sufficient for pre-judgment attachment. The second condition prevents the prejudgment attachment from being used to obtain jurisdiction of the foreign state. From a practical perspective, the vulture creditor would have to demonstrate that it has a well-founded fear that the foreign state's assets will be removed from the court's jurisdiction prior to an order of execution.

Conclusion

The original intention of Congress enacting the FSIA was to make foreign states fully responsible for their commercial activities.[79] Although the initial impression of the FSIA is that it is a suitable repayment vehicle for vulture creditors, this is not so, due to the dualism of immunity and nexus requirements within provisions 1605 and 1610. These provisions make the FSIA unduly restrictive and provide substantial protection to sovereigns.

While a vulture fund may have little difficulty establishing an exception to jurisdictional immunity, the threshold requirement in Section 1610(a) of the FSIA poses particular problems. At the outset, the evidentiary requirements of Section 1610(a) eliminate large classes of property that might be candidates for execution in satisfaction of a judgment against a foreign sovereign.[80] The Ninth Circuit recently emphasized that the statutory structure and construction reflects a pivotal purpose of the FSIA, which is to limit execution against property directly belonging to a foreign state.[81]

It is evident from the cases discussed above that holdout litigation is often not successful due to the restrictiveness of the FSIA. However, the Fifth Circuit's decision in *Connecticut Bank of Commerce v. Republic of Congo* suggests that the FSIA provides no obstacle to garnishment if the foreign state has waived its immunity in the underlying loan contract and if the obligation sought to be garnished satisfies two conditions: (i) the obligation is located in the United States; and (ii) the obligation is used for a commercial activity in the United States. The Ninth Circuit also recently agreed with this approach.[82] Should a vulture fund succeed in attaching royalties that oil companies agreed to pay in exchange for their oil rights, they will reap extraordinary profits at the expense of U.S. companies and U.S. foreign relations.[83] To prevent this from becoming a reality, an option for sovereign debtors is to include a clause in the loan agreement that imposes a cooling off period, during which time "bondholders would be prevented from initiating litigation." At present, it seems that although holdout litigation appears to be gaining momentum in the United States, the restrictive immunity principle engrained in the FSIA will continue to provide protection to sovereign debtors.

Endnotes

1 Christopher Wheeler & Amir Attaran, *Declawing the Vulture Funds: Rehabilitation of a Comity Defence in Sovereign Debt Litigation*, 39 Stanf. J. of Int'l L. 253, 253 (2003).

2 Jill E. Fisch & Caroline M. Gentile, *Vulture or Vanguards?: The Role of Litigation in Sovereign Debt Restructuring*, 53 Emory L. J. 1043, 1070 (2004).

3 Ronald J. Silverman & Mark W. Deveno, *Distressed Sovereign Debt: A Creditor's Perspective*, 11 Am. Bankr. Inst. L. Rev. 179, 185 (2003).

4 Diana B. Henriques, *The Vulture Game*, N.Y. Times 1 (July 19, 1992). The term "vulture" comes from the analogy that these investors "get rich by feeding on the carcasses of corporate America."

5 Deepak Gopinath, *The Debt-Crisis Crisis*, Institutional Investor, 36, 38 (August 2002).

6 Jill E. Fisch & Caroline M. Gentile, *Vulture or Vanguards?: The Role of Litigation in Sovereign Debt Restructuring*, 53 Emory L.J. 1043, 1072-1073 (2004).

7 Michael Waibel, *Opening Pandora's Box: Sovereign Bonds in International Arbitration*, 101(4) Am. J. Int'l L. 711, 713 (Oct. 2007).

8 Argentine Republic v. Amerada Hess, 488 U.S. 428, 434 (1989).

9 Michael L. Morkin, Ethan A. Berghoff & Richard S. Pike, *Doing Business with Foreign Sovereign Entities. Who Are They and What Are the Risks*, 17(2) Business Law Today, (Nov./Dec. 2007).

10 Molly Steele & Michael Heinlen, *Challenges to Enforcing Arbitral Awards Against Foreign States in the United States*, 42 Int'l Law. 87, 88 (2008).

11 Jonathan C. Lippert, *Vulture Funds: The Reason Why Congolese Debt May Force a Revision of the Foreign Sovereign Immunities Act*, 21 N.Y. Int'l L. Rev. 1, 1 (2008).

12 Erwin Chemerinsky, *Against Sovereign Immunity*, 53(5) Stan. L. R. 1201, 1201 (May 2001).

13 11 U.S. 116 (1812).

14 271 U.S. 562 (1926).

15 Jonathan C. Lippert, *Vulture Funds: The Reason Why the Congolese Debt May Force a Revision of the Foreign Sovereign Immunities Act*, 21 N.Y. Int'l L. Rev. 1 (2008).

16 Stella Havkin, *The Foreign Sovereign Immunities Act: The Relationship Between the Commercial Activity Exception and the Noncommercial Tort Exception in Light of* De Sanchez v. Banco Central Di Nicaragua, 10 Hastings Int'l & Comp. L. Rev. 455, 459 (1987).

17 *Letter from Jack B. Tate, Acting Legal Adviser, U.S. Dept. of State, to Acting U.S. Attorney General Philip B. Perlman*, 26 Dep't. St. Bull. 984 (May 1952).

18 Jonathan C. Lippert, *Vulture Funds: The Reason Why Congolese Debt May Force a Revision of the Foreign Sovereign Immunities Act*, 21 N.Y. Int'l L. Rev. 1, 11 (2008).

19 Republic of Austria v. Altmann, 541 U.S. 677, 716 (Kennedy, J., dissenting).

20 42 Fed. Reg. 6366 (February 2, 1977).

21 Gary W Larson, Default on Foreign Sovereign Debt: A Question for the Courts? 18 Ind. L. Rev. 959, 965 (1985).

22 28 U.S.C. § 1604.

23 28 U.S.C. § 1609.

24 For FSIA purposes, 28 U.S.C. § 1603(a) defines "foreign states" as not only states, but also political subdivisions, agencies, and instrumentalities. For the purposes of this paper, only a state will be considered.

25 Jonathan C. Lippert, *Vulture Funds: The Reason Why Congolese Debt May Force a Revision of the Foreign Sovereign Immunities Act*, 21 N.Y. Int'l L. Rev. 1, 12 (2008).

26 28 U.S.C. §1605(a)(2).

27 28 U.S.C. §1605(a)(1).

28 Jonathan C. Lippert, *Vulture Funds: The Reason Why Congolese Debt May Force a Revision of the Foreign Sovereign Immunities Act*, 21 N.Y. Int'l L. Rev. 1, 12 (2008).

29 28 U.S.C. § 1603(d).

30 504 U.S. 607 (1992).

31 *Id.* at 614-615.

32 28 U.S.C. §1605(a)(2).

33 212 F.Supp.2d 30, 34 (D.D.C.2002).

34 *Id.*

35 *Id.*

36 28 U.S.C. §1605(a)(2).

37 212 F.Supp.2d 30, 35 (D.D.C.2002).

38 *Id.*

39 504 U.S. 607, 619 (1992).

40 In *Goodman Holding v. Rafidian Bank*, 26 F.3d 1143 (D.C.Cir.1994), the D.C. Circuit clarified that a "direct effect" in the United States occurs only where payment was "supposed" to have been made or received in the United States.

41 212 F.Supp.2d 30, 36 (D.D.C.2002).

42 290 F.Supp.2d 108 (D.D.C. 2003).

43 290 F.Supp.2d 114 (2003).

44 28 U.S.C. §1605(a)(1).

45 Sampson v. Federal Republic of Germany, 250 F.3d 1145 (7th Cir. 2001).

46 Talbot v. Saipem A.G., 835 F. Supp. 352 (S.D. Tex. 1993).

47 In re Republic of Phillippines, 309 F.3d 1143 (9th Cir. 2002).

48 Calzadilla v. Banco Latina Internacional, 413 F.3d 1285 (11th Cir. 2005).

49 Haven v. Polska, 215 F.3d 727, 46 Fed. R. Serv. 3d 1180 (7th Cir. 2000).

50 28 U.S.C. §1610.

51 28 U.S.C. §1609.

52 Philippine Export & Foreign Loan Guarantee Corp. v. Chuidian, 218 Cal. App.3d 1058, 1093 (6th Dis. 1990).

53 Fidelity Partners, Inc. v. Philippine Export and Foreign Loan Guarantee Corp., 921 F.Supp. 1113, 1117 (S.D.N.Y.1996).

54 S & S Machinery Co. v. Masinexportimport, 802 F.Supp.1109, 1110 (S.D.N.Y.1992).

55 Af-Cap Inc. v. Republic of Congo, 383 F.3d 361, 364 (5th Cir. 2004).

56 *Id.* at 373.

57 28 U.S.C. §1610(a)(2).

58 309 F.3d 240, 251 (5th Cir. 2002).

59 *Id.*

60 507 F.Supp.311, 313 (D.D.C. 1980).

61 659 F.Supp. 606 (D.D.C. 1987).

62 Eastern Timber Corp. v. Government of Republic of Liberia, 659 F.Supp. 606, 610 (D.D.C.1987).

63 475 F.3D 1080 (9th Cir.).

64 309 F.3d 240 (5th Cir. 2002).

65 It is important to note that section 1611 of the FSIA recognizes the following types of property as absolutely immune from attachment: assets held in international organizations, property of military character, and property of a central bank or monetary authority.

66 28 U.S.C. §1610(a)(1).

67 Af-Cap, Inc v. Chevron Overseas (Congo), 475 F.3D 1080, 1087 (9th Cir. 2007).

68 Libra Bank, Ltd. v. Banco Nacional de Costa Rica, S.A., 676 F.2d 47, 49-50 (2nd Cir. 1982).

69 313 F.3d 70 (2d Cir. 2002).

70 *Id.* at 75.

71 Georges R. Delaume, *The Foreign Sovereign Immunities Act and Public Debt Litigation: Some Fifteen Years Later*, 88 Am. J. Int'l L. 257, 270 (1994).

72 Connecticut Bank of Commerce v. Congo 309 F.3d 240 (5th Cir. 2002); Af-Cap Inc. v. Republic of Congo 383 F.3d 361, 366 (5th Cir. 2004).

73 § 1610(a)(2).

74 Jin Sun, *Immunity From Execution of Judgments Under the FSIA – Moderate Balance to the New UN Convention*, 4(5) US-China L. Rev. 13,15 (2007).

75 28 U.S.C. §1610(c).

76 FG Hemisphere Assocs. v. Republique du Congo, 455 F.3d 575 (5th Cir. 2006).

77 Timothy B. Atkeson & Stephen D. Ramsey, *Proposed Amendment of the Foreign Sovereign Immunities Act: The Mathias Bill*, 79 Am. J. Int'l L. 770, 776 (1985).

78 Af-Cap, Inc. v. Chevron Overseas (Congo) Ltd., 475 F.3d 1080 (9th Cir. 2007).

79 Timothy B. Atkeson & Stephen D. Ramsey, *Proposed Amendment of the Foreign Sovereign Immunities Act: The Mathias Bill*, 79 Am. J. Int'l L. 770, 776 (1985).

80 Af-Cap, Inc v. Chevron Overseas (Congo), 475 F.3d 1080, 1089 (9th Cir. 2007).

81 *Id.* at 1088.

82 Af-Cap, Inc v. Chevron Overseas (Congo), 475 F.3d 1080 (9th Cir. 2007).

83 Jonathan C. Lippert, *Vulture Funds: The Reason Why Congolese Debt May Force a Revision of the Foreign Sovereign Immunities Act*, 21 N. Y. Int'l L. Rev.1, 2 (2008).

Judicial Review of Mutual Fund Advisory Fees: Reliance on Markets or Statutory Language?

*By Joanne T. Medero**

Do you own mutual fund shares either directly or through your 401(k) plan? If so, you should be interested in the outcome of the case *Jones v Harris Associates* to be heard early in the 2009-2010 Supreme Court Term. [1] The case involves allegations of "excessive fees" paid to a mutual fund's investment adviser and is notable because the Court rarely takes cases arising under the Investment Company Act ("ICA").[2] The Court was likely motivated to grant review by the differing views of Seventh Circuit Chief Judges Frank Easterbrook and Richard Posner.

I. Case History and Background

The plaintiffs were investors in certain mutual funds advised by Harris Associates, L.P. They alleged that the adviser's compensation was excessive and that, as a result, the adviser had violated ICA Section 36(b).[3] On a motion for summary judgment, the district court dismissed the case, relying on the standard established in a 1982 Second Circuit decision in *Gartenberg v. Merrill Lynch Asset Management*, where the court of appeals found that in order for an adviser to violate Section 36(b), the fee must be so disproportionately large so as to bear no reasonable relationship to the services provided and could not have been the product of arm's length bargaining between the adviser and the mutual fund. The *Gartenberg* court had given considerable weight to whether a fund's board carefully considered the fee and had applied various factors in determining whether a fee is disproportionately large.[4]

On appeal, the Seventh Circuit, in an opinion authored by Chief Judge Easterbrook, affirmed the order of summary judgment but "disapproved" the *Gartenberg* standard, holding that under Section 36(b) a court need only determine whether the fee was negotiated by the investment adviser in a manner consistent with its fiduciary duty to the fund. Stating that it was "skeptical about *Gartenberg* because it relies too little on markets," the opinion went on to say that a "[f]iduciary must make full disclosure and play no tricks but is not subject to a cap on compensation." [5] The panel reasoned that fees are subject to competitive pressure because investors can easily exit a fund when costs are too high relative to return, and what is "excessive" depends on the results available from other investment vehicles, rather than any absolute level of compensation.

Plaintiffs sought rehearing en banc, which was denied as the Seventh Circuit's active judges split five-to-five, with one judge not participating. Joined by four other judges, Judge Posner authored a dissent, arguing that due to the nature of the "captive" relationship of a fund and its directors to an adviser, a court is required in a Section 36(b) case to do more than determine whether the fee negotiations had been open and honest. The dissent criticized the panel for its conclusion that a breach of fiduciary duty could only be inferred from a fee that was "unusual," which the panel would have "applied solely by comparing the adviser's fee with the fees charged by other mutual fund advisers," stating that the comparability approach would allow fees that have resulted from less than arm's length bargaining "to become the industry's floor."[6]

II. ICA Section 36(b)

At its core, the ICA is a regime designed to protect investors in two ways: by providing disclosures about the investment, including historical performance and fees, and by putting in place structural safeguards against the actual and potential conflicts of interest inherent in the structure of mutual funds. Although mutual funds are technically owned by the shareholders who invest in the funds, most mutual funds are created, organized, and managed by external investment advisers. Advisers are compensated for their administrative services and investment management through agreements that must be approved annually by the board of directors of the fund company. Advisory fees are usually calculated as a percentage of the funds' net assets and fluctuate with the value of the funds' portfolio. The ICA provides that a majority of directors must be independent of the adviser and its affiliates, and that advisory contracts must be approved by a majority of these disinterested directors.

It is rare that fund companies fire their advisers. This is not only because of the intertwining of management and services to the fund, but also because a primary reason to invest in a fund is the performance history and reputation of the adviser. To replace the advisory firm with another is tantamount to changing one of the principal factors considered by the fund's shareholders in deciding whether to invest, something directors would naturally be reluctant to do absent extraordinary circumstances.

Congress added Section 36(b) to the ICA in 1970 to impose on advisers a "fiduciary duty with respect to the receipt of compensation for services" and to create a right of action for breach of that fiduciary duty to be exercised by either the SEC or by fund shareholders. In an action under Section 36(b), the approval by the directors of compensation paid to the adviser "shall be given such consideration by the court as is deemed appropriate under all the circumstances." The burden of proof is on the plaintiff to show that there has been a violation of this provision of the ICA. Damages are limited to "actual damages resulting from the breach of fiduciary duty" and cannot exceed the amount of compensation received by the adviser.

While it may seem convoluted that it is the *receipt* of compensation by the adviser that creates the fiduciary duty, rather than the approval of the advisory contract by the directors, this is in fact logical, given the unique structure of mutual funds and the relationship of the adviser to the fund company. At the time of the enactment of Section 36(b), Congress also enacted ICA Section 36(a), which authorizes the SEC—but not by its terms shareholders—to bring actions for

** Joanne Medero is Head of Government Relations and Public Policy for Investment Banking and Investment Management for Barclays PLC. The views expressed are her own.*

breach of fiduciary duty involving personal misconduct against the officers or directors of the fund company or the investment advisory firm.

The legislative history of Section 36(b) shows that Congress was concerned that although as originally enacted the ICA provided a comprehensive construct for the elimination and mitigation of conflicts, it did not provide an effective mechanism by which the fairness of investment advisory contracts could be tested in court. Earlier versions of the legislation that ultimately became Section 36(b) contained language that advisory fees should be "reasonable," but the version enacted eliminated this concept, substituting breach of fiduciary duty as the test instead. It is apparent from a review of this history that the mutual fund industry was concerned that the SEC or the courts would engage in rate-setting, and that the industry clearly preferred the language of Section 36(b) as enacted.[7]

III. Economic Analysis

To date, much of the commentary on the *Harris Associates* case has focused on the differing economic analysis approach in the panel's decision as compared to that in the rehearing dissent, thus setting up Judge Easterbrook against Judge Posner.[8] The panel decision is based on a classical economic analysis that there are thousands of mutual funds available and investors will vote with their feet if the costs relative to performance are too high, so advisers are strongly incented to keep costs low to attract investors. In the dissent, the focus is much more on behavioral economic studies that show that mutual fund investors do not make decisions based on costs. Further, the dissenters also believe the governance structure of the industry is such that directors have "feeble incentives" to police an adviser's compensation. Judge Posner compares the setting of fund advisory fees to excessive executive compensation in publicly traded firms and notes further that mutual funds are a component of the financial services industry, where "abuses have been rampant." In sum, the panel decision concludes that market forces—the competition for more assets—functions well, while the dissent focuses on the distortions created by the mutual fund governance structure, and concludes that there is a market failure that may warrant intervention.

Both the panel decision and the dissent look to "comparability" as a source of information on whether the advisory fees are potentially in violation of Section 36(b). They differ, however, as to what the appropriate comparison should be. The panel would look to mutual funds of similar size with similar investment strategies, while the dissenters' view is that the courts should look to the potential disparity of fees charged by an advisor to its affiliated mutual fund as compared to its unaffiliated institutional clients.[9] This latter approach has been adopted by the Eighth Circuit in a case decided after *Harris Associates* that may further influence the decision of the Court.[10]

IV. Statutory Construction

Actions brought under ICA Section 36(b) are generally referred to as "excessive fee" cases (in an implicit reference to *Gartenberg*), but nowhere in the language of the Section is there any reference to the relative level of fees. Rather, the statute simply states that the investment adviser to a mutual fund shall be deemed to have a fiduciary duty in the receipt of compensation paid by the fund.

Under established principles of statutory construction, the term "fiduciary duty" is to be construed by its plain meaning. But there is no per se law of fiduciary obligations because the nature of fiduciary duty depends on the circumstances and the relationship of the parties involved. As a result, the obligations of a fiduciary may be different depending on whether the relationship is formed under the common law of agency, corporations, wills, or trusts. Regardless of the law under which the relationship formed, a fiduciary is not precluded from earning a fee or other compensation (although it is common to prohibit or limit the ability of a fiduciary to purchase assets from, or sell assets to, the person or entity for which it acts as a fiduciary). In construing the term "fiduciary duty," the panel decision in *Harris Associates* considers the term in relation to the law of trusts, and concludes based on its reading of the *Restatement (Second) of Trusts* that provided that the trustee has fulfilled his obligation of candor in negotiation and honesty in performance, a trustee may negotiate his fee in his own interest and accept what the settlor agrees to pay. The decision does allow, however, that fiduciary compensation could be "so unusual" that a court will infer that deceit must have occurred or the decision-makers have abdicated their responsibilities.[11]

While not addressed in the decision, another provision of the ICA enacted contemporaneously with Section 36(b) supports the panel's analysis of the scope of the adviser's fiduciary obligation. The 1970 amendments added Section 15(c), which requires the approval of the investment advisory contract by a majority the disinterested directors, and stipulates that it shall be "the duty of the directors… to request and evaluate, and the duty of the investment adviser… to furnish, such information as may reasonably be necessary to evaluate the terms of any contract." Reading Section 36(b) together with Section 15(c), it is reasonable to conclude that the fiduciary duty of the adviser is to make all relevant disclosures both proactively and in response to particular queries from the directors. The ICA has thus adopted a common approach to a potential "self-dealing" conflict: disclosure to a competent and disinterested decision-maker.

At the same time, the approval of the compensation by the directors is, in the language of Section 36(b), to be given "such consideration by the court as is deemed appropriate under all the circumstances." This preserves for plaintiffs the ability to show that the directors, in the words of the panel decision, "abdicated" their role.

This analysis needs to be reconciled with Section 36(a)—also added in the 1970 amendments—which addresses breaches of fiduciary duty involving personal misconduct, and Section 36(b), which specifically states that the plaintiff need not prove personal misconduct. The text here provides an explanation: Section 36(a) establishes a cause of action against persons acting in certain listed capacities, such as an officer or director of the fund or an adviser. Section 36(b) is directed to the adviser and affiliated persons of the adviser, which, given the structure of the mutual fund industry, are not individuals but rather corporate

and other legal entities. Section 36(b) does not seek recompense from individuals for receipt of compensation for a fiduciary breach but rather seeks it from the entity which directly received the compensation. It seems logical not to require that the plaintiff prove "personal misconduct" when the action lies against an entity rather than its officers and directors.[12]

V. Conclusion

Decisions under Section 36(b) have illustrated the reluctance of judges to substitute their judgment for that of mutual fund directors and a strong desire to avoid substantive rate setting. The *Gartenberg* decision does so by establishing the factors to be considered by the directors in approving the contract, and *Harris Associates* by its reliance on the market and its reading of fiduciary obligations established by trust law. While it can be argued that *Gartenberg* goes beyond the statutory language of Section 36(b), it can also be said that the panel in *Harris Associates* failed to give the ICA its full effect.

A solution that gives more weight to the statutory language can reconcile one apparent split between the Second and Seventh Circuit. *Gartenberg*, with its emphasis on factors, can be said to set out the type of information that a fiduciary should present as part of full disclosure of its fee arrangements, and what most directors would consider to be important in their evaluation of an advisory contract. Congress did not specify the elements of information that would discharge the fiduciary duty of the adviser, so it would not be appropriate for the Court to establish a mandatory list itself. But it could acknowledge that these factors are relevant to the analysis of the fair disclosure obligation found in *Harris Associates*.

The Court must resolve whether *Gartenberg* is correct that a "disproportionately large" fee violates Section 36(b) or whether the Seventh Circuit is correct in rejecting what it calls a "reasonableness" test. Here, Chief Judge Easterbrook is more true to the statute: Congress imposed a fiduciary duty on an adviser, not a cap on its compensation. Neither the absolute level of the compensation nor the comparative level, whether relative to other mutual fund advisory fees or to fees charged other clients of the adviser, is an appropriate inquiry under Section 36(b), except in the narrow circumstances where the "unusual" fee signals a potential fiduciary breach. In such circumstances, if the plaintiff is unable to show that the disclosures made by the adviser are materially inadequate, the action under Section 36(b) must fail.

Endnotes

1 Jones. v. Harris Associates, L.P., No. 08-586. Oral argument has been set for November 2, 2009.

2 (15 U.S.C. § 80a-1 *et seq.*).

3 527 F. 3d 627 (7th Cir. 2008).

4 694 F.2d 923 (2d Cir. 1982), cert. denied, 461 U.S.906 (1983). These non-exclusive factors include fees charged by other advisers to similar funds, the adviser's costs in providing the services, the nature and quality of the services, the economies of scale in providing services as the fund grows larger to the extent realized by the adviser, and the volume of orders or transactions the adviser must process. These factors have been incorporated into SEC regulations. *See* Disclosure Regarding Approval of Investment Advisory Contracts by Directors of Investment Companies, 69 F.R. 39798 (June 30, 2004).

5 *Harris Associates* at 632.

6 Jones v. Harris Associates, L.P., 537 F.3d 728, 732 (7th Cir. 2008).

7 *See* Investment Company Amendments Act of 1970, S. Rep. No. 91-1382, 91st Cong., 2d Sess (1970).

8 *See, e.g.,* Floyd Norris, "The Supremes Will Decide Which Economics Makes Legal Sense," *New York Times* (March 9, 2009). Of course, there has also been considerable commentary on whether the case means the end of *Gartenberg*, especially since mutual fund advisers rarely have been found to have violated Section 36(b) by courts applying this analysis.

9 *Compare* Harris Associates, 527 F. 3d at 634, *with* Harris Associates, 537 F.3d at 731-732 (Posner, J., dissenting). While both opinions proffer explanations as to why differing advisory fees may be warranted as between a mutual fund and an institutional client—largely based on services and the differing nature of the investor (such as a pension fund), neither opinion focuses on what may also be an important distinction: pension funds generally invest through separately managed accounts or bank maintained collective trust funds. Exempted from regulation under the ICA, separate accounts and collective trust funds avoid the legal, accounting, and other expenses of investment company registration and regulation. Thus, advisers using these accounts and funds can offer lower fees. Any comparability analysis needs to take into account not only the type of investor but also the structure of the investment.

10 Gallus v. Ameriprise Financial, Inc., 615 F.3d 816 (8th Cir. 2009).

11 *Harris Associates*, 537 F.3d at 632.

12 Further support that Section 36(a) is focused on individual conduct is that it does not authorize a private right of action and makes no mention of recoupment of the "actual amount of damages" caused by the fiduciary breach Rather, it provides the SEC the authority to seek an injunction and any other relief as may be appropriate under the circumstances.

Criminal Law and Procedure

Miranda with an English Accent

*by Lauren J. Altdoerffer**

You live in England. You have dabbled in some illegal activities and now find yourself on trial for a violent crime. Your attorney asks if you want to testify. You say yes. Now you are sitting in the witness stand, answering your attorney's questions.

Your attorney asks, "Mr. Smith, where were you on the night of the crime?"

You tell him, "I was in Dumfries, visiting my brother."

"You were nowhere near London on the night of the incident then?" your attorney asks.

"No sir," you reply.

"No further questions, your honor."

Opposing counsel stands up for cross-examination. "Mr. Smith," he says, "You say you were in Dumfries on the night of the crime."

"Yes sir," you reply.

"With your brother?" he asks.

"Yes, sir, with my brother."

The attorney continues, "Now tell me, Mr. Smith, did you mention that you were in Dumfries with your brother to the officer that interrogated you?"

"No, sir," you say, fidgeting slightly.

"And can you tell me, Mr. Smith, did the officer tell you not to tell him that during the interrogation?"

You timidly reply, "Well, not exactly."

"What exactly did the officer tell you, Mr. Smith?"

You pause, knowing where this is going. You've been caught lying. "The officer told me that I did not have to say anything, but that it would harm my defense if I didn't mention straight-away something which I might later rely on in court."

Opposing counsel closes in, "So, Mr. Smith, if you had an alibi when you were questioned, why didn't you mention this to the officer when he interrogated you?"

"I don't know, sir."

Opposing counsel walks behind his desk, and you hear, "No further questions your honor."

With six questions, opposing counsel has presented the jury with a new piece of evidence to consider—your pre-trial silence has created an inference that you are lying.

Miranda's "right to remain silent" makes some Americans uneasy about this line of questioning, but it is standard in the United Kingdom. In 1994, the United Kingdom adopted a rule that if the accused failed to mention a fact during interrogation, only to rely on the fact later during trial, a judge or jury "may draw such inferences from the failure as appear proper[.]"[1]

For the past 15 years, the Criminal Justice and Public Order Act (CJPOA) has allowed the jury to draw an adverse inference from pre-trial silence. The jury is advised of its ability upon instruction from the judge, who may direct the jury that it may, if the jurors think right, make adverse inference from a "no comment" interview. But the jury is not required to make any inference.[2] The instruction usually sounds like this:

> You may draw such a conclusion against [the accused] only if you think it is a fair and proper conclusion and you are satisfied about three things: first, that when he was interviewed he could reasonably have been expected to mention the fact on which he now relies; secondly, that the only sensible explanation for his failure to do so is that he had no answer at the time or none that would stand up to scrutiny; third, that apart from his failure to mention those facts, the prosecution's case against him is so strong that it clearly calls for an answer.[3]

The instructions limit use of the adverse inference. Guilt cannot rest solely on an inference from the defendant's failure to testify,[4] the inference cannot satisfy an element of the prosecution's case,[5] and the jury must consider whether the defendant reasonably relied on counsel's advice to remain silent throughout the interview.[6] All of the safeguards protect a defendant from a false conviction.

While defense protections are important, possible benefits of this type of interrogation are also important. Law enforcement can begin its investigation of the suspect at the source, instead of speculating about a suspect's role in the investigation. This is not possible in the United States, where a suspect knows his silence can stop an investigation in its tracks.

I. United Kingdom Interview and Warnings

The United Kingdom readily accepted an adverse inference from silence as part of its constantly evolving system, which seeks to strike a balance between the accused's right to silence and the government's interest in convicting the guilty. The United Kingdom recognizes a social obligation to aid the police in any type of investigation,[7] and to consistently follow police interview protocol.[8] The protocol, outlined in the Code of Practice for the Detention, Treatment and Questioning of Persons by Police Officers, requires police to administer a caution before questioning if officers intend to uncover evidence for trial.[9]

The caution is administered as an officer informs the suspect of the nature of the offense.[10] It must be given anytime the officer believes he has grounds to arrest a suspect for a criminal offense,[11] unless the accused has been previously cautioned, or his behavior makes an advisory impracticable.[12] The terms of the caution are the same regardless of the offense. Each time a person is arrested the officer cautions the arrestee, "You do not have to say anything. But it may harm your defence if you do not mention when questioned something which you later rely on in Court. Anything you do say may be given in

* *Lauren J. Altdoerffer is an Associate Attorney at the Criminal Justice Legal Foundation.* *Lauren would like to thank Kent Scheidegger, Legal Director, for his comments and advice.*

evidence." The caution preserves silence as evidence and assures that the purpose of the caution has been maintained.[13]

II. The United States' Constitutional Right to Silence

Miranda has been the law of our land since 1966.[14] It is fairly recent, yet we treat it as if it were part of the Bill of Rights' gospel. Practices and laws that do not comport with its maxims are declared unconstitutional, yet the Constitution does not promise a right to remain silent. Only the privilege against compelled testimony at trial is included in its text. The Warren Court conceived *Miranda*'s warnings, and a right to remain silent, in order to prevent compelled testimony in a police station. At the same time, the Court left room for the states to create their own warnings. The states should take the Court up on its offer, and the United Kingdom has provided a model to follow.

Early on, only a defendant at trial enjoyed the benefits of a Fifth Amendment right against self-incrimination. Chief Justice Marshall once observed, "If, in such a case, he say upon his oath that his answer would incriminate himself, the *court* can demand no other testimony of the fact."[15] *Miranda v. Arizona* expanded Fifth Amendment protection beyond the courtroom and into custodial interrogation.[16] Every custodial suspect now has a constitutionally based right to be warned during interrogation,[17] and thus, this right cannot be overturned by an act of Congress.[18] But what about the warnings? The words chosen by the *Miranda* Court are advisory, adopted by the Court "to give concrete constitutional guidelines for law enforcement agencies and courts to follow."[19] Alternative warnings are possible.[20]

To best understand how change is possible, it is necessary to go to the decision adopting *Miranda*'s warnings. After reviewing several examples of coercive custodial interrogations of suspects, the Court held that the best way to prevent compelled self-incrimination at trial was to prevent the prosecution's use of "statements, whether exculpatory or inculpatory, stemming from custodial interrogation of the defendant unless it demonstrates the use of procedural safeguards effective to secure the privilege against self-incrimination."[21] This left a few important questions unanswered, such as, what type of procedural safeguards would satisfy the Fifth Amendment? Does the privilege against self-incrimination also prevent revealing information that doesn't incriminate you? Could another warning prevent self-incrimination at trial and still allow questioning during interrogation? *Miranda* encourages states to experiment with their own procedures to incorporate the same protection;[22] therefore, the answer to the third question is "yes."

The standard *Miranda* warning was created by the Court, but can be altered slightly from jurisdiction to jurisdiction.[23] Modifications currently assure that the accused understands that he can remain silent. For example, in some jurisdictions each sentence may be followed by the question, "Do you understand?" This is just one more protection to insure that the accused voluntarily surrendered his silence. Each jurisdiction can reform the warning so long as warnings advise the accused "of [his] right to silence and to assure a continuous opportunity to exercise it[.]"[24] As the Court stated in *Miranda*, "Our decision in no way creates a constitutional straitjacket which will handicap efforts at reform, nor is it intended to have this effect."[25] The United States could reform its warnings to resemble the United Kingdom's warnings. These reformed warnings would still advise the accused of his right to silence while granting him the opportunity to exercise it.

Miranda's decision is really the means to an end. It insures that the accused has a "real understanding and intelligent exercise of the privilege" to remain silent.[26] Phrases "that anything said can and will be used against the individual," and the "right to have counsel present," are tools that prevent compulsory self-incrimination.[27] But before *Miranda*, a defendant had no reason to expect he was always entitled to silence. For a long time legal scholars proposed the judge be allowed to examine the accused.[28] *Griffin v. California*[29] silenced those arguments, but post-*Miranda*, new proposals have sought to correct the misconception that the Fifth Amendment created a blanket protection of a right to silence.[30]

One example of such a proposal was provided by Judge Friendly in an article that explains how *Miranda*'s warnings actually act as a police straitjacket during interrogation.[31] To Judge Friendly, extending the privilege to police investigation because silence would protect the "healthy and conservative goals" of a criminal trial was "like prohibiting graduate students from looking at secondary sources for fear this will tempt them from original research and thus corrupt their morals."[32] Judge Friendly found scarce logic to support *Miranda*'s protection pre-trial and proposed that the Fifth Amendment's protection from compulsion never prohibit

> [c]omment by the judge at any criminal trial on previous refusal by the defendant to answer inquiries relevant to the crime before a grand jury or similar investigating body, or before a judicial officer charged with the duty of presiding over his interrogation, provided that he shall have been afforded the assistance of counsel when being so questioned and shall have then been warned that he need not answer; that if he does answer, his answer may be used against him in court; and that if he does not answer, the judge may comment on his refusal.[33]

The proposal permitted judicial comment only when the defendant was afforded counsel, received warnings during questioning, and chose silence.[34]

The United Kingdom's version of the right to silence allows a judge to comment on the defendant's pre-trial silence, so long as the accused had the opportunity to have counsel present during his pre-trial interview, and the interrogating officers reasonably followed the standard outlined in the Code.[35] Arguably, if we adopted Judge Friendly's proposal to change *Miranda* warnings, we would attain the preferred results of the United Kingdom. The Supreme Court has cleared a path toward this result.[36]

Doyle v. Ohio was the high-water mark for excluding silence as evidence. *Doyle* barred disclosure of a defendant's pre-trial silence at trial once the defendant had received his *Miranda* warnings. The defendants in *Doyle* had been arrested and convicted for selling ten pounds of marijuana to a government

informant.[37] The defendants were given separate trials, and each defendant offered testimony claiming he had been framed.[38] The prosecution asserted that the defendants had sold the drugs to an informant, but both defendants claimed they had wanted to buy drugs from the informant and that the informant had set them up.[39] On cross-examination, the prosecutor questioned each defendant as to why they had not claimed they had been the victims of a "frame-up" upon arrest.[40] The defendants were convicted.

On appeal, the State argued "that the discrepancy between an exculpatory story at trial and silence at the time of arrest gives rise to an inference that the story was fabricated somewhere along the way."[41] Therefore, the Fifth Amendment allowed it to cross-examine defendants about their silence for the limited purpose of impeachment.[42] The Supreme Court rejected this argument.[43] According to the Court, cross-examination under these circumstances violated the Due Process Clause of the Fifth Amendment.[44] While the Court recognized that impeachment served an important purpose,[45] it held that "[s]ilence in the wake of [*Miranda*'s] warnings may be nothing more than the arrestee's exercise of these *Miranda* rights."[46] Without much explanation, the Court found that the *Miranda* warnings, and not the Fifth Amendment, implied a defendant's silence could carry no penalty.[47]

Doyle presumes a lot about the defendant's reliance on the warnings:

> If (a) the defendant is advised that he may remain silent, and (b) he does remain silent, then we (c) presume that his decision was made in reliance on the advice, and (d) conclude it is unfair in certain cases, though not others, to use his silence to impeach his trial testimony."[48]

This presumption extends beyond our understanding that the privilege of silence may be waived in matters to which the defendant testifies.[49] It would become irrelevant if the *Miranda* warnings were changed. If *Miranda*'s promised "right to remain silent" were altered to advise that any silence might be used to impeach the defendant later, the defendant's silence could be commented on at trial.[50]

The Supreme Court appears to have recognized this and has limited the reach of *Doyle*. *Portuondo v. Agard* hinted that given the right circumstances "there might be reason to reconsider *Doyle*."[51] The Court has also allowed a prosecutor to cross-examine an un-*Mirandized* defendant about his failure to offer his exculpatory testimony post-arrest because "*Doyle* [only] bars the use against a criminal defendant of silence maintained after receipt of governmental assurances."[52] *Doyle*'s dissent provides the basis for this argument.[53] If we assume that the government did not assure anything, but warned that failure to disclose an alibi now could harm claims of an alibi at trial, then the defendant did not rely on a government assurance that silence would protect him. Adoption of Britain's warnings would negate *Doyle*'s presumptions.

Once amended warnings remove the presumption that post-arrest silence is inadmissible, a judge may advise the jury that it may make inferences from contradictory testimony offered at trial. *Portuondo v. Agard* prohibits a prosecutor from urging the jury to infer something from the defendant's refusal to testify but allows the jury, "in evaluating the relative credibility of a defendant," to consider something that would be "natural and irresistible" for a juror to consider.[54] If a jury were told that a defendant had been accused of a crime, and informed that evidence showing his innocence would help him, yet he remained silent and offered an alibi at trial, it would be natural for the jury to conclude that the defendant lied on the stand. It is a comment "in accord with [the Court's] longstanding rule that when a defendant takes the stand" he may constitutionally have his credibility as a witness impeached.[55]

Silence should be evidence when a valid inference follows from it. There is nothing strange about treating silence as an incriminating act, providing evidence of guilt. In *Raffel v. United States*, a conspiracy suspect declined to testify in his own defense in the first trial, but took the stand in his second trial.[56] When he testified at his second trial, he denied making incriminating statements to the arresting prohibition agent who had testified against him at both trials.[57] The trial court questioned Raffel about his choice not to deny the statements at the first trial.[58] The questions caused Raffel to explain why he had remained silent at the first trial.[59] The U.S. Supreme Court found the questioning to be consistent with the Fifth Amendment prohibition against self-incrimination, reasoning that the government can consider a defendant's decision to testify as an all-or-nothing proposition exposing him to impeachment.[60] "The safeguards against self-incrimination are for the benefit of those who do not wish to become witnesses in their own behalf and not for those who do."[61]

A similar situation arose in *Baxter v. Palmigiano*.[62] In *Baxter*, a prisoner in a disciplinary proceeding was advised that "he had a right to remain silent during the hearing but that if he remained silent his silence would be held against him."[63] Relying on *Miranda,* the lower court decided inmates are entitled to representation when charges involve conduct punishable as a state crime.[64] The Supreme Court declined to extend *Miranda* that far.[65] Instead, the Supreme Court found that where a prisoner had been advised of his right to remain silent, and advised that his silence could be used against him, there was not a Fifth Amendment violation[66]:

> [Palmigiano] remained silent at the hearing in the face of evidence that incriminated him; and, as far as this record reveals, his silence was given no more evidentiary value than was warranted by the facts surrounding his case. This does not smack of an invalid attempt by the state to compel testimony without granting immunity or to penalize exercise of the privilege.[67]

Baxter was again discussed in *McKune v. Lile*, a case upholding Kansas' rule that prisoners involved in the Sexual Abuse Treatment Program (SATP) admit responsibility for prior criminal acts, without promises of government immunity.[68] If a prisoner refused to admit responsibility, he could not participate; and his prison privilege status was negatively affected.[69] Lile claimed this was compulsion, but the Court found valid reasons for requiring admission, and denying immunity.[70] The Constitution does not require individuals to be left with the impression that society will not punish them for serious past offenses.[71] Instead, the "constitutional guarantee is

only that the witness not be *compelled* to give self-incriminating testimony."[72] "Determining what constitutes unconstitutional compulsion involves a question of judgment: Courts must decide whether the consequences of an inmate's choice to remain silent are closer to the physical torture against which the Constitution clearly protects or the *de minimis* harms which it does not."[73] In Lile's case, loss of the privilege did not rise to the level of unconstitutional compulsion[74]: "Although a defendant may have a right, even of constitutional dimension, to follow whichever course he chooses, the constitution does not by that token always forbid requiring him to choose."[75]

The Fifth Amendment does not prohibit informed choice; rather, "it prohibits only the compulsion of such testimony."[76] Likewise, the Constitution does not always prohibit requiring a defendant to choose, as "[t]here is a difference between the sorts of penalties that would give a prisoner a reason not to violate prison disciplinary rules and what would compel him to expose himself to criminal liability."[77] The difference is clear if one acknowledges that revised warnings would not compel exposure of crime. They would simply relay the consequences of silence during interrogation when the accused acts as his own witness and offers exculpatory evidence at trial. Silence is not conclusive evidence of guilt. The jury may draw natural inferences from the defendant's silence. Silence correctly becomes one more item to be weighed in the evidentiary balance.[78]

III. Benefits of Revision

History shows that we can revise *Miranda*'s warnings, but why should we? That's easy. Revision will advance our justice system's search for the truth. Consider the benefits: a revised warning encourages the guilty to confess, instead of encouraging career criminals to remain silent; police can follow up on leads, collect more evidence, and arrest the guilty party; crimes will be solved faster, meaning fewer victims will have to wait for resolution; and juries will have more evidence to consider. These are good things, as they will help restore balance to the criminal justice system.

(1) Confessions

Interrogation is critical in police investigations. Without interrogation, "those who were innocent might be falsely accused, those who were guilty might wholly escape prosecution, and many crimes would go unsolved."[79] The criminal justice system should encourage freely given confessions instead of rewarding the criminal exploiting his privilege to remain silent—at the expense of public safety—during police interrogation.[80] A confession that does not create the specter that "adverse consequences can be visited upon the [defendant] by reason of further testimony" does not implicate the Fifth Amendment because "there is no further incrimination to be feared."[81] Asking the accused to give up exonerating evidence does not trigger adverse inferences unless the accused lies or contradicts himself on the stand.

This is consistent with our original understanding of what the Fifth Amendment protected. Historically, we sought to prevent compulsion, and even now,

> [t]he ultimate test remains that which has been the only clearly established test in Anglo-American courts for two hundred years: the test of voluntariness. Is the confession the product of an essentially free and unconstrained choice by its maker? If it is, if he has willed to confess, it may be used against him. If it is not, if his will has been overborne and his capacity for self-determination critically impaired, the use of his confession offends due process.[82]

If a confession met this test, it was considered reliable admissible evidence.[83] *Miranda* limited the use of confessions by informing the suspect that he had a right to remain silent, and if he was not so informed, his confession could not be used as evidence.[84] The warnings hindered the search for the truth. For example, after *Miranda*'s pronouncement, Baltimore reportedly saw the number of suspects willing to confess drop from 20%-25% to 2%.[85] The positive results observed in the United Kingdom after it adopted its revised warning demonstrates how the search for truth may be better served with different *Miranda* warnings.[86]

Before CJPOA, suspects enjoyed a broad right to silence, and studies attempted to estimate how often this right was invoked.[87] One study conducted before CJPOA went into effect interviewed police officers from ten different police stations and found that the accused offered "no comment" 10% of the time.[88] Suspects selectively answered questions 13% of the time.[89] A follow-up study found a decreased reliance on silence during police interviews.[90] The percentage of suspects who refused to answer every question fell from 10% to 6%, and the number of suspects who answered selective questions fell from 13% to 10%.[91] Within two years, the number of suspects who answered all questions during the police interviews increased 7%.[92]

Even those with legal advice were willing to talk after the warnings had been given.[93] Prior to 1994, 20% of suspects receiving legal advice had refused to answer questions.[94] In 2000, this number dropped to 13%.[95] The study's authors hypothesized that legal advisers must be counseling their clients to provide an account to police if they can,[96] as some defense lawyers view the warnings as a defensive tool[97]:

> In a way it's probably helped us because it's thrown the emphasis back onto the police in that we obviously require a disclosure before we advise clients. "We're not going to answer your questions, because it is on tape that you're not prepared to disclose what your evidence is. Therefore how can we advise the clients in the proper manner?" So that straight away throws the emphasis back on the officer.[98]

Officers may now be required to lay out evidence before the accused is required to answer questions, but the United Kingdom has found this tit-for-tat process beneficial, since the most frequent invokers of the right of silence, the serious offenders, are the ones who offer up statements post-CJPOA.[99] These statements aid law enforcement because even "cock and bull" information gives police "something concrete to check up on," and allows inferences to be drawn at trial if the suspect changes his story.[100] A good example was offered by one legal adviser:

> I can think of two or three particular villains that I regularly used to represent and they were always "no comment" pre the Act. Since the Act one of them has moved from being

> a professional burglar to the drug scene and on the two occasions where he's been interviewed in my presence, he has given a limited interview, he has given an explanation for items being in his possession to avoid the special warning and he has given an explanation for his general conduct. But the [professional criminals] in this area tend to give a statement… on tape, rather than submit themselves to questions.[101]

United Kingdom studies have not found compulsion, but have demonstrated that police must establish a stronger case before they bring a suspect to the police station. The warning itself is designed to inform the accused of his rights during interrogation, and it allows the accused to choose how to proceed. He may request an attorney or choose to answer a limited number of questions, and if he chooses to speak, the danger of impeachment is minimized. Evidence does not support a conclusion that the warnings compel confessions.[102] A study shows that CJPOA caused the proportion of silent suspects to fall but did not increase the percentage of suspects making admissions.[103] Admissions remained consistently at 55%.[104]

Police may surrender more of their case evidence, but they obtain more in the process. This helps prosecutors decide whom to charge and meet evidentiary burdens. CJPOA's warnings help unravel the truth behind a crime.

(2) Restoring the balance

Critics will argue that an adverse inference from silence allows police to drum up false charges against a defendant. One study, however, found that the inference actually reduced the number of silent suspects eventually charged with an offense.[105] This makes sense, because silence will neither save an already weak case, nor alleviate the prosecution's burden of proving guilt beyond a reasonable doubt.[106] Choosing silence during interrogation may not hurt the defendant, and more often than not, will place heavy burdens on a prosecutor.[107] Justice—punishment of the guilty while the innocent walk free—cannot be achieved if law enforcement is unnecessarily saddled with heavy burdens, and the jury is prohibited from inferring anything. To lessen the burden, a jury should be permitted to draw natural inferences from a suspect's conduct during interrogation. If the accused chooses to remain silent, a judge should be permitted to instruct on this silence. It will provide a little more information to help determine who to believe.

This would prevent silence from allowing the defendant to engage in gamesmanship at trial. In areas without an "ambush" statute, a silent defendant could present a surprise alibi at trial. Before the United Kingdom adopted CJPOA's warnings, English courts saw defendants present defenses for the first time at trial in 7% to 10% of its cases.[108] The use of "ambush" defenses decreased after the warnings were implemented.[109] One observer noted: "If you've got a sophisticated criminal who's come up with a defence or explanation for his conduct very late in the day—in other words, a surprise defence—catching the prosecution completely unawares, I think that's where the provisions are really useful."[110]

The provisions are useful because they permit a jury to draw inferences from the defendant's new alibi and allow the prosecution to respond. A prosecutor may now tell a more complete story to the jury. It begins at the time of the alleged crime, goes through the physical evidence obtained by the prosecution, relates the facts of the defendant's interrogation, and concludes with the defendant's conduct at trial. A prosecutor cannot do this if the suspect's pre-trial silence is universally excluded. Revising *Miranda*'s warnings so that they no longer overprotect a defendant's right to silence,[111] and allow reliable evidence, will restore a "carefully crafted balance designed to fully protect *both* the defendant's and society's interests."[112]

Conclusion

The United Kingdom revised its "right to silence" warning 15 years ago. In the years since, it has seen little change in the rate of confession, but has observed that a suspect is willing to volunteer exonerating facts during interrogation. This is consistent with our understanding that the Fifth Amendment protects against self-incrimination. It is consistent with the understanding that our Supreme Court appears to be restoring. *Griffin, Doyle*, and *Mitchell* extended a pro-defendant stance, but *Portuondo v. Agard* advised that silence might be used to impeach a defendant who testifies at trial. The current *Miranda* warnings do not advise the defendant that he might be impeached. If we were to adopt the United Kingdom's warnings, this could further the truth-seeking function of trial. It is a win-win situation.

"The Fifth Amendment guarantees that no person shall be compelled to give evidence against himself, and so is violated whenever a truly coerced confession is introduced at trial, whether by way of impeachment or otherwise."[113] But the Fifth Amendment is not necessarily violated if voluntary silence is introduced for impeachment purposes.[114] Silence that is probative of the defendant's character, the plausibility of his story, or even his guilt should not be excluded when no reasonable person questions its reliability. Revising *Miranda*'s warnings to allow an adverse inference from silence will correct the presumption that any comment on pre-trial silence violates some Fifth Amendment "right to silence." Revision will promote an understanding that inadmissibility should not be automatic, but should be done only when the accused is truly compelled to self-incrimination.

Endnotes

1 Criminal Justice and Public Order Act 1994, s. 34 (d)(1) (available at: http://www.opsi.gov.uk/acts/acts1994/ukpga_19940033_en_1). Section 34 provides:

(1)Where, in any proceedings against a person for an offence, evidence is given that the accused

(a) at any time before he was charged with the offence, on being questioned or cautioned by a constable trying to discover whether or by whom the offence had been committed, failed to mention any fact relied on in his defence in those proceedings; or

(b) on being charged with offence or officially informed that he might be prosecuted for it, failed to mention any such fact, being a fact which in the

circumstances existing at the time the accused could reasonably have been expected to mention when so questioned, charged or informed, as the case by be, subsection (2) below applies.

(2) Where this subsection applies –

(a) a magistrates' court inquiring into the offence as examining justices;

(b) a judge, in deciding whether to grant an application made by the accused under—

(I) section 6 of the Criminal Justice Act of 1987,

(ii) paragraph 5 of Schedule 6 to the Criminal Justice Act 1991 (application for dismissal of charge of violent or sexual offence involving child in respect of which notice of transfer has been given under section 53 of that Act);

(c) the court, in determining whether there is a case to answer; and

(d) the court or jury, in determining whether the accused is guilty of the offence charged, may draw such inferences from the failure as appear proper.

(2A) Where the accused was at an authorised place of detention at the time of the failure, subsection (1) and (2) do not apply if he had not been allowed an opportunity to consult a solicitor prior to being questioned, charged or informed as mentioned in subsection (1) and (2) above do not apply if he had not been allowed an opportunity to consult a solicitor prior to being questioned, charged or informed as mentioned in subsection (1) above.

(3) Subject to any directions by the court, evidence tending to establish the failure may be given before or after evidence tending to establish the fact which the accused is alleged to have failed to mention.

(4) This section applies in relation to questioning by persons (other than constables) charged with the duty of investigating offences or charging offenders as it applies in relation to questioning by constables; and in subsection (1) above 'officially informed' means informed by a constable or any such person.

(5) This section does not—

(a) prejudice the admissibility in evidence of the silence or there reaction of the accused in the face of anything said in his presence relating to the conduct in respect of which he is charged, in so far as evidence thereof would be admissible apart from this section; or

(b) preclude the drawing of any inference from any such silence or other reaction of the accused which could properly drawn apart from this section.

2 *See* R. v. Breen, [2003] EWCA Crim 3817, ¶ 34-40, 2003 WL 23014902; Regina v. Rae, [2006] EWCA Crim. 734, ¶ 23-24, 2006 WL 1457366; R. v. Steele, Whomes, Correy, [2006]EWCA Crim. 195, ¶ 56-67, 2006 WL 1732536.

3 R v. Steele, Whomes, Correy, ¶ 632.

4 *See* CJPOA §38(3); R. v. Cowan, (1996) Q.B. 373 (U.K.).

5 *See* CJPOA §34(2)(d); R v. Cowan, (1996) Q.B. 373, 373-374 (U.K.).

6 R v. Steel, Whomes, Correy, ¶ 63.

7 *See* Rice v. Connolly [1966] 2 Q.B. 414, 419.

8 *See* Police and Criminal Evidence Act of 1984, Codes of Practice For the Detention, Treatment and Questioning of Persons by Police Officers (1991) Code C, ¶ 10(a), 33-34.

9 Code C, ¶ 10.1, 33.

10 Code C, §11(a), ¶ 11.1A, 37.

11 *Id.* §10(a), ¶ 10.3, 34

12 *Id.* at §10(a), ¶ 10.4, 34.

13 *Id.* §10(b), ¶ 10.5, ¶ 10.7, Note 10D, 34-35, 37.

14 *See Miranda*, 384 U.S. 436 (1966); Dickerson v. United States, 530 U.S. 428 (2000).

15 United States v. Burr, 25 F. Cas. 38, 40 (No. 14,692e) (CC Va. 1807).

16 *Miranda*, 384 U.S. 436, 442-444 (1966).

17 *Dickerson,* 530 U.S., at 437-438.

18 *Id.* at 444.

19 *Miranda*, 384 U.S., at 441-442; *Dickerson* 530 U.S., at 439.

20 *Id.* at 444 (allowing alternative *Miranda* warnings so long as alternative "inform[s] accused persons of their right of silence and to assure a continuous opportunity to exercise it...").

21 *Id.*

22 *Id.* at 467.

23 *See id.*

24 *Id.* at 444.

25 *Id.* at 467.

26 *Id.* at 469.

27 *Id.*

28 *See, e.g.,* P. Kauper, *Judicial Examination of the Accused—A Remedy for the Third Degree*, 30 Mich. L. Rev. 1224 (1932).

29 380 U.S. 609 (1965).

30 *See* H.J. Friendly, *The Fifth Amendment Tomorrow: The Case for Constitutional Change*, 37 U. Cin. L. Rev. 671 (1968).

31 *Id.* at 691.

32 *Id.* at 690-691.

33 *Id.* at 721-722.

34 *Id.*

35 *See* footnote 3, *supra*.

36 *See* Office of Legal Policy, *Report to Attorney General on Adverse Inferences From Silence*, 22 U. Mich. J. L. Reform 1005, 1007, 1013-1014 (1989) (arguing for broader disclosure of pretrial silence and advising that *Doyle v. Ohio* does not apply silence "if the defendant had been put on notice that his failure to talk could be used against him").

37 Doyle v. Ohio, 426 U.S. 610, 611 (1976).

38 *Id.* at 613.

39 *Id.*

40 *Id.* at 613-614.

41 *Id.* at 616.

42 *Id.*

43 *Id.* at 617.

44 *See id.* at 618.

45 *Id.* at 617 n. 7.

46 *Id.* at 617.

47 *See id.* at 618 n.9.

48 *Id.* at 620 (Stevens, J., *dissenting*) (footnote omitted).

49 *See* Brown v. United States, 356 U.S. 148, 154-155 (1958); Mitchell v. United States, 526 U.S. 314, 321 (1999).

50 For example, it might be treated as an inconsistent statement or an adoptive admission. *See Doyle*, 426 U.S. at 622 (Stevens, J., dissenting); F.R.E. 801(d)(2)(B); *see also* F.R.E. 801 (d)(1); L. McClelland, *Silence as an Admission in Criminal Trial in Pennsylvania*, 53 Dickinson L. Rev. 318 (1949); D. Barrett, *Admissibilty of Accusatory Statement as Adoptive Admissions when Defendant is Under Arrest,* 35 Cal. L. Rev. 128 (1947).

51 529 U.S. 61, 74 (2000).

52 Fletcher v. Weir, 455 U.S. 603, 606-607 (1982) (quoting Anderson v. Charles, 447 U.S. 404, 408 (1980)).

53 *See Doyle*, 426 U.S. at 620 (Stevens, J., dissenting).

54 Portuondo v. Agard, 529 U.S. at 67-68.

55 *Id.* at 69; *see also* Kansas v. Ventris, 555 U.S. ___, 129 S.Ct. 1841, 1846 (2009) (allowing "tainted evidence" for impeachment at trial); *cf.* Jenkins v. Anderson, 447 U.S. 231, 246 (1980) (Marshall, J., dissenting).

56 271 U.S. 494, 495 (1926).

57 *Id.*

58 *Id.* n.*.

59 *Id.*

60 *Id.* at 499.

61 *Id.*

62 425 U.S. 308 (1976).

63 *Id.* at 312.

64 *Id.* at 315.

65 *Id.*

66 *Id.* at 318.

67 *Id.*

68 536 U.S. 24, 34-35 (2002) (plurality).

69 *Id.* at 30-31.

70 *Id.* at 31-35.

71 *See id.* at 35-36.

72 *Id.* at 36 (quoting United States v. Washington, 431 U.S. 181, 188 (1977)) (emphasis in original).

73 *Id.* at 41.

74 *Id.*

75 *Id.* (quoting McGuatha v. California, 402 U.S. 183, 213 (1971)).

76 *Id.* at 49 (O'Connor, J., concurring).

77 *Id.* at 52 (O'Connor, J., concurring).

78 *See, e.g., Baxter,* 425 U.S., at 318.

79 Schneckloth v. Bustamonte, 412 U.S. 218, 225 (1973).

80 *See* Watts v. Indiana, 338 U.S. 49, 57-60 (1949) (Jackson, J., concurring).

81 Mitchell v. United States, 526 U.S. 314, 326 (1999).

82 *Schneckloth,* 412 U.S. at 225-226 (quoting Culombe v. Connecticut, 367 U.S. 568, 602 (1961)).

83 *See* Oregon v. Elstad, 470 U.S. 298, 306-307 (1985).

84 *Id.* at n.1.

85 *See* M. O'Neill, *Undoing* Miranda, 2000 B.Y.U. L. Rev. 185, 210-233 (2000) (describing Senate Judiciary Committee hearings that uncovered *Miranda*'s harmful effects on law enforcement).

86 *See* T. Bucke, R. Street & D. Brown, *The Right of Silence: the Impact of the Criminal Justice and Public Order Act 1994*, Home Officer Research Study (2000).

87 *Id.* at 30 (comparing Leng 1993 and Association of Police Officers studies).

88 C. Phillips & D. Brown, with the assistance of Z. James and P. Goodrich, (1998), *Entry into the Criminal Justice System: a Survey of Police Arrests and Their Outcomes,* Home Office Research Study No. 185, at 75.

89 *Id.*

90 Bucke, Street & Brown, *supra* note 86, at 31.

91 *Id.*

92 *Id.*

93 *Id.* at 32.

94 *Id.* at 32-33.

95 *Id.*

96 *Id.* at 32.

97 *See id.* at 23.

98 *Id.* (quoting a legal adviser).

99 *Id.* at 36-37.

100 *Id.* at 35.

101 *Id.* at 37.

102 *Id.* at 34.

103 *Id.*

104 *Id.*

105 *Id.* at x.

106 *See id.*

107 *See* Gallup Poll, Public Opinion, 1993, at 231-232 (1994). A 1993 Gallup poll asked a randomly selected nationwide sample: "Do you believe the criminal justice system makes it too hard for the police and prosecutors to convict people accused of crimes, or not?" 70% of the sample agreed, and only 27% answered in the negative. (Cited in B. Ingraham, *The Right of Silence, the Presumption of Innocence, the Burden of Proof, and a Modest Proposal: A Reply to O'Reilly*, 86 J. Crim. L. & Criminology 559, 560 n.4 (1996).

108 Bucke, Street & Brown, *supra* note 86, at 59 n.46 (citing M. Zander & P. Henderson, "Crown Court" study (1993)).

109 *Id.* at 59.

110 *Id.*

111 *See* Duckworth v. Eagan, 492 U.S. 195, 209 (1989) (O'Connor, J., concurring).

112 Moran v. Burbine, 475 U.S. 412, 433 n. 4 (1986) (emphasis in original).

113 Kansas v. Ventris, 555 U.S. ___, 129 S.Ct. 1841, 1845(2009).

114 *See id.* at 1846.

Whither the Rule of Lenity

*by Dan Levin & Nathaniel Stewart**

The "rule of lenity" "requires ambiguous criminal laws to be interpreted in favor of the defendants subjected to them."[1] Although long a favorite of defense attorneys, actual applications of the rule, at least at the Supreme Court level,[2] have been relatively rare. This is perhaps somewhat surprising as the rule's roots in due process principles, and potential application where a strict construction of a statute results in an ambiguity, could lead both traditionally liberal and traditionally conservative Justices to favor its use. In 2008, in *United States v. Santos,* the Supreme Court issued a plurality opinion holding that a key term in a federal money laundering statute was ambiguous and applied the rule of lenity to resolve the ambiguity in the defendants' favor. The plurality involved just such a coalition of conservative and liberal Justices (Justices Scalia, Thomas, Ginsburg, and Souter; with Justice Stevens writing separately and agreeing that the rule should apply), raising the question of whether the rule may be entering a period of somewhat greater application.

As noted, to date, the Supreme Court has applied the rule sparingly and "only when, after consulting the traditional canons of statutory construction, [the court is] left with an ambiguous statute."[3] As Justice Thomas noted in *Staples v. United States*, "[t]hat maxim of construction [the rule of lenity] is reserved for cases where, '[a]fter seiz[ing] every thing from which aid can be derived,' the Court is 'left with an ambiguous statute.'"[4] Similarly, the Court has described the rule as "appl[ying] only when the equipoise of competing reasons cannot otherwise be resolved...."[5]

Determining when the traditional canons have failed and an ambiguous statute remains, however, enjoys little consensus among members of today's Supreme Court. Because it is currently used only as an interpretative tool-of-last-resort, it is not surprising that the rule of lenity has not ultimately served to "break the tie" in many cases. After all, the Court may choose to interpret a criminal statute using any number or combination of the canons of construction in order to avoid declaring a statute hopelessly ambiguous.[6] Since the 2006-2007 term, for example, the rule of lenity has been mentioned or discussed in a majority, dissenting, or concurring opinion fewer than a dozen times,[7] and it has been applied and broken the tie in the defendant's favor only once—in the *Santos* case.[8]

Although it has rarely decided a case, the rule of lenity has been more frequently cited in dissenting opinions arguing that the statutory provision at issue is ambiguous enough to warrant the rule's application. Maybe this, too, should be expected in light of the rule's broad implications for and effect on criminal statutes.[9]

Because the rule of lenity can be applied in a manner that protects defendants' due process rights and also in a manner based on strict statutory construction, it is perhaps not surprising that at times it results in interesting coalitions that cross the Court's traditional conservative-liberal lines. As noted, Justice Scalia's plurality opinion in *Santos*[10] was joined by Justices Thomas, Ginsburg, and Souter; Justice Stevens concurred in the judgment and wrote separately, also endorsing application of the rule. Justice Alito filed a dissenting opinion joined by Chief Justice Roberts, Justices Kennedy and Breyer.

Other discussions of the rule of lenity in the most recent three terms have included the following: Justice Ginsburg acknowledging that the statutory definition in question was "not a model of the careful drafter's art" and yet declining to apply the rule,[11] while Chief Justice Roberts, joined by Justice Scalia, considered the case "a textbook case" for the rule of lenity;[12] Justice Souter and Justice Ginsburg joining with the traditionally more conservative members in declining to apply the rule of lenity,[13] while Justice Stevens[14] and Justice Breyer[15] each wrote dissenting opinions calling for its application; and Justice Scalia writing an opinion, joined by Justice Stevens and Justice Ginsburg, calling for application of the rule of lenity, in a case in which neither the majority opinion, authored by Justice Alito, nor Justice Thomas's separate dissent, even discussed the rule.[16]

A brief review of *Santos* and several of the recent cases discussing the rule highlights the confusing difficulty the Court faces in determining when to apply the rule of lenity. The Justices seem to agree on the rule's purpose and that the rule is one of "last resort," to be used when all other attempts to interpret the text have failed. But *when* those attempts have failed, and *when* the rule must be employed, remains murky and uncertain.

I. United States v. Santos

Santos and one of his collectors were convicted of money laundering charges related to their long-standing illegal lottery scheme.[17] Santos employed several operatives to manage an illegal lottery, including "runners" to collect bets at bars and restaurants and "collectors" who would deliver those bets to him.[18] Financial transactions between Santos and his employees and lottery winners formed the basis for money laundering charges and subsequent convictions under 18 U.S.C. § 1956(a)(1),[19] which criminalizes financial transactions involving "proceeds" of certain types of unlawful activities.[20]

The convictions were vacated by the lower court on grounds that "proceeds" means "profits" rather than "gross receipts," and the Government had failed to prove that the funds involved in the transactions represented "profits" from the lottery.[21] After the Seventh Circuit Court of Appeals affirmed the holding, the Supreme Court granted certiorari to determine the narrow question of whether "proceeds" means "profits" or "gross receipts," i.e. whether the government has to prove that the underlying criminal activity was profitable or

** Dan Levin and Nathaniel Stewart are white collar attorneys with White & Case LLP in Washington, DC.*

just that the money used in the transaction was a product of the criminal activity.[22]

The plurality held in favor of the two defendants by settling on the narrower of the two meanings. The money laundering statute did not define "proceeds," and, as Justice Scalia noted, "[w]hen a term is undefined, we give it its ordinary meaning."[23] But "proceeds" is equally capable of two "ordinary meanings"—either "receipts" or "profits."[24] In such a case, as Justice Scalia explained, the rule of lenity "requires ambiguous criminal laws to be interpreted in favor of the defendants subjected to them," since no individual should be held criminally liable for statutory offenses not clearly prescribed.[25] Thus, according to the plurality, "[b]ecause the 'profits' definition of "proceeds" is always more defendant-friendly than the 'receipts' definition, the rule of lenity dictates that it should be adopted."[26] Justice Scalia went on to argue that "[w]hen interpreting a criminal statute, we do not play the part of mind reader," and, quoting Justice Frankfurter, stated: "'When Congress leaves to the Judiciary the task of imputing to Congress an undeclared will, the ambiguity should be resolved in favor of lenity.'"[27]

The Government had made two primary arguments in favor of the "gross receipts" interpretation. First, gross receipts would more "accurately reflect the scale of the criminal activity" and thus better serve the purpose of the money laundering statute.[28] The plurality rejected this argument out of concern that such a broad interpretation would effectively "merge" any illegal gambling offense into a much more severe money laundering offense, "because paying a winning bettor is a transaction involving receipts that the defendant intends to promote the carrying on of the lottery."[29] Second, the Government argued for the "receipts" interpretation "because—quite frankly—it is easier to prosecute."[30] Justice Scalia rejected this position because it "[e]ssentially... asks us to resolve the statutory ambiguity in light of Congress's presumptive intent to facilitate money-laundering prosecutions," a position which "turns the rule of lenity upside-down. We interpret ambiguous criminal statutes in favor of defendants, not prosecutors."[31]

Concurring in the judgment as the decisive fifth vote, Justice Stevens noted at the outset that "[w]hen Congress fails to define potentially ambiguous statutory terms, it effectively delegates to federal judges the task of filling gaps in a statute."[32] Justice Stevens argued that Congress has, in other contexts, and could have here "defined 'proceeds' differently when applied to different specified unlawful activities," and therefore judges may do the same "as long as they are conscientiously endeavoring to carry out the intent of Congress."[33] Thus, Justice Stevens would not pick a single definition of "proceeds," but would define the term differently depending on the type of unlawful activity that produces the funds in question.[34] Ultimately, Justice Stevens concluded that the rule of lenity required the narrower "profits" interpretation for money laundering charges based on illegal gambling transactions because the statutory text and its legislative history did not clearly indicate congressional intent regarding the defendants' gambling operation.[35]

Justice Alito's dissent, joined by Chief Justice Roberts, Justice Kennedy, and Justice Breyer, argued that Congress intended the term "proceeds" to be defined as "gross receipts" for any unlawful activity under the statute.[36] The dissent focused on the legislative history of the statute and cited similar definitions in other statutes, including every state money laundering statute, as well as the prosecutorial burdens created by the plurality's definition, and concluded that the "meaning of 'proceeds' in the money laundering statute emerges with reasonable clarity when the term is viewed in context, making the rule of lenity inapplicable."[37]

II. Other Recent Rule of Lenity Cases

In two cases since *Santos*, application of the rule of lenity has been rejected, making clear that even if *Santos* may herald some greater receptivity to the rule, it is still likely to be sparingly applied.

A. United States v. Hayes

In *United States v. Hayes*, Justice Ginsburg, writing for the majority, reversed the Fourth Circuit's application of the rule of lenity, and held that the rule did not apply because the statute's "text, context, purpose, and what little drafting history there is all point in the same direction."[38] *Hayes* concerned the Gun Control Act of 1968,[39] which prohibits a person convicted of a "misdemeanor crime of domestic violence" from possessing a firearm. The defendant was charged with three counts of possessing firearms after being convicted of a crime of domestic violence, but he moved to dismiss the indictment on grounds that the state statute "under which he was convicted in 1994... was a generic battery proscription, not a law designating a domestic relationship between offender and victim as an element of the offense."[40] The district court denied defendant's motion to dismiss, and he pleaded guilty and appealed. The Fourth Circuit reversed the conviction, holding that the predicate offense for a conviction under § 922(g)(9) must "have as an element a domestic relationship between the offender and the victim."[41]

The Supreme Court's decision turned on whether "misdemeanor crime of domestic violence" in § 921(a)(33)(A) requires that "the predicate misdemeanor identify as an element of the crime a domestic relationship between the aggressor and victim."[42] Section 921(a)(33)(A) provides:

> The term "misdemeanor crime of domestic violence" means an offense that—
>
> (i) is a misdemeanor under Federal, State, or Tribal law; and
>
> (ii) has, as an element, the use or attempted use of physical force, or the threatened use of a deadly weapon, committed by a current or former spouse, parent, or guardian of the victim, by a person with whom the victim shares a child in common, by a person who is cohabitating with or has cohabitated with the victim as a spouse, parent, or guardian, or by a person similarly situated to a spouse, parent, or guardian of the victim.

The majority first focused on the text of the statute and observed that

> as an initial matter... § 921(a)(33)(A) uses the word "element" in the singular, which suggests that Congress

> intended to describe only one required element.... Had Congress meant to make the latter as well as the former an element of the predicate offense, it likely would have used the plural "elements," as it has done in other offense-defining provisions.[43]

Justice Ginsburg then approached the text's syntax and found that

> [t]reating the relationship between aggressor and victim as an element of the predicate offense is also awkward as a matter of syntax. It requires the reader to regard "the use or attempted use of force, or the threatened use of a deadly weapon" as an expression modified by the relative clause "committed by." In ordinary usage, however, we would not say that a person "commit[s]" a "use." It is more natural to say that a person "commit[s]" an "offense."[44]

The majority went on to note that "[h]ad Congress placed the 'committed by' phrase in its own clause, set off from clause (ii) by a semi-colon or a line break, the lawmakers might have better conveyed that 'committed by' modifies only 'offense' and not 'use' or 'element.'"[45]

Furthermore, the Court rejected the Fourth Circuit's application of the "rule of the last antecedent," "under which 'a limiting clause or phrase... should ordinarily be read as modifying only the noun or phrase that it immediately follows.'"[46] According to the Court,

> [a]pplying the rule of the last antecedent here would require us to accept two unlikely premises: that Congress employed the singular "element" to encompass two distinct concepts, and that it adopted the awkward construction "commit" a "use." ... "Committed" retains its operative meaning only if it is read to modify "offense."[47]

Thus, the Court's textual analysis concluded that "[m]ost sensibly read, then, § 921(a)(33)(A) defines 'misdemeanor crime of domestic violence' as a misdemeanor offense that (1) 'has, as an element, the use [of force],' and (2) is committed by a person who has a specified domestic relationship with the victim."[48]

The majority then considered the statute's purpose and the "practical considerations" that "strongly support" its reading of the statute,[49] and found that "[b]y extending the federal firearm prohibition to persons convicted of 'misdemeanor crime[s] of domestic violence,'"[50] Congress sought to prevent domestic abusers who are not charged with or convicted of felonies from possessing firearms. The majority argued that "[c]onstruing § 922(g)(9) to exclude the domestic abuser convicted under a generic use-of-force statute (one that does not designate a domestic relationship as an element of the offense) would frustrate Congress' manifest purpose,"[51] which, the Court believed, was to "keep[] firearms out of the hands of domestic abusers" even if those abusers are not charged with or convicted of felonies.[52] The majority then noted that "[g]iven the paucity of state and federal statutes targeting *domestic* violence, we find it highly improbable that Congress meant to extend 922(g)(9)'s firearm possession ban only to the relatively few domestic abusers prosecuted under laws rendering a domestic relationship an element of the offense."[53]

The majority opinion concluded with a brief look at the scant legislative history of the statute, consisting of an earlier version of the law and a floor statement by the bill's sponsoring Senator, which included the statement that:

> Convictions for domestic violence-related crimes often are fore crimes, such as assault, that are not explicitly identified as related to domestic violence. Therefore, it will not always be possible for law enforcement authorities to determine from the face of someone's criminal record whether a particular misdemeanor conviction involves domestic violence, as defined in the new law.[54]

The Court acknowledged that "[t]he remarks of a single Senator are 'not controlling,' but, ... the legislative record is otherwise 'absolutely silent.'"[55]

Rejecting the defendant's contention that the statute's ambiguity called for application of the rule of lenity, the majority held that although the statute's definition of "misdemeanor crime of domestic violence" "is not a model of the careful drafter's art," the statute was not so ambiguous as to allow for the rule of lenity to apply.[56]

Whereas the *Hayes* majority did not think the statute ambiguous enough to apply the rule of lenity, the dissent, written by Chief Justice Roberts and joined by Justice Scalia, considered this "a textbook case for application of the rule of lenity."[57] The dissent rejected the majority's reading on textual, structural, and practical grounds, and concluded that the statute is so ambiguous that the rule of lenity should be applied.

Like the majority opinion, the dissent started with the statute's text and framed the question as "whether the definition of 'misdemeanor crime of domestic violence' in § 921(a)(33)(A) includes misdemeanor offenses with no domestic-relationship element."[58] The Chief Justice began by disagreeing with Justice Ginsburg's reading of the text, noting that "[t]he majority would read the 'committed by' phrase in clause (ii) to modify the word 'offense' in the opening clause of subparagraph (A), leapfrogging the word 'element' at the outset of clause (ii)."[59] Under the majority's reading, "[i]ndividuals convicted under generic use-of-force statutes containing no reference to domestic violence would therefore be subject to prosecution under § 922(g)(9)."[60] The dissent found this reading incongruous and preferred the Fourth Circuit's more "natural reading," which held that "'committed by' modifies the immediately preceding phrase: 'the use or attempted use of physical force, or the threatened use of a deadly weapon.'"[61] "Read this way," wrote the Chief Justice, "a domestic relationship is an element of the prior offense."[62]

The dissent also analyzed the structure of the statute to decipher its meaning, and concluded that "[t]he most natural reading of the statute... is that the underlying misdemeanor must have as an element the use of force committed by a person in a domestic relationship with the victim."[63] Chief Justice Roberts argued that "[t]he fact that Congress included the domestic relationship language in the clause of the statute designating the element of the predicate offense strongly

suggests that it is in fact part of the required element."[64] He contended that the majority's reading "requires restructuring the statute and adding words. The majority first must place the 'committed by' phrase in its own clause—set off by a line break, a semicolon, or '(iii)'—to indicate that 'committed by' refers all the way back to 'offense.'"[65] The dissent noted several other textual revisions required by the majority's reading and argued that they "are not insignificant revisions; they alter the structure of the statute[,] ... [which] is often critical in resolving verbal ambiguity."[66]

Turning to the majority's arguments concerning the statute's sparse legislative history—a single floor statement by a single Senator—the dissent stated that "[s]uch tidbits do not amount to much," especially when, as here, "the statement was delivered the day the legislation was passed and *after* the House of Representatives had passed the pertinent provision."[67] Thus, the dissent dismissed the relevance and "value of such statements due to their inherent flaws as guides to legislative intent, flaws that persist... in the absence of other indicia."[68]

Chief Justice Roberts concluded by turning to the rule of lenity: "Taking a fair view, the text of 921(a)(33)(A) is ambiguous, the structure leans in the defendant's favor, the purpose leans in the Government's favor, and the legislative history does not amount to much," thereby making this "a textbook case" for applying the rule of lenity.[69] Moreover, he wrote, "[i]t cannot fairly be said here that the text 'clearly warrants' the counterintuitive conclusion that a 'crime of domestic violence' need not have domestic violence as an element."[70]

B. Dean v. United States

Whereas Chief Justice Roberts argued in favor of the rule of lenity in his *Hayes v. United States* dissent, he authored the opinion rejecting the rule's application in *Dean v. United States*.[71] In *Dean*, the Chief Justice was joined by the traditionally more conservative justices along with Justices Ginsburg and Souter, while Justices Stevens and Breyer (who had both joined Justice Ginsburg in rejecting the rule of lenity arguments in *Hayes*) each wrote separately in calling for the rule of lenity to be applied.

Dean concerned the meaning of 18 U.S.C. 924(c)(1)(A), which imposes extra punishment for discharging a firearm during a "crime of violence or drug trafficking crime." The question in *Dean* was whether the statute requires that the defendant intended to discharge the firearm during the commission of his crime. The Court held that it does not.[72]

During the course of an armed bank robbery, Dean's gun accidentally discharged as he was removing money from the teller's drawer. The defendant was convicted of discharging a firearm during an armed robbery, in violation of 924(c)(1)(A)(iii), and was sentenced to a mandatory minimum of 10 years in prison.[73] Dean appealed, arguing that the gun fired accidentally and that "the sentencing enhancement... requires proof that he intended to discharge the firearm."[74] The Eleventh Circuit affirmed Dean's conviction and sentence.

In affirming the court of appeals, the majority rejected Dean's argument that "any doubts about the proper interpretation of the statute should be resolved in his favor under the rule of lenity."[75] After analyzing the statute's text, structure, and practical application, the majority determined that the statute is not so "grievously ambiguous" as to warrant the rule of lenity.[76]

The majority began by observing that the statute's text on its face "does not require that the discharge be done knowingly or intentionally, or otherwise contain words of limitation," and the Court refrained from reading words or elements into the statute.[77] The Court then turned to the structure of the statute and found that it too did not support Dean's contention that the sentence enhancement included an intent requirement. The majority noted that whereas subsection (ii) of the statute "expressly included an intent requirement" for the 7-year mandatory minimum sentence if a criminal brandishes a firearm,[78] "Congress did not, however, separately define 'discharge' to include an intent requirement."[79] The Court rejected Dean's argument that "even if the statute is viewed as silent on the intent question, that silence compels a ruling in his favor."[80] Dean argued that there is a presumption in criminal cases that "the Government [must] prove the defendant intended the conduct made criminal."[81] Chief Justice Roberts acknowledged that "[i]t is unusual to impose criminal punishment for the consequences of purely accidental conduct," but explained that "it is not unusual to punish individuals for the unintended consequences of their *unlawful* acts," citing the felony-murder rule, whereby "[i]f a defendant commits an unintended homicide while committing another felony, the defendant can be convicted of murder," as an example.[82] The majority observed that Dean was already guilty of illegal conduct, and that conduct was not accidental:

> The fact that the actual discharge of a gun under 924(c)(1)(A)(iii) may be accidental does not mean that the defendant is blameless. The sentence enhancement in subsection (iii) accounted for the risk of harm resulting from the manner in which the crime is carried out, for which the defendant is responsible.[83]

Taken together, the majority was convinced that the law's text and structure were sufficiently clear and unambiguous to deny Dean's claim that the rule of lenity should apply.

Justice Stevens and Justice Breyer, however, were not so convinced. Justice Stevens argued that the structure and history of the statute indicate that "Congress intended § 924(c)(1)(A)(iii) to apply only to intentional discharges" because it may be inferred that "Congress intended to impose increasingly harsh punishment for increasingly culpable conduct."[84] Justice Stevens points to the escalating sentences in subsections (i) – (iii) and argues that by implication, the 5-year to 7-year to 10-year progression should correspond to escalating degrees of culpability. Because the accidental discharge caused no harm, and because the defendant did not act intentionally in firing his gun, he therefore lacked a more culpable mens rea and should not be punished more harshly. Rather than read subsection (iii) as a strict-liability offense, Justice Stevens contended that the Court should have applied the "common-law presumption that provisions imposing criminal penalties require proof of mens rea," which, he argued, was "bolstered by the fact that we have long applied

the rule of lenity—which is similar to the mens rea rule in both origin and purpose—to provisions that increase criminal penalties as well as those that criminalize conduct."[85] He stated that he would apply this presumption in order to "avoid the strange result of imposing a substantially harsher penalty for an act caused not by an 'evil-meaning mind' but by a clumsy hand."[86]

Justice Breyer largely adopted the points made by Justice Stevens without much additional explanation, and then focused more narrowly on the rule of lenity, which he argued "tips the balance against the majority's position."[87] Justice Breyer believed "the discharge provision here is sufficiently ambiguous to warrant the application of that rule [of lenity]," but he offered little analysis for that view.[88] Instead, Justice Breyer argued that the rule of lenity should be applied because "in the case of a mandatory minimum [sentence], an interpretation that errs on the side of *exclusion* (an interpretive error on the side of leniency) still *permits* the sentencing judge to impose a sentence similar to... the statutory sentence even if that sentence... is not legislatively *required*."[89] In contrast, "an interpretation that errs on the side of *inclusion* requires imposing 10 years of additional imprisonment on individuals whom Congress would not have intended to punish so harshly."[90] Such an "inclusive" error would remove discretion from the sentencing judge and perhaps "depart dramatically" from what Congress had intended.[91]

III. The Newest Justice's Views

Justice Sonia Sotomayor has joined the Supreme Court for the 2009-2010 Term, replacing Justice Souter, who joined in the *Santos* opinion applying the rule and also penned only one recent dissenting opinion favoring its use.[92] It is natural to ask whether this change in personnel may herald any greater—or lesser—receptivity to the use of the rule.

Of course practice on a lower court is not always a reliable predictor of practice on the Supreme Court. But it is worth noting that during her tenure on the lower courts, then-Judge Sotomayor authored seven opinions in which the rule was at issue—six during her time on the Second Circuit Court of Appeals,[93] and one as a district court judge in the Southern District of New York.[94]

In each of the six circuit court decisions in which she discussed the rule of lenity, Judge Sotomayor rejected the arguments in its favor, finding the statute in question was sufficiently clear using the traditional canons of interpretation. As discussed above, even in otherwise difficult cases the rule of lenity rarely proves dispositive, so it is not surprising that Judge Sotomayor did not find the challenged statutes so "grievously ambiguous" as to require the judicial rule's application. In four of those six cases, Judge Sotomayor declined the defendant's invitation to apply the rule in summary fashion, with relatively little discussion or explanation.[95] In two related circuit court opinions, however, she discussed the rule of lenity and its meaning at some length before concluding that it did not apply.

In *Sash v. Zenk* (*Zenk I)*,[96] and in the petition for rehearing that case (*Zenk II*),[97] Judge Sotomayor, writing for unanimous panels, held that the rule of lenity did not apply to the calculation of credits awarded to federal prisoners for good behavior, governed by 18 U.S.C. § 3624(b). In *Zenk I*, the court explained that "[t]he rule of lenity has two purposes: first, to ensure that the public receives fair notice of what behavior is criminal and what punishment applies to it; and second, to ensure that legislatures and not courts define criminal activity."[98] But because the statute at issue was not a criminal statute, the rule of lenity was irrelevant.[99]

On a petition for rehearing, *Zenk II* addressed the defendant's arguments that the court had erred in its earlier analysis because the Supreme Court had previously held that sentencing credit calculations were "criminal for purposes of an *ex post facto* analysis."[100] The court in *Zenk II* sought to clarify its earlier holding "to avoid any confusion,"[101] and drew the distinction between the rule of lenity and *ex post facto* doctrine. Acknowledging that the two rules are related and that "both are concerned with notice and fair warning,"[102] Judge Sotomayor distinguished between their purposes:

> The rule of lenity concerns situations in which a legislature fails to give notice of the scope of punishment by leaving "a grievous ambiguity or uncertainty in the language and structure of the [statute]....," while the *ex post facto* doctrine "concerns situations in which the legislature gives adequate notice, but then affirmatively changes its instructions in a way that disadvantages the defendant."[103]

Accordingly, the court observed, the rule of lenity is the narrower doctrine, and "should be more narrowly applied."[104] Judge Sotomayor went on to explain:

> The reason the *ex post facto* doctrine is broader than the rule of lenity in the area of sentencing administration is that there is a greater potential for unfairness when a legislature changes the law pertaining to a criminal offender's sentence than when the legislature merely leaves a question open for future regulation by an administrative agency....
>
> The rule of lenity, however, deals with different concerns and employs a different analysis, and so it is not remarkable that the scopes of these doctrines should also differ or that we should consider a particular statute to be "criminal" in a way that implicates one doctrine but not the other.[105]

In light of her view that the rule of lenity should be "narrowly applied," it is perhaps unsurprising that Judge Sotomayor declined to apply the rule in each of her opinions on the Court of Appeals.

In contrast, as a district judge in *United States v. Westcott*,[106] Judge Sotomayor found that the "defendant's reading of [the statute] is as reasonable as the government's, and that the rule of lenity therefor[e] requires that the provision be applied according to defendant's interpretation."[107] *Westcott* concerned a defendant who had been convicted of robbery and then deported from the United States to Jamaica. He illegally reentered the United States several years later and pleaded guilty to reentering "after being deported subsequent to the

commission of an aggravated felony, in violation of 8 U.S.C. § 1326(b)(2)."[108]

Section 501 of the Immigration Act of 1990 amended the definition of "aggravated felony" to include robbery, and the issue in *Westcott* was whether the "effective date" provision in § 501(b) expanded the definition of aggravated felony to reach the defendant's earlier robbery conviction and therefore made him an aggravated felon. Section 501(b) states:

> Effective date—The amendments made by subsection (a) shall apply to offenses committed on or after the date of enactment of this Act, except that the amendments made by paragraphs (2) and (5) of subsection (a) shall be effective as if included in the enactment of 7342 of the Anti-Drug Abuse Act of 1988.

The expanded definition of "aggravated felony" to include "any crime of violence" was made in paragraph 3 of 501(a), and was therefore not one of the enumerated paragraphs to be effective as if enacted in the Anti-Drug Abuse Act. But, as the Government argued, the "offense" at issue in the case was the "defendant's illegal reentry, which occurred 'after the date of the enactment of [the Immigration] Act.'"[109] Thus, according to the Government, the expanded definition of aggravated felony would apply and reach the defendant's robbery conviction. The defendant countered that the "offenses" referred to in § 501(b) "are limited to those offenses delineated as aggravated felonies in § 501(a) of the Act," which would therefore not reach back to his earlier robbery conviction.[110]

Judge Sotomayor recognized that several circuit courts were divided on this issue, with the Fourth and Fifth Circuits adopting the Government's view and the Ninth Circuit, in an en banc decision, unanimously taking the defendant's position.[111] In holding that the rule of lenity should apply in this case, Judge Sotomayor acknowledged that the positions taken by the Fourth and Fifth Circuits were "plausible interpretation[s] of the Immigration Act," but that the "structure and language of section 501 of the Immigration Act" support the Ninth Circuit's "natural and reasonable reading" of the statute.[112] Judge Sotomayor began by analyzing the text of the statute and found that "[i]n short, section 501(b) can reasonably and naturally be construed to provide that most of those crimes set forth in section 501(a)—including crimes of violence—are aggravated felonies only to the extent that they occurred after November 29, 1990."[113] Then, finding "no real guidance" in the statute's legislative history, the judge noted that although the Government had offered a plausible interpretation of § 501(b), it provided "no arguments which unambiguously preclude the Ninth Circuit's reading of that same provision."[114] This is perhaps a strange formulation of the Government's burden, requiring it to present arguments which "unambiguously preclude" another court's reading of a statute, but she went on to explain that for each of the Government's interpretations there was an equally plausible alternative way to read the text.[115] She concluded, therefore, that "the Ninth Circuit has identified an interpretation of section 501 which favors [the] defendant, which appears reasonable, and which cannot be rejected through the applicable tools of statutory construction."[116] Thus, in cases where the statute is ambiguous and capable of two reasonable, competing meanings, the rule of lenity "'requires the sentencing court to impose the lesser of two penalties'"[117] and "assures defendant the benefit of the doubt."[118] On appeal, the Second Circuit affirmed the defendant's sentence without reaching the rule of lenity question.[119]

As noted, divining how any Justice will act based on their prior record is, at best, a hazardous task. Justice Sotomayor's lower court decisions, however, suggest that she takes the rule seriously while applying it, in her own words, "narrowly." If so, her addition to the Court in place of Justice Souter is unlikely to mark any significant departure in the application of the rule.

IV. Conclusion

The rule of lenity will undoubtedly remain a favorite among defense counsel. Given its roots in due process notice principles, it is perhaps somewhat surprising that it has not received more traction in coalitions of traditionally conservative, strict constructionist Justices and more liberal Justices. It is possible that *Santos* may signal a greater willingness of such coalitions of Justices to apply the rule in the future, and the Supreme Court's 2009-2010 Term includes at least three cases in which the rule may be discussed or applied.[120] *Hayes* and *Dean* strongly suggest, however, that there has been no radical change yet and that, at least for the foreseeable future, the rule is still likely to be sparingly used.

Endnotes

1 United States v. Santos, 128 S.Ct. 2020, 2025 (2008).

2 This article is limited to the Supreme Court's recent application of the rule of lenity and does not examine how frequently the rule is used in the lower courts.

3 United States v. Shabani, 513 U.S. 10, 17 (1994); *see also* Callanan v. United States, 364 U.S. 587, 596 (1961) ("The rule comes into operation at the end of the process of construing what Congress has expressed.").

4 Staples v. United States, 511 U.S. 600 n.17 (1994) (denying the rule's application) (quoting Smith v. United States, 508 U.S. 223, 239 (1993)).

5 Johnson v. United States, 529 U.S. 694, 713 n.13 (2000).

6 The use of these other canons may help explain the relative rarity of the rule's use at the Supreme Court level. For example, in *McNally v. United States*, 483 U.S. 350 (1987), the Court limited the reach of the mail fraud statute using language sounding much like the rule of lenity but without actually invoking the rule:

> The Court has often stated that when there are two rational readings of a criminal statute, one harsher than the other, we are to choose the harsher only when Congress has spoken in clear and definite language. As the Court said in a mail fraud case years ago: "There are no constructive offenses; and before one can be punished, it must be shown that his case is plainly within the statute." Rather than construe the statute in a manner that leaves its outer boundaries ambiguous and involves the Federal Government in setting standards of disclosure and good government for local and state officials, we read § 1341 as limited in scope to the protection of property rights. If Congress desires to go further, it must speak more clearly than it has.

Id. at 360 (internal citations omitted). Indeed, the dissent appears to have assumed that the majority was applying the rule (or doctrine) of lenity:

> To support its crabbed construction of the Act, the Court makes a straightforward but unpersuasive argument. Since there is no explicit,

> unambiguous evidence that Congress actually contemplated "intangible rights" when it enacted the mail fraud statute in 1872, the Court explains, any ambiguity in the meaning of the criminal statute should be resolved in favor of lenity. The doctrine of lenity is, of course, sound, for the citizen is entitled to fair notice of what sort of conduct may give rise to punishment. But the Court's reliance on that doctrine in this case is misplaced for several reasons.

Id. at 482-83.

7 *See, e.g.,* Boyle v. United States, 129 S.Ct. 2237 (2009) (majority refusing to apply the rule of lenity); Dean v. United States, 129 S.Ct. 1849 (2009) (Stevens, J., and Breyer, J., dissenting) (each calling for the rule of lenity).

8 United States v. Santos, 128 S.Ct. 2020 (2008). The Court has found it somewhat easier to conclude that the rule does not apply in a particular case, but even in those instances unanimity has been elusive. In *Boyle v. United States*, 129 S.Ct. 2237 (2009), for example, the Court refused to apply the rule of lenity, but the dissent did not comment on this point. As far as the authors are aware, Justice Ginsburg's opinion in *Burgess v. United States* is the only recent unanimous decision in which the Court held that the rule of lenity could not be invoked.

9 This raises the additional question as to whether the rule retains much force as a practical matter for deciding cases in the defendant's favor, or whether it is largely a tool used by the dissent to construe the statute in question as too ambiguous to be read as the majority has read it. This question lies beyond the scope of this article.

10 128 S.Ct. 2020 (2008).

11 United States v. Hayes, 129 S.Ct. 1079, 1089 (2009).

12 *Id*. at 1093 (Roberts, C.J., dissenting).

13 Dean v. United States, 129 S.Ct. 1849, 1856 (2009).

14 *Id*. at 1859 (Stevens, J., dissenting).

15 *Id*. at 1860 (Breyer, J., dissenting).

16 James v. United States, 127 S.Ct. 1586 (2007).

17 *Id.* at 2022.

18 *Id.* at 2023. The co-defendant in the case was one of Santos' collectors.

19 *Id*.

20 Money Laundering Control Act of 1986, 18 U.S.C. § 1956(a)(1) (2006).

21 *Santos*, 128 S.Ct. at 2023.

22 *Id.* at 2022.

23 *Id*. at 2024.

24 *Id*.

25 *Id*. at 2025.

26 *Id*.

27 *Id*. at 2026 (quoting Bell v. United States, 349 U.S. 81, 83 (1955)).

28 *Id*.

29 *Id.*

30 *Id*. at 2028.

31 *Id*. Within one year of the *Santos* decision, Congress responded with the Fraud Enforcement and Recovery Act of 2009 (FERA), Pub. L. No. 111-21, 123 Stat. 1617. Designed to bolster the government's anti-fraud capabilities, FERA, *inter alia*, amended the Money Laundering Control Act's definition of "proceeds" (§ 2(f)(1)(9)) to specifically include "gross receipts" of unlawful activity, as well as "any property derived from or obtained or retained, directly or indirectly, through some form of unlawful activity."

32 *Santos*, 128 S.Ct. at 2031 (Stevens, J., concurring).

33 *Id*. at 2032 (Stevens, J., concurring).

34 *Id*.

35 *Id.* Justice Stevens agreed with Justice Alito that "the legislative history of § 1956 makes it clear that Congress intended the term 'proceeds' to include gross revenues from the sale of contraband and the operation of organized crime syndicates involving such sales." *Id*. But he disagreed with the dissent because "that history sheds no light on how to identify the proceeds of many other types of specified unlawful activities." *Id*.

36 *Id.* at 2044.

37 *Id*. at 2054 (Alito, J., dissenting).

38 United States v. Hayes, 129 S.Ct. 1079, 1089 (2009).

39 18 U.S.C. § 921(g)(9).

40 *Hayes*, 129 S.Ct. at 1083.

41 United States v. Hayes, 482 F.3d 749, 751 (4th Cir. 2007).

42 *Hayes*, 129 S.Ct. at 1082.

43 *Id*. at 1084.

44 *Id*. at 1085.

45 *Id*.

46 *Id*. at 1086.

47 *Id.* at 1086-87.

48 *Id*. at 1087.

49 *Id*.

50 *Id*.

51 *Id*.

52 *Id*.

53 *Id.* at 1087-88.

54 *Id*. at 1088.

55 *Id*.

56 *Id*. at 1089.

57 *Id*. at 1093 (Roberts, C.J., dissenting).

58 *Id.* at 1089 (Roberts, C.J., dissenting).

59 *Id*.

60 *Id*.

61 *Id*.

62 *Id*.

63 *Id*. at 1091

64 *Id*.

65 *Id.*

66 *Id*.

67 *Id*. at 1092.

68 *Id*.

69 *Id*. at 1093.

70 *Id*.

71 129 S.Ct. 1849, 1856 (2009).

72 *Dean*, 129 S.Ct. at 1853.

73 *Id*.

74 *Id*.

75 *Id*. at 1856.

76 *Id*. The Court quotes *Muscarello v United States*, 524 U.S. 125, 138 (1998) ("The simple existence of some statutory ambiguity, however, is not sufficient to warrant application of that rule, for most statutes are ambiguous to some degree.").

77 *Dean*, 129 S.Ct. at 1853.

78 *Id.* (observing that Congress defined "brandish" to mean "to display all or part of the firearm, or otherwise make the presence of the firearm known to another person, *in order to intimidate* that person." 924(c)(4) (emphasis

added)).

79 *Id*. at 1854.

80 *Id*. at 1855.

81 *Id*.

82 *Id*.

83 *Id*.

84 *Id*. at 1857 (Stevens, J., dissenting).

85 *Id*. at 1858-59.

86 *Id*. at 1859.

87 *Id*. at 1860 (Breyer, J., dissenting).

88 *Id*. at 1861.

89 *Id*. at 1860.

90 *Id*.

91 *Id*. at 1861.

92 *See* United States v. Rodriguez, 128 S.Ct. 1783, 1793 (2008) (Souter, J., dissenting).

93 *See* United States v. Giordano, 442 F.3d 30 (2d Cir. 2006); Sash v. Zenk, 439 F.3d 61 (2d Cir. 2006) (rehearing); Sash v. Zenk, 428 F.3d 132 (2d Cir. 2005); United States v. Maloney, 406 F.3d 149 (2d Cir. 2005); United States v. Reinoso, 350 F.3d 51 (2d Cir. 2003); United States v. Figueroa, 165 F.3d 111 (2d Cir. 1998).

94 United States v. Westcott, 966 F.Supp. 186 (S.D.N.Y. 1997).

95 *See* United States v. Giordano, 442 F.3d 30, 40 (2d Cir. 2006) ("Because we find that the statute unambiguously reaches intrastate use of a telephone, we decline Giordano's invitation to apply the rules of lenity and constitutional avoidance to guide our interpretation."); United States v. Maloney, 406 F.3d 149, 153 n.6 (2d Cir. 2005) (rejecting the rule of lenity argument in a footnote); United States v. Reinoso, 350 F.3d 51, 55-56 (2d Cir. 2003) (summarily rejecting the rule of lenity argument in a section-closing paragraph); United States v. Figueroa, 165 F.3d 111, 119 (2d Cir. 1998) (dismissing the rule of lenity argument in the opinion's final paragraph).

96 428 F.3d 132 (2d Cir. 2005).

97 Sash v. Zenk, 439 F.3d 61 (2d Cir. 2006) (rehearing).

98 *Zenk*, 428 F.3d at 134.

99 *Id*.

100 *Zenk*, 439 F.3d at 63.

101 *Id*.

102 *Id*. at 64.

103 *Id*.

104 *Id*. at 65.

105 *Id*. at 66.

106 966 F.Supp. 186 (S.D.N.Y. 1997).

107 *Westcott*, 966 F.Supp. at 188.

108 *Id*. at 187.

109 *Id*. at 188.

110 *Id*.

111 *Id*. at 189 (citing and discussing United States v. Garcia-Rico, 46 F.3d 8 (5th Cir. 1995) (applying amended definition of aggravated felony retroactively); United States v. Campbell, 94 F.3d 125 (4th Cir. 1996) (applying amended definition of aggravated felony retroactively); and United States v. Gomez-Rodriguez, 96 F.3d 1262 (9th Cir. 1996) (rejecting retroactive application of the 1990 amendments)).

112 *Westcott*, 966 F.Supp. at 189, 190.

113 *Id*. at 190.

114 *Id*. at 191.

115 *Id*.

116 *Id*.

117 *Id*. (quoting United States v. Canales, 91 F.3d 363, 367 (2d Cir. 1996)).

118 *Id*.

119 United States v. Westcott, 159 F.3d 107 (2d Cir. 1998).

120 *See* Skilling v. United States (08-1394) (discussing the scope of federal law punishing a corporate executive's failure to provide "honest services"); Weyhrauch v. United States (08-1196) (discussing the rule of lenity in the context of the "intangible right to honest services"); Black v. United States (08-876) (discussing the rule of lenity in the context of the "honest services" provisions of 18 U.S.C. § 1346).

Environmental Law & Property Rights

Ripening Federal Property Rights Claims

*By J. David Breemer**

The Fifth Amendment to the United States Constitution provides that "private property [shall not] be taken for public use, without just compensation." This guarantee promises compensation for property owners subject to laws and regulations that invade, or excessively limit the use of, property.[1] However, securing that promise in the courts is procedurally difficult. Most importantly, before a court will even consider whether a local or state government has caused a compensable taking of private property, the property owner must demonstrate that his takings claim is "ripe" for judicial review.[2]

In general, the ripeness doctrine is designed to prevent courts from entangling themselves in abstract disputes, that is, disputes where there is as yet no concrete conflict of rights or injuries. The goals of the ripeness doctrine are to ensure that judicial power complies with the constitutional demand that such power address an actual "case and controversy," and to ensure that judicial resources are not needlessly or prematurely expended on hypothetical grievances.[3] The Supreme Court has established specific ripeness rules for federal takings claims. These rules place strong, though not insurmountable, barriers between property owners and their constitutional right to secure just compensation for a taking of their property. In some jurisdictions, the ripeness rules for takings claims may apply to other property rights claims. Therefore, to understand modern federal takings law, and the state of property rights in general, one must understand takings ripeness doctrine. This requires familiarity with the Court's 1985 decision in *Williamson County Regional Planning Commission v. Hamilton Bank*,[4] as it applies to property rights claims filed in state and federal courts.

I. THE WILLIAMSON COUNTY FRAMEWORK

A. Overview

The Supreme Court granted certiorari in *Williamson County* to consider "whether federal, state, and local governments must pay money damages to a landowner whose property allegedly has been 'taken' temporarily by the application of government regulations."[5] However, the Court never reached this issue. Instead, it focused on the ripeness of the property owner's claims. The *Williamson County* Court initially ruled that a federal takings claim is unripe until the "government entity charged with implementing the regulations has reached *a final decision regarding the application of the regulations to the property at issue*."[6] The Court ruled that there was no final decision, and thus no ripe takings claim, because the property owner had failed to apply for available regulatory variances that might have loosened some of the challenged land use restrictions.[7] The Court applied the same analysis to the owner's due process claim.[8]

Although the *Williamson County* Court's final-decision ripeness analysis effectively decided the case against the property owner, it did not stop there. Instead, it articulated and applied a second, more novel ripeness requirement. Specifically, the Court ruled that a federal takings claim is "complete" and therefore ripe after a property owner has first sought, and has been denied, just compensation through a state's available and "adequate procedures."[9] The Court did not exhaustively define the types of "state procedure" that might ripen a federal takings claim, but it ultimately held that the *Williamson County* property owner's takings claim was premature because it had failed to seek compensation through Tennessee's inverse condemnation procedure before raising its takings claim in federal court.[10] The Court did not apply the state procedures requirement to a due process claim that sought invalidation of the offending regulation and damages, rather than "just compensation."[11]

Williamson County thus established two hurdles for a property owner wishing to obtain just compensation for an alleged invasion of property under the Takings Clause of the Fifth Amendment. First, the owner must demonstrate that the alleged invasion arose from a final agency decision (the final decision requirement).[12] Second, the owner must show that he has sought, but been denied, just compensation in a state court procedure (the state procedures requirement).[13]

B. The State Procedures Prong Applies Only to Ripeness in Federal Courts

Williamson County did not distinguish between state and federal courts in articulating its two ripeness requirements, but the decision left some question as to whether the state procedures prong applied equally in state and federal courts. Some of the reasoning behind the state procedures rule suggests that completed state litigation is a ripeness predicate in any judicial forum.[14] Yet the claim in *Williamson County* arose from and was held unripe in federal court,[15] not a state court, and precedent cited by the *Williamson County* Court for the state procedures rule is directed toward federal court ripeness.[16]

In the aftermath of *Williamson County*, state courts diverged on the application of the state procedures rule. Some held that it required a property owner to complete state law takings litigation before raising a federal takings claim in the state court.[17] In *San Remo Hotel v. City and County of San Francisco*, the Court resolved this conflict. *San Remo* held that *Williamson County* does not require a federal takings claimant to fully and unsuccessfully litigate a state law damages action prior to filing a federal claim in *state* court.[18] A property owner may file a federal takings claim in state court simultaneously with any other claims.[19] This means that *Williamson County*'s "state procedures" rule applies only in federal courts, while the final decision requirement applies in both state and federal forums.[20]

* *Principal Attorney, Pacific Legal Foundation.*

II. FINAL DECISION RIPENESS

According to *Williamson County*, a final decision exists when the government reaches "a final, definitive position regarding how it will apply the regulations at issue to the particular land in question."[21] The *Williamson County* decision further noted that, sometimes, a property owner will have to unsuccessfully seek a variance from a challenged property restriction to obtain a final decision.[22]

Since *Williamson County*, lower courts have had ample opportunity to construe and apply the final decision requirement to both takings and due process claims. Unfortunately, many courts have failed to understand the limits of the requirement. For instance, in the takings context, some courts have misconstrued finality ripeness to require a decision that denies a property owner all or substantially all economically beneficial use of property.[23] This is wrong. Construing finality to demand a particular level of impact on property use effectively converts a ripeness standard into a test for whether a property owner has a valid claim on the merits.[24] Final decision ripeness is not concerned with whether a property owner has a winning claim; it is simply concerned with ensuring that a land use decision is concrete enough to allow a court to even consider whether it rises to the level of a taking.[25] In short, a final decision for takings or due process ripeness exists whenever any challenged land use restriction—whatever its impact—is concretely applied to the subject property.[26]

The variance principle is also misunderstood. Many courts assume that this principle requires pursuit of available administrative appeals, like an application to a Zoning Board of Appeals,[27] but it does not. Unlike a valid variance mechanism, which potentially limits the reach of an initial land use regulation or decision, an administrative appeal remedies such a decision. *Williamson County* expressly *rejected* exhaustion of administrative or state law judicial remedies as a final decision predicate.[28]

Even when there is a potential variance procedure, a landowner is not always required to pursue it. This is because finality ripeness ultimately derives from the exhaustion of agency discretion over application of a particular land use regulation,[29] not on whether the landowner has submitted multiple applications.[30] The variance requirement does not apply if the agency has no real discretion to grant one, or the agency's decision as to restricting a particular land use is already reasonably certain.[31] A landowner also need not pursue a variance if it is part of a futile or unfair process.[32]

When constrained by its purpose and *Williamson County*'s guidelines, the final decision ripeness requirement makes some sense. It essentially mirrors and implements the general and established ripeness requirement that an issue be "fit for review" before it can be adjudicated.[33] But when *Williamson County*'s self-imposed limits on final decision ripeness are neglected, the doctrine promotes bureaucratic obfuscation, delay, and indecision, and judicial timidity in declaring property rights violations. This state of affairs can eviscerate property rights as completely as the most onerous overt restriction.[34]

III. THE STATE PROCEDURES FEDERAL COURT "RIPENESS" REQUIREMENT

A. The State Procedures Requirement Ripens Nothing

A property owner wishing to pursue his federal right to just compensation in federal court, rather than in state court, must contend with *Williamson County*'s state procedures ripeness predicate, as well as the final decision rule. The rationale behind requiring state litigation for ripeness has always been suspect,[35] but, as originally articulated, the rule simply seemed to establish a temporary hurdle for federal takings claims.[36] As described in *Williamson County*, the state procedures ripeness principle anticipates that federal courts will hear a federal takings claim after the would-be claimant is denied monetary relief in state court.[37] It is at this point that the alleged federal taking can be deemed to be "without just compensation," and therefore federally actionable.

Unfortunately, the state procedures doctrine has never functioned as advertised because the requirement clashes with res judicata doctrines that bar plaintiffs from splitting their claims between separate lawsuits.[38] More to the point, the principles of claim and issue preclusion hold that a plaintiff may not go to federal court with a claim or issue that was or could have been litigated on the merits in a previous lawsuit. Given these rules, the very act of ripening a federal takings claim for federal court through a state court suit actually extinguishes any possibility for subsequent federal review.[39]

Because of its unstated effect in extinguishing federal review at the moment of ripeness, the state procedures rule has been accurately described a "trap" for unwary property owners. Many lower courts have expressed discomfort with the interaction of the state procedures rule and res judicata, but the Supreme Court approved of it in *San Remo*.[40] There, the Court confirmed that federal takings claims "ripened" for federal review are paradoxically barred from that forum.[41] Notably, four concurring justices in *San Remo* expressed a desire to overrule the state litigation aspect of *Williamson County* ripeness in an appropriate future case.[42] But despite being presented with repeated petitions on the issue since *San Remo*, the Court has strangely declined to follow up.

Until the Supreme Court overturns the state litigation rule, the bottom line is that most as-applied federal takings claims must be litigated in state court, or not at all. Notably, facial takings claims may not be subject to the same limits.[43] Many circuits allow facial takings claims to be raised in federal court without compliance with *Williamson County*'s state compensation litigation requirement.[44]

B. The State Procedures Requirement Should Not Bar Other Property Rights Claims from Federal Court

While *Williamson County* articulated the state procedures requirement exclusively for takings claims, some courts have held that it also applies to due process and equal protection-based property rights claims.[45] There is, however, no doctrinal basis for this extension of *Williamson County*. In *Williamson County*, the Supreme Court clearly derived the state litigation rule from the special nature of the "Just Compensation Clause" of the Fifth Amendment.[46] Property rights claims that do not

invoke the Just Compensation Clause logically do not trigger *Williamson County*'s demand for state court compensation litigation.[47] *Williamson County* itself confirmed these limits by declining to apply the state litigation rule to a due process claim in that case that sought invalidation and damages, not "just compensation."[48]

Nevertheless, some courts have indirectly subjected due process and/or equal protection claims to the state procedures barrier by subsuming those claims in the protections (and ripeness rules) of the Takings Clause.[49] But this reasoning is no longer viable after the Supreme Court's decision in *Chevron v. Lingle*.[50]

Lingle made clear that due process and equal protection violations are not addressed or remedied by the Takings Clause.[51] Since property rights claims arising under the Due Process and Equal Protection Clauses are not covered by the Takings Clause, there is no authority for treating them as unripe takings claims.[52] Beginning with the Ninth Circuit's 2007 decision in *Crown Point Development LLC v. City of Sun Valley*,[53] federal courts have explicitly held that, after *Lingle*, due process claims offer a distinct avenue for constitutional relief from onerous land use regulation.

A proper understanding of the injuries addressed by due process or equal protection land use claims confirms they cannot be analyzed as an adjunct of a takings claim subject to *Williamson County*. The core injury in a takings claim is the government's refusal to provide just compensation.[54] But due process and equal protection injuries have nothing to do with a failure to provide just compensation;[55] they hinge on irrational treatment and failure of notice and/or hearing.[56] Property rights claims asserting non-takings injuries are subject to independent treatment. Under traditional due process and equal protection law, there is no place for state court litigation as a ripeness predicate.[57] The state litigation ripeness rule is accordingly limited to as-applied federal takings claims.

CONCLUSION

Williamson County and progeny establish strong ripeness barriers to judicial review of property rights claims. However, when correctly applied, these barriers are not insurmountable. The final decision ripeness requirement is not markedly different than traditional ripeness doctrine; if there is a concrete land use decision and injury to property, finality exists, and a claim is fully justiciable in state court. In federal courts, however, the claimant must generally do more. With some exceptions, such a litigant must also satisfy Williamson County's state compensation procedures ripeness prong. While this ripeness rule is indeed a powerful bar to federal review of takings claims, it should not apply to or hinder other property rights claims, such as those arising under due process or equal protection guarantees.

Endnotes

1 *See* Lingle v. Chevron U.S.A. Inc., 544 U.S. 528, 538 (U.S. 2005).

2 *See* Williamson County Regional Planning Comm'n v. Hamilton Bank of Johnson City, 473 U.S. 172, 186, 192-94 (1985); Palazzolo v. Rhode Island, 533 U.S. 606, 626 (2001).

3 County Concrete Corp. v. Town of Roxbury, 442 F.3d 159, 164 (3rd Cir. 2006) ("The ripeness doctrine serves 'to determine whether a party has brought an action prematurely and counsels abstention until such time as a dispute is sufficiently concrete to satisfy the constitutional and prudential requirements of the doctrine.'") (citations omitted).

4 473 U.S. 173 (1985).

5 Williamson County, 473 U.S. at 185.

6 *Id*. at 186.

7 *Id*. at 186-188.

8 *Id*. at 200.

9 *Id*. at 194.

10 *Id*. at 196.

11 *Id*. at 198-200.

12 *Id*. at 186-88, 200.

13 *Id*. at 194.

14 473 U.S. at 195 n.13 ("[B]ecause the Fifth Amendment proscribes takings without just compensation, no constitutional violation occurs until just compensation has been denied. The nature of the constitutional right therefore requires that a property owner utilize procedures for obtaining compensation before bringing a § 1983 action.").

15 *Id*. at 175.

16 *See id*. at 195 (citing *Parratt v. Taylor*, 451 U.S. 527 (1981)).

17 *See* San Remo Hotel v. City and County of San Francisco, 545 U.S. 323, 351 n.2 (Rehnquist, C.J., concurring) (citing cases).

18 *Id*. at 346 (*Williamson County* "does not preclude state courts from hearing simultaneously a plaintiff's request for compensation under state law and the claim that, in the alternative, the denial of compensation would violate the Fifth Amendment of the Federal Constitution.").

19 *Id.*

20 *See, e.g.,* Kitchen v. City of Newport News, 657 S.E.2d 132, 139 (Va. 2008) ("Based on *San Remo Hotel*, we are persuaded that contrary to the City's contentions, Kitchen was not required to seek a remedy under state law [] without success before the circuit court could consider his Fifth Amendment claim").

21 473 U.S. at 191.

22 *Id*. at 190-92.

23 *See, e.g.,* Adrian v. Town of Yorktown, 2007 WL 1467417, *8 (S.D.N.Y. 2007) ("In order for a decision to be deemed final, the decision must deny Plaintiffs 'all reasonable beneficial use of [their] property.'").

24 *Compare Lingle*, 544 U.S. at 538 (A regulatory taking occurs when regulations "deprive an owner of 'all economically beneficial us[e]' of her property.") (quoting Lucas v. South Carolina Coastal Council, 505 U.S. 1003, 1019 (1992)). Thus, if courts construe ripeness to require a denial of all economic use, they wrongly demand a meritorious claim.

25 *Williamson County*, 473 U.S. at 191 (ripeness requires "a final, definitive position regarding how it will apply the regulations at issue to the particular land in question"); *County Concrete Corp.*, 442 F.3d at 164 (once a "decision maker has arrived at a definitive position on the issue" the property owner has been inflicted with "an actual, concrete injury") (quoting *Williamson*, 473 U.S. at 192)).

26 *Williamson*, 473 U.S. at 191.

27 *See, e.g.,* Tobin v. Centre Tp., 954 A.2d 741, 748 (Pa. Cmwlth. 2008) (takings claim "not ripe because they had not exhausted the statutory remedy" of invalidating the subject ordinance).

28 *Williamson County*, 473 U.S. at 193; *see also* DLX, Inc. v. Kentucky, 381 F.3d 511, 517 (6th Cir. 2004) ("administrative exhaustion is explicitly not a component of a federal takings claim"); Maguire Oil Co. v. City of Houston, 243 S.W.3d 714, 720 (Tex.App.-Hous. (14 Dist.), 2007) ("[A]n

[administrative] appeal... would not be determinative of ripeness.").

29 Suitum v. Tahoe Regional Planning Agency, 520 U.S. 725, 738 (1997).

30 *Palazzolo*, 533 U.S. at 620-22 (once agency discretion is utilized, a claim is ripe, as "[r]ipeness doctrine does not require a landowner to submit applications for their own sake.... [but a landowner] is required to explore development opportunities... only if there is uncertainty as to the land's permitted use.").

31 *Id.*

32 *Palazzolo*, 533 U.S. at 621 ("Government authorities, of course, may not burden property by imposition of repetitive or unfair land-use procedures in order to avoid a final decision.") (citing Monterey v. Del Monte Dunes at Monterey, Ltd., 526 U.S. 687, 698 (1999)).

33 *See Williamson County*, 473 U.S. at 191 (explaining that "until the administrative agency has arrived at a final, definitive position regarding how it will apply the regulations at issue to the particular land in question," the Court cannot evaluate the takings issues of the "economic impact of the challenged action and the extent to which it interferes with reasonable investment-backed expectations") (citations omitted).

34 *See, e.g.,* Hallco Texas, Inc. v. McMullen County, 221 S.W.3d 50, 63 (Tex. 2006) (Hecht, J., dissenting) ("This case illustrates how the government can use this ripeness requirement to whipsaw a landowner. The government can argue either that there was no request for a variance when there should have been, or that the request was not specific enough, or that it was not reasonable enough, or that there was insufficient time to consider it—and therefore the landowner's regulatory-takings claim is premature, unripe, and should be dismissed. Or else it can argue that a request for a variance would be a waste of time, or that none was authorized, or that the landowner should have known his ridiculous proposal would never be seriously considered—and therefore his claim is late, barred, and should be dismissed. One way or the other, the result is the same. Ripening a regulatory-takings claim thus becomes a costly game of "Mother, May I," in which the landowner is allowed to take only small steps forwards and backwards until exhausted."). *See generally* Donald J. Kochan, *Ripe Standing Vines and the Jurisprudential Tasting of Matured Legal Wines: Property and Public Choice in the Permitting Process* (June 10, 2009), BYU Journal of Public Law, Vol. 24, No. 1, 2009 (forthcoming).

35 *See San Remo Hotel*, 545 U.S. 348-49 (Rehnquist, C.J., concurring); J. David Breemer, *You Can Check Out But You Can Never Leave: The Story of* San Remo Hotel—*The Supreme Court Relegates Federal Takings Claims to State Courts Under a Rule Intended to Ripen the Claims for Federal Review*, 33 B.C. Envtl. Aff. L. Rev. 247, 291-99, 305 (2006).

36 *DLX, Inc.*, 381 F.3d at 518 (*Williamson County* "clearly contemplates that after a state just-compensation proceeding, a federal-court action will be filed.").

37 *Id.*

38 *See San Remo*, 545 U.S. at 351 (Rehnquist, C.J., concurring) ("[O]ur holding today [enforcing res judicata over ripe claims] ensures that litigants who go to state court to seek compensation will likely be unable later to assert their federal takings claims in federal court"); *DLX, Inc.*, 381 F.3d at 519-20 ("The availability of federal courts to hear federal constitutional takings claims has often seemed illusory, because under *Williamson County* takings plaintiffs must first file in state court, as DLX did, before filing a federal claim, and because in deciding that federal claim, preclusive effect must be given to that prior state-court action under 28 U.S.C. & sect; 1738 according to the res judicata law of the state.").

39 *Id.*; *see also* Rockstead v. City of Crystal Lake, 486 F.3d 963, 968 (7th Cir. 2007) ("Although the Williamson line of cases that requires the property owner to seek compensation in the state courts speaks in terms of 'exhaustion' of remedies, that is a misnomer. For if... the property owner goes through the entire state proceeding [on the merits], and he loses, he cannot maintain a federal suit.").

40 545 U.S. at 338.

41 *Id*; *see also id*. at 351 (Rehnquist, C.J., concurring).

42 *Id.* at 352 (Rehnquist, C.J., concurring).

43 *Suitum*, 520 U.S. at 736 ("[F]acial challenges to regulation are generally ripe the moment the challenged regulation is passed").

44 *See* Asociacion De Subscripcion Conjunta Del Seguro De Responsabilidad Obligatorio v. Galarza, 484 F.3d 1, 14 (1st Cir. 2007) (as "a facial statutory challenge, its takings claim need not be brought first to a Commonwealth body, either administrative or judicial"); Quicken Loans, Inc. v. Wood, 449 F.3d 944, 953 (9th Cir. 2006) (same).

45 *See* River Park, Inc. v. City of Highland Park, F.3d 164, 167 (7th Cir. 1994); Bateman v. City of West Bountiful, 89 F.3d 704, 709 (10th Cir. 1996); Bigelow v. Michigan Dept. of Natural Resources, 970 F.2d 154 (6th Cir. 1992).

46 473 U.S. at 194 n.13 ("[B]ecause the Fifth Amendment proscribes takings without just compensation, no constitutional violation occurs until just compensation has been denied. The nature of the constitutional right therefore requires that a property owner utilize procedures for obtaining compensation before bringing a § 1983 action.").

47 *Murphy*, 402 F.3d 342, 349-50 (2005) (holding that "because we are not confronted with such a [takings] claim, this [state compensation] aspect of *Williamson County* is not implicated" with respect to ripening a due process claim); Herrington v. Sonoma County, 834 F.2d 1488, 1499 n.10 (9th Cir.1987) (the requirement to seek "'just compensation' from state entities.... does not apply to the due process and equal protection claims at issue here"); McKenzie v. City of White Hall, 112 F.3d 313, 317 (1997) (same).

48 473 U.S. at 200.

49 *Bateman*, 89 F.3d at 709; Miller v. Campbell County, 945 F.2d 348, 352-353 (10th Cir. 1991); Buckles v. Columbus Municipal Airport Authority, 90 Fed. Appx. 927, 931 (6th Cir. 2004); Montgomery v. Carter County, Tennessee, 226 F.3d 758, 769 (6th Cir. 2000).

50 544 U.S. 528 (2005).

51 *Id*. at 542 ("[A] regulation that fails to serve any legitimate governmental objective may be so arbitrary or irrational that it runs afoul of the Due Process Clause.... But such a test is not a valid method of discerning whether private property has been 'taken' for purposes of the Fifth Amendment").

52 *See Lanier*, 520 U.S. at 272 n.7 (explaining that a substantive due process claim is barred only when the behavior challenged by the claim is "covered by a specific constitutional provision").

53 506 F.3d 851, 854 (9th Cir. 2007); *see also* A Helping Hand, L.L.C. v. Baltimore County, 515 F.3d 356, 369 (4th Cir. 2008) (following *Crown Point*); Action Apartment Ass'n, Inc. v. Santa Monica Rent Control Bd., 509 F.3d 1020, 1024 (9th Cir. 2007) (same).

54 *Williamson County*, 473 U.S. at 194-95.

55 *Lingle*, 544 U.S. at 543 (if a land use decision "is so arbitrary as to violate due process – that is the end of the inquiry. No amount of compensation can authorize such action.").

56 Samaad v. City of Dallas, 940 F.2d 925, 932 (5th Cir. 1991) (equal protection injury is distinct from a takings injury); *Warren*, 411 F.3d 697, 708 (2005) ("[Plaintiffs'] procedural due process claim is not ancillary to their takings claim, but addresses a separate injury-the deprivation of a property interest without a predeprivation hearing.").

57 Zinermon v. Burch, 494 U.S. 113, 125 (1990) (holding that neither substantive due process claim nor procedural due process claim against non-random deprivations of property are subject to a state court exhaustion rule); Village of Willowbrook v. Olech, 528 U.S. 562, 564 (2000) (equal protection land use claim adjudicated on merits free of any ripeness or exhaustion hurdle).

The Supreme Court and the Judicial Takings Doctrine

By Steven Geoffrey Gieseler and Nicholas M. Gieseler**

This Fall, the United States Supreme Court will hear arguments in *Stop the Beach Renourishment v. Florida Department of Environmental Protection*.[1] Those with an interest in property rights jurisprudence will pay particular attention, as the case will be the Supreme Court's first dealing with the Takings Clause since the Court decided three such cases in 2005.[2] But *Stop the Beach Renourishment* is even more notable, perhaps, in that it holds the potential for the Supreme Court to furnish, for the first time, a majority opinion dealing with the doctrine known as "judicial takings."

As its name indicates, the judicial takings doctrine posits that the judiciary, just like the executive and legislative branches, may be held responsible for violations of constitutionally-protected property rights. Specifically, *Stop the Beach Renourishment* presents the question of whether a state court decision can so radically depart from settled background principles of state property law that it constitutes a taking, or violates due process guarantees, under the Federal Constitution.[3] Despite occasional language on the matter in concurrences and dissents, no majority opinion of the Supreme Court has formally addressed this judicial takings question. The Court has sanctioned, and applied, a similar approach in a variety of other contexts, holding that state court decisions, just like actions of the other two branches, can violate persons' constitutional rights. With the grant of certiorari in *Stop the Beach Renourishment*, the judicial takings doctrine is primed, finally, to have its day in court.

Stop the Beach Renourishment involves a Florida statute aimed at stemming beach erosion.[4] The Florida Legislature enacted the Beach and Shore Preservation Act to restore beaches eroded by a series of hurricanes that hit the state.[5] The Act carried out its objective by authorizing government officials to place large quantities of sand on the eroded beach, both landward and seaward of the line dividing privately-owned property from publicly-owned beach. Problems arose not just because the law authorized the government to physically occupy, via the sand placement, private property, but also because the Act granted the title to the new strip of beach to the State of Florida.[6]

This provision has profound implications. For over a century, under Florida common law, owners of littoral property—land lying directly adjacent to the water—have possessed rights unique to the littoral context. For instance, these owners maintained the exclusive right to directly access the water from their property, the right to new land formed by accretion, and the right to an unobstructed view of the water.[7] But because the Florida renourishment law severs direct contact with the water, these common law rights vanish.

** The Gieseler Brothers are attorneys with the Pacific Legal Foundation. PLF filed amicus briefs in support of the property owners in Stop the Beach Renourishment in the Florida Supreme Court and the U.S. Supreme Court.*

Recognizing the effects this scheme would have on their property values, as well as the constitutional issues involved, affected landowners brought suit challenging the validity of the Act as applied to their property. After a series of administrative actions, a Florida appeals court held that the Act resulted in an uncompensated taking of property, in the form of the vanquished littoral rights.[8] But the Florida Supreme Court reversed this decision, writing that because "the exact nature of [littoral] rights has rarely been described in detail," it was within the court's ambit to redefine them.[9] Some might argue that the court did so by overlooking over one hundred years of Florida common law to declare that "there is no independent right of contact with the water."[10]

This tactic elicited a heated dissent from Justice Fred Lewis, who described the majority opinion as having "butchered" Florida law to create a "dangerous precedent… based upon infirm, tortured logic and a rescission from existing precedent."[11] The crux of Judge Lewis's dissent was that the court's holding "simply erased" settled fundamental principles of Florida property law, and with them the entirety of the beachfront owners' littoral rights.[12] Though not explicitly identified as such, Judge Lewis's dissent reflects the very foundations of the judicial takings doctrine.

I. Judicial Takings in the Supreme Court

A. Of Concurrence and Dissent

The Fourteenth Amendment to the United States Constitution, incorporating the Fifth Amendment's protections, forbids states from taking private property for public use without just compensation.[13] The most obvious of these prohibited takings occurs when a government entity physically confiscates, occupies, invades, or takes title to private property.[14] So too must government pay just compensation when it regulates property to the extent that it is taken for constitutional purposes.[15]

While these takings of private property typically arise from legislative or administrative acts, the question remains whether actions of state courts can give rise to similar government liability. This question of judicial takings—so named though such scenarios implicate both the Takings Clause and guarantees of due process—has been asked of the Supreme Court long before *Stop the Beach Renourishment*. Fifteen years ago, in *Stevens v. City of Cannon Beach*, the Court denied a petition for writ of certiorari filed by the owners of beachfront property in Oregon.[16] The petitioners in *Stevens* alleged that the Oregon Supreme Court's application of the doctrine of customary use effected a taking of their private property, without just compensation, in violation of the Fifth and Fourteenth Amendments.

In dissenting from the Supreme Court's denial, Justice Scalia, joined by Justice O'Connor, invoked the Court's opinion in *Lucas v. South Carolina Coastal Council* for the proposition that certain principles inherent in the right to security in private property are so fundamental as to require payment when they are abrogated by state action.[17] In Justice Scalia's

reading, this holds true whether the state actor applying such restrictions is the executive, the legislature, or the judiciary: "No more by judicial decree than by legislative fiat may a State transform private property into public property without just compensation."[18]

Justice Scalia's dissent recognized the general rule that "the Constitution leaves the law of real property to the States."[19] However, "just as a State may not deny rights under the Federal Constitution through pretextual procedural rulings, neither may it do so by invoking nonexistent rules of state substantive law."[20] Justice Scalia concluded that he would grant the petition to determine whether the lower court's ruling violated the property owners' due process rights.[21] He also wrote that he would apply this theory of constitutional protection to takings claims in general.[22]

There are other Supreme Court decisions that mirror Justice Scalia's view of the validity of the judicial takings doctrine. The clearest and most influential opinion of the kind is Justice Stewart's concurrence in *Hughes v. Washington*.[23] In *Hughes*, an owner of upland property sought a determination of the ownership of accretions that had gradually formed along her beachfront property.[24] The land was conveyed to the landowner prior to the formation of what became the State of Washington.[25] At the time of the conveyance, the common law rule was that an owner of property bordering the ocean had the right to include within his title any accretion gradually built up by the movement of the tides.[26]

The Supreme Court considered the issue of who owned the accreted land—the state or the upland private owner—and held that the upland owner was to remain the sole owner of the property.[27] In his concurrence, Justice Stewart emphasized that property owners have valid claims under the Takings Clause where state courts suddenly depart from settled property law to the detriment of private owners:

> To the extent that the decision of the Supreme Court of Washington on that issue arguably conforms to reasonable expectations, we must of course accept it as conclusive. But to the extent that it constitutes a sudden change in state law, unpredictable in terms of the relevant precedents, no such deference would be appropriate. For a State cannot be permitted to defeat the constitutional prohibition against taking property without due process of law by the simple device of asserting retroactively that the property it has taken never existed at all.[28]

Justice Lewis, dissenting from the Florida Supreme Court's opinion in *Stop the Beach Renourishment*, shared many of the concerns identified by Justice Stewart in his concurrence in *Hughes*. Justice Lewis wrote that the majority's decision summarily altered the definition of littoral property that had governed in Florida for nearly a century: "In this State, the legal essence of littoral or riparian land is contact with the water. Thus, the majority is entirely incorrect when it states that such contact has no protection under Florida law and is merely some 'ancillary' concept that is subsumed by the right of access."[29] Justice Lewis recognized that it was the Florida Supreme Court's novel interpretation of the state statute in question, and not necessarily the statute itself, that violated the well-established constitutional rights of the beachfront property owners. Citing these owners' fundamental right to have their property maintain contact with the water, Justice Lewis wrote that "[t]he majority now avoids this inconvenient principle of law—and firmly recognized and protected property right[s]" by ignoring decades of settled state law on the matter.[30]

B. State Courts and Federal Rights

Justice Stewart's concurrence in *Hughes* finds analogues in other Supreme Court decisions holding that sudden judicial departures from settled state law violate citizens' rights as guaranteed by the Federal Constitution. For example, in *Webb's Fabulous Pharmacies v. Beckwith*,[31] the Court considered a Florida Supreme Court decision upholding as constitutional a state statute permitting counties to seize the interest accruing on an interpleader fund paid into by private citizens and maintained by county courts.[32] As with the dissent in the *Stop the Beach Renourishment* case, the Court's analysis focused not as much on the relevant Florida statute as on the Florida Supreme Court's opinion interpreting that statute. The U.S. Supreme Court found that the Florida court's holding was unconstitutional, and that "[n]either the Florida Legislature by statute, nor the Florida courts by judicial decree, may accomplish the result the county seeks simply by recharacterizing the principal as 'public money' because it is held temporarily by the court."[33] The Court concluded with a statement precisely on point for the property owners in *Stop the Beach Renourishment*: "a State, by *ipse dixit*, may not transform private property into public property without compensation...."[34]

Similarly, in *Bouie v. City of Columbia*, the Supreme Court was faced with a constitutional challenge to a state court decision that departed significantly from established jurisprudence governing a basic right.[35] In *Bouie*, the South Carolina Supreme Court applied an entirely new construction of a criminal trespass statute in order to uphold the convictions of two alleged trespassers.[36] This interpretation was such a departure from settled state law that the Supreme Court held it amounted to the imposition of an ex post facto law in violation of the petitioners' due process rights.[37] In the *Bouie* Court's view, a state may not avoid constitutional restrictions on its power merely by delegating the restriction to the courts instead of having them instituted by the elected branches: "If a state legislature is barred by the Ex Post Facto Clause from passing such a law, it must follow that a State Supreme Court is barred by the Due Process Clause from achieving precisely the same result by judicial construction."[38]

These are but two examples of Supreme Court decisions that extend constitutional protections, and prohibitions, to judicial actions. There are perhaps scores of similar opinions. In numerous civil rights cases of the late 1950s and early 1960s, the Court found state court actions to violate rights guaranteed by the Federal Constitution. For example, in *NAACP v. Ala. ex rel. Flowers*, the Supreme Court found injunctions against NAACP operations, as issued by Alabama courts, to violate the Due Process Clause.[39] In *NAACP v. Ala. ex rel. Robinson*, the Court reached a similar decision, writing that "[i]t is not of moment that the State has here acted solely through its judicial branch, for whether legislative or judicial, it is still the application of

state power which we are asked to scrutinize."[40] The Supreme Court's First Amendment jurisprudence is particularly notable for holding state judiciaries accountable for constitutional violations, where state courts "assert[ ] retroactively" that private actors had no right to exercise their First Amendment liberties.[41] Thus, if sometime in 2010 the Supreme Court issues an opinion recognizing and validating the judicial takings doctrine in *Stop the Beach Renourishment*, the decision will not be an outlier, but rather one comporting with the body of the Court's jurisprudence.

II. The Theoretical Basis for the Judicial Takings Doctrine

In the first opinion to incorporate the federal Takings Clause against the states, the Supreme Court made explicit that incorporation applies to state courts as well as state legislatures and executives.[42] In *Chicago Burlington & Quincy Railroad*, a just compensation case, the Court held that a state court decision "whereby private property is taken for the state or under its direction for public use, without compensation made or secured to the owner, is, upon principle or authority, wanting in the due process of law required by the fourteenth amendment of the constitution of the United States...."[43] As Roderick Walston writes in the Utah Law Review, the *Chicago Burlington* opinion, aside from its more famous incorporation holding, also "held that the state must pay compensation even when the alleged 'taking' is the result of judicial rather than legislative action.... [T]he final judgment of a state court is the act of the state for due process and takings purposes."[44] In property rights cases, then, the idea that state courts can violate due process guarantees, and take private property, is as old as the incorporation doctrine itself.

With regard to sudden changes in substantive law, the Supreme Court captured the modern theoretical basis for the judicial takings doctrine in its opinion in *Lucas*. In that case, the Court recognized that certain basic principles of property ownership are so fundamental as to be beyond the reach of the state, unless the state is willing to pay the owner for his property.[45] The *Lucas* Court arrived at that holding in part by way of negative examples; that is, by pointing to certain uses of property that, historically, *never* were lawful (and thus, the regulation of which could not require just compensation), the Court distinguished those incidents of ownership that *always* were lawful.[46]

Despite this distinction, the Court's *Lucas* opinion does not establish which aspects of property fit into which category. According to W. David Sarratt, writing in the Virginia Law Review, "[t]his ambiguity has provided states with a loophole in the *Lucas* rule large enough to circumvent the rule entirely, provided that state courts are willing to be rather creative in defining background legal principles."[47] Sarratt continues, observing that even post-*Lucas*, "[s]tates may thus attempt to avoid compensation altogether by announcing that under their background principles of state law, the property owner never had the property right she claims has been taken. Of course, state courts can pull off this ploy better than state legislatures."[48]

It is this reality that makes the theory of judicial takings so crucial to protecting property rights. As Sarratt notes, the job of defining what constitutes a *Lucas* background principle, existing perhaps for centuries, is more appropriate for the judiciary. This is because "legislatures are presumed to act prospectively, saying what the law shall be, while courts are presumed to decide questions retrospectively, saying what the law is and has been."[49] But this dexterity, uncabined by federal review, is not without peril. Sarratt writes:

> [W]hen state courts are understood to wield the power not only to declare the law, but also to make it, the *Lucas* rule's background-principles exception invites state courts to reshuffle property rights in ways that state legislatures cannot, potentially allowing the state to avoid paying compensation for takings of property.[50]

Premising his conclusion in part on the Supreme Court's holding in *Erie R.R. Co. v. Tompkins* ("[W]hether the law of the State shall be declared by its Legislature in a statute or by its highest court in a decision is not a matter of federal concern."),[51] Sarratt views federal application of the judicial takings doctrine as necessary.[52] Since *Erie* stands for the proposition that state courts are permitted to "make real law on behalf of the state,"[53] a state court's departure from established law must be treated by federal courts "as wielding real lawmaking power–including the ability to take property."[54]

Justice Scalia recognized in his *Stevens* dissent the real potential for state courts to overstep their bounds in their roles as "definers" of background principles. "Our opinion in *Lucas*, for example, would be a nullity if anything that a state court chooses to denominate background law–regardless of whether it is really such–could eliminate property rights."[55] The only answer to this conundrum–state courts have the authority and particular ability to define background principles, but in doing so they can eviscerate private property with nothing more than the metaphorical stroke of a pen–is for federal courts, including the Supreme Court, to review these definitions by way of the judicial takings doctrine.

As Professor David J. Bederman has written, the judiciary's ability to wield its power to make law is pronounced in property rights cases involving beach property, where judge-made law of custom governs.[56] Citing "custom as an end-run around *Lucas*,"[57] Bederman writes of examples of state supreme courts having "obliterated constitutional requirement[s] (whether articulated in a takings or due process idiom)"[58] relating to property rights by invoking common law principles of custom that the courts themselves have developed.[59] The danger in beach cases like *Stevens* and *Stop the Beach Renourishment*, generally, is that in the absence of federal review, state courts are free to fashion whatever rules they choose without being cabined by constitutional boundaries.

Writing in the Virginia Law Review in 1990, Barton H. Thompson, Jr., authored what is widely recognized as the "seminal article on the judicial takings problem."[60] Thompson's article identifies the rationales federal courts might use to apply the judicial takings doctrine, focusing specifically on the importance of legal determinacy. He notes that sudden changes from established precedent often are a signal that state courts have abdicated their roles as the "generally less political" branch of government: "Justice Stewart's suggestion that judicial

changes in property law are takings only when 'sudden' and quite unpredictable may have been designed partially to ferret out the more questionable judicial changes."[61] Thompson observes that while slow, gradual changes in the common law assure property owners that legal considerations, not political ones, dictated a ruling, "[n]o such assurances accompany a sudden and quite startling judicial shift in property rights."[62]

Thompson's article is among the earliest, if not the first, to address scholarly opposition to the judicial takings doctrine. He writes that the "few scholars to have seriously addressed the issue have generally argued that it would be catastrophic to subject the courts to the same constitutional constraints as the legislative and executive branches, but with little illumination as to why."[63] Thompson preemptively rebuts any textualist objection to the effect that the Fifth or Fourteenth Amendments do not apply to state courts, noting that while nothing in the Constitution's language compels such application, the "broad language certainly does not preclude application to judicial changes in property rights."[64] As outlined in Part I above, this textual imprecision has done nothing to preclude the Supreme Court's application of due process guarantees to the state courts, a point Thompson expounds upon in his article.[65]

Thompson further cites opposition to the judicial takings doctrine that asserts the Just Compensation Clause necessarily excludes the judiciary, for the courts have no "fiscal purse" of their own.[66] He parries this argument by writing "it is worth noting that the executive branch also lacks a separate purse and yet there is no doubt that the fifth amendment applies to at least some executive takings."[67]

The main focus of Thompson's seminal article, though, is the "normative pulls and counterpulls that have shaped our takings jurisprudence."[68] To this end, he returns repeatedly to the dangers of allowing state courts, charged with defining themselves what is and isn't property under *Lucas*, also to serve as final arbiter of the validity of these definitions under the Federal Constitution. Thompson argued, two years before *Lucas*, that state courts were too eager, and too able, to take private property without repercussion. Thompson wrote that

> [c]ourts have the doctrinal tools to undertake many of the actions that legislatures and executive agencies are constitutionally barred from pursuing under the takings protections–and pressure is mounting for courts to use these tools. Indeed, while paying lip service to stare decisis, the courts on numerous occasions have reshaped property law in ways that sharply constrict previously recognized private interests. Faced by growing environmental, conservationist, and recreational demands, for example, state courts have recently begun redefining a variety of property interests to increase public or governmental rights, concomitantly shrinking the sphere of private dominion.[69]

In the nearly two decades since Thompson wrote these words, the pressures he identified have only increased. Perversely, the Supreme Court's opinion in *Lucas*, offering greater protections for owners of private property, perhaps has left state courts even more free to effect takings of private property, as they step in where legislatures and executives now are more afraid to tread. By agreeing to hear *Stop the Beach Renourishment*, the Supreme Court appears ready to define these boundaries for such state court action, and to finally and unambiguously answer the question of the validity of the judicial takings doctrine.

CONCLUSION

As Thompson wrote in 1990, "[t]he United States Supreme Court's reluctance to apply the takings protections to courts proved particularly puzzling [ ] when one compares the Court's treatment of other constitutional restrictions that, unlike the takings protections, are essentially noneconomic."[70] "Even where the Court has concluded that a specific noneconomic protection does not directly apply to the judiciary, the Court has sometimes extended the protection to judicial actions, using a more general constitutional provision."[71] In mere months, we should know, going forward, whether the Supreme Court will extend such constitutional protection to owners of property threatened by the actions of state courts.

Endnotes

1 Case No. 08-1151.

2 Lingle v. Chevron U.S.A. Inc., 544 U.S. 528 (2005); San Remo Hotel, L.P. v. City and County of San Francisco, Cal., 545 U.S. 323 (2005); Kelo v. City of New London, Conn., 545 U.S. 469 (2005).

3 The relevant Question Presented, one of three in the case as accepted by the Court, reads as follows: "The Florida Supreme Court invoked 'nonexistent rules of state substantive law' to reverse 100 years of uniform holdings that littoral rights are constitutionally protected. In doing so, did the Florida Court's decision cause a 'judicial taking' proscribed by the Fifth and Fourteenth Amendments to the United States Constitution?"

4 Walton County v. Stop the Beach Renourishment, Inc., 998 So. 2d 1102 (Fla. 2008).

5 Fla. Stat. ch. 161 (2003).

6 Fla. Stat. § 161.191(1).

7 Bd. Of Trustees v. Medeira Beach Nominee, Inc., 272 So. 2d 209, 214 (Fla. 2d DCA 1973) (beachfront owners "have the exclusive right of access over their own property to the water.").

8 Save Our Beaches, Inc. v. Florida Dep't of Envtl. Prot., 2006 WL 1112700 (July 6, 2006).

9 *Stop the Beach Renourishment*, 998 So. 2d at 1111.

10 *Id.* at 1119.

11 *Id.* at 1121 (Lewis, J., dissenting).

12 *Id.*

13 Chicago, B. & Q. R. Co. v. City of Chicago, 166 U.S. 226, 242-43 (1897).

14 *Kelo*, 545 U.S. 469; Loretto v. Teleprompter Manhattan CATV Corp., 458 U.S. 419 (1982).

15 *Lingle*, 544 U.S. 528, 538-541.

16 Stevens v. City of Cannon Beach , 510 U.S. 1207, 114 S. Ct. 1332 (1994).

17 *Stevens*, 114 S. Ct. at 1334 (Scalia, J., dissenting) (citing *Lucas*, 505 U.S. 1003 (1992)).

18 *Id.*

19 *Id.*

20 *Id.* (citation omitted)

21 *Id.* at 1335-36

22 *Id.*

23 389 U.S. 290, 296-97 (1967) (Stewart, J., concurring).

24 *Id.* at 291.

25 *Id.*

26 *Id.*

27 *Id.* at 294.

28 *Id.* at 296-97 (Stewart, J., concurring).

29 *Stop the Beach Renourishment*, 998 So. 2d at 1122.

30 *Id.* at 1123.

31 449 U.S. 155 (1980).

32 *Id.* at 155-56.

33 *Id.* at 164.

34 *Id.*

35 378 U.S. 347 (1964).

36 *Id.* at 362.

37 *Id.*

38 *Id.* at 353-54.

39 377 U.S. 288, 306-08 (1964).

40 357 U.S. 449, 463 (1958).

41 Org. for a Better Austin v. Keefe, 402 U.S. 415 (1971) (state court injunction forbidding distribution of pamphlet violated First Amendment); *see also* Carroll v. President & Comm'rs of Princess Anne, 393 U.S. 175 (1968) (state court injunction against rally violated right to free speech); N.Y. Times Co. v. Sullivan, 376 U.S. 254 (1964) (vacating state court restriction on publication).

42 David J. Bederman, *The Curious Resurrection of Custom: Beach Access and Judicial Takings*, 96 Colum. L. Rev. 1375, 1435 (1996); Roderick E. Walston*, The Constitution and Property: Due Process, Regulatory Takings, and Judicial Takings*, 2001 Utah L. Rev. 379, 425-26 (2001) (citing *Chicago B & Q R. Co.*, 166 U.S. 226).

43 166 U.S. at 241

44 Walston, *supra* note 42, at 425-26.

45 Lucas, 505 U.S. at 1029-30.

46 *Id.* at 1030-31.

47 W. David Sarratt, *Judicial Takings and the Course Pursued*, 90 Va. L. Rev. 1487, 1489 (2004).

48 *Id.* at 1490.

49 *Id.* at 1491.

50 *Id.*

51 304 U.S. 64, 78 (1938).

52 Sarratt, *supra* note 47, at 1496-97.

53 *Id.* at 1496.

54 *Id.* at 1497.

55 *Stevens*, 114 S. Ct. at 1334 (Scalia, J., and O'Connor, J., dissenting).

56 Bederman, *supra* note 42, at 1438-39.

57 *Id.* at 1442.

58 Whether judicial takings challenges are properly grounded in takings or due process law, or both, is unsettled even among those who are proponents of the general doctrine. Justice Scalia's *Stevens* dissent, for example, contemplates the doctrine applying to both takings and due process under some circumstances. 114 S. Ct. at 1335-36. Others, such as Roderick E. Walston, argue persuasively that the doctrine applies only to due process claims. Walston, *supra* note 42, at 1434-36.

59 Bederman, *supra* note 42, at 1442-43.

60 Sarratt, *supra* note 47, at 1494 (citing Barton H. Thompson, Jr., *Judicial Takings*, 76 Va. L. Rev. 1449 (1990).

61 Thompson, *supra* note 60 at 1496-97.

62 *Id.* at 1497.

63 *Id.* at 1453.

64 *Id.* at 1456.

65 *Id.* at 1456-57.

66 *Id.* at 1456.

67 *Id.* at 1456, n.22.

68 *Id.* at 1453.

69 *Id.* at 1451.

70 *Id.* at 1456.

71 *Id.* at 1457.

Financial Services & E-Commerce

The Stop Trading on Congressional Knowledge Act

*By Stephen M. Bainbridge**

Over time, nobody beats the market. This basic premise of efficient capital markets theory has been confirmed in numerous academic studies.[1] The only important exception to the rule traditionally has been corporate insiders trading in their own corporation's stock.[2] The obvious and generally accepted explanation for insiders' results is their access to and use of material nonpublic information about the company.[3]

A 2004 study of the results of stock trading by United States Senators during the 1990s, however, found that the Senators on average beat the market by 12% a year.[4] In sharp contrast, U.S. households on average underperformed the market by 1.4% a year and even corporate insiders on average beat the market by only about 6% a year during that period.[5] A reasonable inference is that some Senators had access to—and were using—material nonpublic information about the companies in whose stock they trade:

> Looking at the timing of cumulative returns, the senators also appeared to know exactly when to buy or sell their holdings. Senators would buy stocks just before the shares suddenly would outperform the market by more than 25%. Conversely, senators would sell stocks that had been beating the market by about 25% for the past year just when the shares would fall back in line with the market's performance.
>
> The researchers say senators' uncanny ability to know when to buy or sell their shares seems to stem from having access to information that other investors wouldn't have. "I don't think you need much of an imagination to realize that they're in the know," says Alan Ziobrowski, a business professor at Georgia State University in Atlanta and one of the four authors of the study.[6]

Members of Congress can obtain material nonpublic information in many ways. They can learn inside information when, for example, a company confidentially discloses it during the course of a Congressional hearing or investigation. In most cases, however, members of Congress likely trade on the basis of market information.

> "Market information" refers to information that affects the price of a company's securities without affecting the firm's earning power or assets.… Examples include information that an investment adviser will shortly issue a "buy" recommendation or that a large stockholder is seeking to unload his shares or that a tender offer will soon be made for the company's stock.[7]

In the present context, examples of market information readily available to members of Congress include knowing that "tax legislation is apt to pass and which companies might benefit," being aware "that a particular company soon will be awarded a government contract or that a certain drug might get regulatory approval…."[8]

Analysis of the legality of such trading must begin by recognizing that the term "insider trading" is a misnomer in two relevant senses. First, trading on the basis of either inside or market information is a potential breach of the federal securities laws. The mere fact that information may have originated outside the company is irrelevant, so long as the information is material to the company's stock price.[9] Second, the federal securities laws' prohibition of so-called "insider" trading in fact extends to many corporate outsiders.[10] Accordingly, Congressmen, their staffers, and other government officials and employees are not exempt from liability for trading on the basis of material nonpublic information simply because they are not corporate insiders.

Under current law, however, although congressional staffers and other government officials and employees could be prosecuted successfully for insider trading under the federal securities laws, the quirks of the applicable laws almost certainly would prevent members of Congress from being successfully prosecuted.[11] To address that anomaly, Congressmen Louise Slaughter (D-NY) and Brian Baird (D-WA) have introduced The Stop Trading on Congressional Knowledge Act ("STOCK Act").[12] If adopted, the Act "will prohibit Members of Congress and their staff from using nonpublic information they are able to obtain through their official positions to enrich their personal portfolios."[13]

Part I of this Article explains why members of Congress are effectively immune from insider trading liability under the current federal securities laws. Part II sets out the policy justifications for extending those laws to include members of Congress. Finally, Part III critiques the STOCK Act's approach to banning congressional insider trading.

I. Current Law

Although the modern insider trading prohibition technically is grounded in the federal securities statutes and regulations, most notably Rule 10b-5,[14] promulgated by the SEC pursuant to the authority granted it by Section 10(b) of the Securities Exchange Act,[15] the prohibition in fact is the product of a series of judicial decisions creating a quasi-common law in the interstices of Rule 10b-5.[16] As the prohibition evolved, two conceptually distinct theories emerged pursuant to which liability for insider trading can be imposed:

> [T]he Supreme Court has recognized two theories of insider trading liability: the "classical theory" and the "misappropriation theory." The classical theory generally only imposes liability when a trader or tipper is an insider of the traded-in corporation. The classical

* *William D. Warren Professor of Law, UCLA School of Law.*

> insider-trader thus breaches a fiduciary duty owed to the corporation's shareholders. The misappropriation theory, however, creates liability when a tipper or trader misappropriates confidential information from his source of the information. The misappropriator thus breaches a fiduciary duty owed to the source.[17]

As we shall see below, in the vast majority of cases only the latter theory will be relevant to insider trading by members of Congress and other governmental officials.

A. The Classical Theory

The modern federal insider prohibition began taking form in *S.E.C. v. Texas Gulf Sulphur Co.*[18] The *TGS* opinion rested on a policy of equality of access to information. The court contended that the federal insider trading prohibition was intended to assure that "all investors trading on impersonal exchanges have relatively equal access to material information." Put another way, the majority thought Congress intended "that all members of the investing public should be subject to identical market risks." Accordingly, under *TGS* and its progeny, virtually anyone who possessed material nonpublic information was required either to disclose it before trading or abstain from trading in the affected company's securities. If the would-be trader's fiduciary duties precluded him from disclosing the information prior to trading, abstention was the only option.

In *Chiarella v. U.S.*,[19] the United States Supreme Court rejected the equal access policy. Instead, the Court made clear that liability could be imposed only if the defendant was subject to a duty to disclose prior to trading. In turn, the requisite duty to disclose arises out of a fiduciary relationship between the inside trader and the persons with whom he trades. *Chiarella* thus made clear that the disclose or abstain rule is not triggered merely because the trader possesses material nonpublic information. When a securities fraud action is based upon nondisclosure, there can be no fraud absent a duty to speak, and no such duty arises from the mere possession of nonpublic information.[20] As the court explained in its subsequent decision in *Dirks v. S.E.C.*[21]:

> We were explicit in Chiarella in saying that there can be no duty to disclose where the person who has traded on inside information "was not [the corporation's] agent, ... was not a fiduciary, [or] was not a person in whom the sellers [of the securities] had placed their trust and confidence." Not to require such a fiduciary relationship, we recognized, would "[depart] radically from the established doctrine that duty arises from a specific relationship between two parties" and would amount to "recognizing a general duty between all participants in market transactions to forgo actions based on material, nonpublic information."[22]

The substantial limitation on the scope of insider trading liability imposed by *Chiarella* posed the question whether anyone other than classical insiders such as directors, officers, and perhaps large shareholders could be held liable for dealing on the basis of insider information. *Dirks* confirmed that the classical theory reached at least two other categories of potential defendants. First, certain outsiders with especially close relationships with the issuing corporation can become constructive insiders:

> Under certain circumstances, such as where corporate information is revealed legitimately to an underwriter, accountant, lawyer, or consultant working for the corporation, these outsiders may become fiduciaries of the shareholders. The basis for recognizing this fiduciary duty is not simply that such persons acquired nonpublic corporate information, but rather that they have entered into a special confidential relationship in the conduct of the business of the enterprise and are given access to information solely for corporate purposes.… For such a duty to be imposed, however, the corporation must expect the outsider to keep the disclosed nonpublic information confidential, and the relationship at least must imply such a duty.[23]

Second, the *Dirks* court held that the insider trading applies not only when an insider—whether classical or constructive—trades, but also when such an insider tips inside information to an outsider who then trades.[24] The court held that a tippee's liability is derivative of that of the tipper, "arising from [the tippee's] role as a participant after the fact in the insider's breach of a fiduciary duty."[25] A tippee therefore can be held liable only when the tipper breached a fiduciary duty by disclosing information to the tippee, and the tippee knows or has reason to know of the breach of duty.[26]

What *Dirks* proscribes thus is not merely a breach of confidentiality by the insider, but rather the breach of a fiduciary duty of loyalty to refrain from profiting on information entrusted to the tipper.[27] Looking at objective criteria, courts must determine whether the insider-tipper personally benefited, directly or indirectly, from his disclosure.[28] The most obvious example of a benefit is the outright sale of information for cash. Non-pecuniary gains by the insider can also qualify, however.[29] Liability could be imposed, for example, on a corporate CEO who discloses information to a wealthy investor not for any legitimate corporate purpose, but solely to enhance his own reputation.[30] Likewise, liability could be imposed where the tip is a gift because that transaction is regarded as analogous to the situation in which the tipper trades on the basis of the information and then gives the tippee the profits.[31]

Cases in which members of Congress or other government officials qualify as classical insiders or constructive insiders present no enforcement difficulties under current law. Nothing in the existing rules precludes their application in such cases. Such cases, however, presumably are quite rare. According to the House Ethics Manual, for example, members of Congress and their senior staff may not, inter alia, "serve for compensation as an officer or member of the board of an association, corporation, or other entity."[32] Opportunities to serve as a classical insider thus are unlikely to arise.

In contrast, it seems plausible that Congressmen or other government officials might sometimes receive tips from corporate insiders. Such a tip would be the functional equivalent of a bribe. Nothing in current law would prohibit prosecution of both tipper and tippee in such cases. Instead, it would be treated the same way as gifts of information.

Suppose the insider claimed that he gave the tip not for

personal benefit, however, but so that the company would benefit. In effect, the tipper claims, he bribed the Congressman so the Congressman would do a favor for the company. The logic of *Dirks* suggests there could be no insider trading liability in such a case.[33]

B. The Misappropriation Theory

As defined by the Supreme Court in *U.S. v. O'Hagan*,[34] in which the Court endorsed the misappropriation theory as a valid basis for insider trading liability, the misappropriation theory is another misnomer. It does not deal with theft of inside information—or, at least, not directly—but rather holds that a fiduciary's undisclosed use of information belonging to his principal, without disclosure of such use to the principal, for personal gain constitutes fraud in connection with the purchase or sale of a security and thus violates Rule 10b-5.[35]

The Court acknowledged that misappropriators have no disclosure obligation running to the persons with whom they trade.[36] Instead, it grounded liability under the misappropriation theory on deception of the source of the information; the theory addresses the use of "confidential information for securities trading purposes, in breach of a duty owed to the source of the information."[37] Under this theory, "a fiduciary's undisclosed, self serving use of a principal's information to purchase or sell securities, in breach of a duty of loyalty and confidentiality, defrauds the principal of the exclusive use of that information."[38] So defined, the Court held, the misappropriation theory satisfies § 10(b)'s requirement that there be a "deceptive device or contrivance" used "in connection with" a securities transaction.[39]

Where members of Congress, congressional staffers, or other government officials obtain material nonpublic information in the course of their duties and then use it to trade in the stock of the relevant issuer, their conduct could be colloquially described as a theft of the information, but any potential insider trading liability under the misappropriation theory would require proof of a fiduciary duty between the official and the source of the information. To be sure, two recent cases hold that a fiduciary relationship is not essential to misappropriation liability. In *S.E.C. v. Cuban*,[40] a district court held that a non-fiduciary who had agreed contractually both to keep information confidential and not to use the information for personal gain could be held liable for misappropriation-based insider trading liability. In *S.E.C. v. Dorozhko*,[41] the Second Circuit held that an alleged computer hacker who supposedly broke into the computer system of a company called IMS Health, Inc., and used the information he learned in doing so to purchase put options on the company's stock had committed a deceptive act in connection with the purchase or sale of a security. Because the case purportedly involved a material misrepresentation (namely, the hacker's disguising of his identity in breaching the company's computer network), the court opined that showing of a fiduciary duty is unnecessary.

The *Cuban* case seems more important for present purposes. If government ethics rules banning the use of nonpublic information for personal gain are deemed to constitute the requisite agreement, the *Cuban* case provides a precedent for imposing liability. The *Cuban* decision, however, is inconsistent with the well-accepted proposition that a fiduciary relationship is required.[42]

Assuming that the misappropriation theory in fact requires a breach of a duty of disclosure arising out of a fiduciary relationship or similar relationship of trust and confidence, an important distinction arises between members of Congress and other government officials. The Standards of Ethical Conduct For Employees of the Executive Branch provide that "[p]ublic service is a public trust, requiring employees to place loyalty to the Constitution, the laws and ethical principles above private gain."[43] Accordingly, an employee of the Executive Branch should be deemed an agent of the government or, at least, to stand in a similar relationship of trust and confidence with the government.[44] The Standards further provide: "An employee shall not engage in a financial transaction using nonpublic information, nor allow the improper use of nonpublic information to further his own private interest or that of another, whether through advice or recommendation, or by knowing unauthorized disclosure."[45]

Turning to Congress, both members of a Congressman's staff and committee staffers are employees of their respective houses.[46] They are subject to an ethical obligation never to "use any information received confidentially in the performance of governmental duties as a means for making private profit."[47]

These employment relationships should suffice for congressional staffers to be deemed to have an agency or other relationship of trust and confidence with their employing agency. In *S.E.C. v. Cherif*,[48] for example, the court held that "a person violates Rule 10b-5 and Section 10(b) by misappropriating and trading upon material information entrusted to him by virtue of a fiduciary relationship such as employment."[49] Put into *O'Hagan*'s terminology, "a [staffer's] undisclosed, self serving use of [congressional] information to purchase or sell securities, in breach of a duty of loyalty and confidentiality, defrauds [Congress]."[50]

Of whom are members of Congress agents or fiduciaries, however? With whom do they have the requisite relationship of trust and confidence out of which the requisite duty to disclose before trading arises? The only logical candidate is the electorate. Although there is some precedent in other contexts for the proposition that "a public official… owe[s] a fiduciary duty to the public to make governmental decisions in the public's best interest,"[51] the predominant view, as stated by former SEC enforcement official Thomas Newkirk, is that "[i]f a congressman learns that his committee is about to do something that would affect a company, he can go trade on that because he is not obligated to keep that information confidential…. He is not breaching a duty of confidentiality to anybody."[52] To be sure, if the *Cuban* decision discussed above[53] becomes widely accepted, and the congressional ethics manuals' prohibition of insider trading by members is deemed to provide the requisite agreement, liability might be imposed on members who violate that obligation.[54] As we have seen, however, that decision remains highly controversial.[55]

An apt precedent for treating stock trading by congressional staffers and members of Congress differently is provided by *U.S. v. Carpenter*.[56] R. Foster Winans wrote the widely read "Heard on the Street" column for the Wall Street Journal, which

provides investing information and advice. Because that column apparently had a short lived effect on the price of the stocks it covered, someone who knew the column's contents in advance could profit by trading in the affected stocks. Although Wall Street Journal policy stated that prior to their publication the contents of columns were the Journal's confidential property, Winans, before publication, disclosed the contents of his columns to several friends who then traded in the affected stocks. Winans and his friends were convicted of mail and wire fraud and insider trading under Rule 10b-5 pursuant to the misappropriation theory.[57]

In *Carpenter*, the Second Circuit held that Winans and his fellow conspirators committed illegal insider trading by "secreting, stealing, purloining or otherwise misappropriating material non-public information in breach of an employer-imposed fiduciary duty of confidentiality," on the basis of which they then traded in the stock of issuers mentioned in Winans' columns.[58] In dicta, the court indicated that the Wall Street Journal could have traded on the basis of the information in question:

> Appellants argue that it is anomalous to hold an employee liable for acts that his employer could lawfully commit. Admittedly, … the Wall Street Journal or its parent, Dow Jones Company, might perhaps lawfully disregard its own confidentiality policy by trading in the stock of companies to be discussed in forthcoming articles.… Although the employer may perhaps lawfully destroy its own reputation, its employees should be and are barred from destroying their employer's reputation by misappropriating their employer's informational property.… Here, appellants, constrained by the employer's confidentiality policy, could not lawfully trade by fraudulently violating that policy, even if the Journal, the employer imposing the policy, might not be said to defraud itself should it make its own trades.[59]

Nothing in *O'Hagan* is inconsistent with the distinction drawn in *Carpenter*. The misappropriation theory bans undisclosed trading by an agent or other fiduciary in breach of a duty of loyalty to the principal; it does not ban trading by the principal in the same information, even if the agent in question developed the information for the principal. As an employer, a member of Congress is free to trade; as an employee, the staffer is not.

II. Policy

Over 40 years ago, Henry Manne observed that "the federal government is the largest producer of information capable of having a substantial effect on stock-market prices."[60] To the extent the government does not itself generate such information, vast amounts of information must be disclosed to the federal government before it becomes public.[61] Congressmen are especially well-positioned to receive nonpublic information, Manne argued.[62] In addition to their direct interactions with nongovernmental information sources, they are also "focal points for receiving information produced or learned in all the various executive departments and agencies" that report to them.[63]

The argument for prohibiting insider trading by members of Congress and other government employees is straightforward: "the ability of elected officials to profit on the basis of material nonpublic information creates perverse incentives for these officials, and introduces innumerable distortions and the potential for immeasurable harm in a legal system in which public trust and confidence is critical."[64] As Larry Ribstein observes:

> Congress's insider trading is bad because it gives our lawmakers the wrong incentives. Do we really want to give Congress more reasons to hurt and help particular firms?
>
> In fact, Congress's trading is worse than trading by corporate insiders, which at least might be rationalized as a way to let employees cash in on their productive efforts. It's far worse than the usual trading on non-public information by outsiders without any breach of duty, which may encourage socially productive investigation and monitoring....[65]

Congressional insider trading thus is undesirable, in the first instance, because it creates incentives for members and staffers to steal proprietary information for personal gain. The massive increase in federal involvement in financial markets and corporate governance as a result of the financial crisis of 2008 has made opportunities to steal such information even more widely available to government officials. Second, it gives members and staffers incentives to game the legislative process so as to maximize personal trading profits. Indeed, some members of Congress are so prominent that their pronouncements could move the market, allowing them to profit even further from trading on undisclosed information. Third, inside information can be utilized as a pay-off device."[66] Fourth, it gives members and staffers incentives to help or hurt firms, which distorts market competition.

In sum, the point hardly requires belaboring. There is no plausible justification for allowing members of Congress or other governmental actors to use material nonpublic information they learn as a result of their position for personal stock trading gains. To the contrary, the policy arguments all come down on the side of banning such trading.

III. The Stop Trading on Congressional Knowledge Act

Congressmen Brian and Slaughter introduced versions of the STOCK Act in the 109th,[67] 110th,[68] and now the 111 1th Congresses.[69] According to Congressman Brian, the current version of the Act would:

- Prohibit Members or employees of Congress from buying or selling stocks, bonds, or commodities futures based on nonpublic information they obtain because of their status;
- Prohibit Executive Branch employees from buying or selling stocks, bonds or commodities futures based on nonpublic information they obtain because of their status;
- Prohibit those outside Congress from buying or selling stocks, bonds, or commodities futures based on nonpublic information obtained from within Congress or the Executive Branch;

- Prohibit Members and employees from disclosing any non-public information about any pending or prospective legislative action obtained from a member or employee of Congress for investment purposes;

- Require Members of Congress and employees to report the purchase, sale or exchange of any stock, bond, or commodities future transaction in excess of $1,000 within 90 days. Members and employees who choose to place their stock in holdings in blind trusts or mutual funds would be exempt from the reporting requirement….[70]

A. The Prohibition on Trading and Tipping

The STOCK Act is not self-executing. To the contrary, it mostly dumps the problem into the SEC's lap by directing the Commission to undertake a number of rulemaking proceedings.

For present purposes the key operative provision is Section 2(a), in which the SEC is directed to adopt a rule prohibiting:

> [A]ny person from buying or selling the securities of any issuer while such person is in possession of material nonpublic information, as defined by the Commission, relating to any pending or prospective legislative action relating to such issuer if—
>
> (1) such information was obtained by reason of such person being a Member or employee of Congress; or
>
> (2) such information was obtained from a Member or employee of Congress, and such person knows that the information was so obtained.[71]

A rule comporting with this provision would ban members of Congress and congressional staffers from trading on the basis of material nonpublic information obtained by virtue of their position. It also would ban the tippee of a member or staffer from trading so long as the tippee knew the information was obtained from a member or staffer. The provision thus solves the doctrinal problems associated with prosecuting members of Congress who commit insider trading. Other provisions of the Act do likewise with respect to federal government employees generally.

Note, however, that there is a key limitation on the scope of the prohibition authorized by the Act; namely, the information must relate to a "pending or prospective legislative action," which action in turn must relate to the issuer of the securities traded. As to the former aspect, how broadly will "legislative action" be interpreted? As to the latter, information about one issuer may often allow one to profit by trading in the securities of another company.

Consider the following cases:

- After Congress defeats proposed legislation that would have sharply increased Acme's costs of doing business, Acme's CEO gives a key Congressman a hot tip on Acme stock as a pay off. There was a legislative action, but it was in the past and, accordingly, is neither pending nor prospective.

- A member of Congress learned from a Cabinet member that a government agency was about to enter a large procurement contract. There is no "pending or prospective" legislative action, but there is valuable material nonpublic information on which the member could trade.

- The CEO of Acme is an avid hunter. Congress is considering legislative action that would ban hunting of the CEO's favorite game animal. The CEO of Acme gives a key Congressman a hot tip on Acme stock as a bribe to oppose the hunting law. This is perhaps the most egregious form of Congressional insider trading, yet it would not seem to relate to "such issuer" and thus would not be prohibited.

- During a confidential committee investigation, a member of Congress learns that Acme is about to announce a major new discovery. The member infers that Ajax—Acme's major competitor—will take a serious hit. The member shorts Ajax stock. Technically, the member has not traded in the stock of "such issuer."

Three other problems with the present statutory language deserve mention. First, the Act applies only to "the securities of any issuer." The rulemaking authorized by the Act thus could not proscribe trading in third-party derivatives, such as options. Second, while the Act authorizes a ban on tippee trading, it does not expressly authorize a regulatory ban on tipping by members or staffers.[72] There is no reason members and staffers should be allowed to tip with impunity. Finally, Rule 14e-3 prohibits tippees from trading on the basis of material nonpublic information about a tender offer not only if the tippee knows the information came from one of the specifided sources, but also if the tippee "has reason to know" that it came from a proscribed source. The STOCK Act should do likewise with respect to information obtained from a member or staffer.

Congress could solve these problems by broadening the grant of rulemaking authority given the SEC by the Act even further, so as to allow the SEC to address harms related to those to which the Act is addressed. We have known about the need to address congressional insider trading at least since Manne's 1966 book, however, and the SEC has been "happily complicit with" the failure to address it.[73] Accordingly, it seems safe to assume that the SEC will be loath to bite the hand that feeds its budget by taking an aggressive stance in the Act's mandated rulemaking proceedings. Where there are known gaps, such as those identified here, Congress therefore should fill them expressly.

B. Reporting Provision

The STOCK Act would require that a member of Congress disclose to the Clerk of the House or Secretary of the Senate, as the case may be, transactions of $1,000 or more in "any stocks, bonds, commodities futures, or other forms of securities that are otherwise required to be reported under" the Ethics in Government Act.[74] A member has up to 90 days after the transaction is effected to disclose it.[75] This compares quite unfavorably with the two days corporate insiders have to report transactions covered by Section 16 of the Securities and Exchange Act.[76] A shorter disclosure window is in order.

IV. Conclusion

Insider trading by corporate insiders has been banned for over four decades.[77] Throughout that period, we have known that insider trading by members of Congress was a potential problem that arguably presented even more serious policy concerns than trading by classic insiders. Congressional insider trading creates perverse legislative incentives and opens the door to serious corruption. Yet, both Congress and the SEC have turned a blind eye.

The STOCK Act would fix the doctrinal obstacles to prosecuting members of Congress who commit insider trading. If passed, it also might finally give the SEC political cover to actually bring such cases. Although the present Act still needs work, it is long overdue.

Endnotes

1 Bob Ryan, Corporation Finance and Valuation 84 (2006) ("The empirical evidence is absolutely solid, fund managers cannot out perform the market….").

2 Hasan Nejat Seyhun, Investment Intelligence From Insider Trading 312 (2000).

3 *Id.* at 74.

4 Alan J. Ziobrowski et al., Abnormal Returns from the Common Stock Investments of the U.S. Senate, 39 J. Fin. & Quant. Anal. 661 (2004).

5 Jane J. Kim, U.S. Senators' Stock Picks Outperform the Pro's, Wall. St. J., Oct. 26, 2004, available at http://tinyurl.com/nrwm6r.

6 *Id.* The extent of Congressional trading on material nonpublic information is uncertain. "Just over a third of the senators bought or sold individual stocks in any one year in the study, and the vast majority of stock transactions were less than $15,000." *Id.*

7 U.S. v. Chiarella, 588 F.2d 1358, 1365 n.8 (2d Cir. 1978), rev'd on other grounds, 445 U.S. 222 (1980).

8 Kim, *supra* note 5.

9 *See* Chiarella v. U.S., 445 U.S. 222, 240 n.1 (1980) (Burger, C.J., dissenting) ("Academic writing in recent years has distinguished between 'corporate information'—information which comes from within the corporation and reflects on expected earnings or assets—and 'market information.' … It is clear that § 10(b) and Rule 10b-5 by their terms and by their history make no such distinction."). *See also* Dirks v. S.E.C., 463 U.S. 646, 656 n.15 (1983) (citing Chief Justice Burger's dissent approvingly).

10 *See* S.E.C. v. Tome, 638 F. Supp. 596, 616 (S.D.N.Y. 1986) ("The term 'insider trading' actually is a misnomer, only imperfectly describing the proscribed conduct, since liability under the securities laws can extend to those who are not insiders, as that term is commonly understood …").

11 Insider trading potentially can give rise to both civil and criminal actions. The SEC is authorized by Section 21A(a)(1) of the Securities Exchange Act to bring a civil action against insider traders in a United States district court, seeking a variety of civil penalties, including disgorgement and fines of up to three times the defendant's profits. 15 U.S.C. § 78u-1(a)(1). Under Section 32 of the Exchange Act, the Department of Justice may bring felony criminal charges against those who willfully violate, inter alia, the rules against insider trading. 15 U.S.C. § 78ff(a). Unless the context requires otherwise, the term "prosecute" will be used herein, for purposes of semantic convenience, to refer to both SEC civil and DOJ criminal proceedings.

12 H.R. 682, 111th Cong., 1st Sess. (2009).

13 Press Release, Brian Introduces Legislation to Prohibit Insider Trading on Capitol Hill (Jan. 27, 2009), available at http://tinyurl.com/ne77mw.

14 17 CFR § 240.10b-5.

15 15 U.S.C. § 78j(b). Other bases of insider trading liability, such as Rule14e-3's restrictions on the use of information about a tender offer, 17 C.F.R. § 240.14e-3, or the federal mail and wire fraud statutes, 18 U.S.C. § 1341 and 1343, are outside the scope of this article.

16 S.E.C. v. Cuban, 2009 WL 2096166 (N.D. Tex. 2009) (noting that insider trading generally is "defined by judicial opinions construing Rule 10b-5").

17 S.E.C. v. Rocklage, 470 F.3d 1, 5 (1st Cir. 2006).

18 401 F.2d 833 (2d Cir. 1968), cert. denied, 394 U.S. 976 (1969).

19 445 U.S. 222 (1980).

20 *Id.* at 232.

21 463 U.S. 646 (1983).

22 *Id.* at 254-55.

23 *Id.* at 655 n.14.

24 *Id.* at 659-60.

25 *Id.* at 659 (quoting *Chiarella*, 445 U.S. at 230).

26 *Id.* at 660.

27 *See* Kevin V. Haynes, Insider Trading Under Rule 10b-5, in Fundamentals of Securities Law, WL SP057 ALI-ABA 411 ("In *Dirks* the Court alludes to the inherent unfairness of insiders' trading on information that was intended to be available only for corporate purposes. This suggests that the insider breaches a fiduciary duty to refrain from self-dealing when he trades on inside information."). *See also* Keith Valory, Comment, The Misappropriation Theory of Insider Trading: What Constitutes a "Similar Relationship of Trust and Confidence?," 39 Santa Clara L. Rev. 287, 296 (1998) ("It is clear from the Court's rulings on the Classical Theory of insider trading that it is most concerned with curtailing deceptive self-dealing in material, non-public corporate information.").

28 *See Dirks*, 463 U.S. at 663 (stating that the test of whether a disclosure by an insider amounts to a breach of fiduciary duty focuses on "objective criteria, i.e., whether the insider receives a direct or indirect personal benefit from the disclosure").

29 *See* S.E.C. v. Yun, 327 F.3d 1263, 1275 (11th Cir. 2003) (holding that "the gain does not always have to be pecuniary. A reputational benefit that translates into future earnings, a quid pro quo, or a gift to a trading friend or relative all could suffice to show that the tipper personally benefitted.").

30 *Id.*

31 *Id.*

32 Committee on Standards of Official Conduct, Restrictions on Outside Employment Applicable to Members and Senior Staff, available at http://tinyurl.com/mz3eyv.

33 *Cf.* 65 Fed. Reg. 51,716 n.7 (2000) (SEC release acknowledging how selective disclosures to analysts was viewed as protected from insider trader liability because tipper received no personal benefit but rather provided the tip so as to benefit corporation).

34 521 U.S. 642 (1997).

35 *Id.* at 652.

36 *Id.* at 653.

37 *Id.* at 652.

38 *Id.*

39 *Id.* at 656-57.

40 Civil Action No. 3:08-CV-2050-D (N.D. Tex. July 17, 2009).

41 No. 08-0201-cv (2d Cir. July 22, 2009).

42 *See, e.g.,* Regents of the Univ. of Cal. v. Credit Suisse First Boston (USA), Inc., 482 F.3d 372, 389 (5th Cir. 2007) (holding that "the [Supreme] Court . . . has established that a device, such as a scheme, is not 'deceptive' unless it involves breach of some duty of candid disclosure").

43 5 C.F.R. § 2635.101.

44 Joseph Kalo, Deterring Misuse of Confidential Government Information:

A Proposed Citizens' Action, 72 Mich. L. Rev. 1577, 1581 (1974) ("The application of fiduciary duties to activities of government employees is not novel.").

45 5 C.F.R. § 2634.703(a). Nonpublic information is defined for this purpose as "information that the employee gains by reason of Federal employment and that he knows or reasonably should know has not been made available to the general public." *Id.* § 2634.703(b).

46 Email from Mary Baumann, U.S. Senate Historical Office, to Jenny Lentz, UCLA School of Law Library, Friday, July 31, 2009 (copy on file with author).

47 House Ethics Manual 249 (2008).

48 933 F.2d 403 (7th Cir. 1991), cert. denied, 502 U.S. 1071 (1992).

49 *Id.* at 410. See also SEC v. Clark, 915 F.2d 439, 453 (9th Cir. 1990) ("[A]n employee's knowing misappropriation and use of his employer's material nonpublic information regarding its intention to acquire another firm constitutes a violation of § 10(b) and Rule 10b-5.").

50 U.S. v. O'Hagan, 521 U.S. 642, 652 (1997). See generally Andrew George, Public (Self)-Service: Illegal Trading on Confidential Congressional Information, 2 Harv. L. & Poly' Rev. 161, 165-66 (2008) (concluding that staffers can be held liable either for trading on or tipping of material nonpublic information learned on the job).

51 United States v. Woodard, 459 F.3d 1078, 1086 (11th Cir. 2006).

52 Brody Mullins, Bill Seeks to Ban Insider Trading by Lawmakers and Their Aides, Wall St. J., Mar. 28, 2006, at A1.

53 *See supra* text accompanying notes 40-42.

54 *See* George, *supra* note 50, at 166-68 (arguing that the House and Senate ethics manuals constitute the requisite agreement).

55 *See supra* text accompanying note 42.

56 791 F.2d 1024 (2d Cir. 1986), aff'd on other grounds, 484 U.S. 19 (1987).

57 The Supreme Court affirmed on all counts, but the securities fraud convictions were affirmed only by an evenly divided Court (4-4). 484 U.S. 19 (1987). By long-standing tradition, a decision by an evenly divided court affirms the lower court result but has no precedential or stare decisis effect, which left the legitimacy of the misappropriation theory uncertain until the *O'Hagan* decision validated it.

The federal mail and wire fraud statutes, 18 U.S.C. §§ 1341 and 1343, respectively prohibit the use of the mails and "wire, radio, or television communication" for the purpose of executing any "scheme or artifice to defraud." The mail and wire fraud statutes protect only property rights, McNally v. U.S., 483, U.S. 350 (1987), but confidential business information is deemed to be property for purposes of those statutes. Carpenter v. U.S., 484 U.S. 19, 25 (1987). Hence, the Supreme Court held, the Wall Street Journal owned the information used by Winans and his co-conspirators and, moreover, that their use of the mails and wire communications to trade on the basis of that information constituted the requisite scheme to defraud.

58 *Carpenter*, 791 F.2d at 1032.

59 *Id.* at 1033-34.

60 Henry G. Manne, Insider Trading and the Stock Market 171 (1966).

61 *Id.* at 172.

62 *Id.* at 179.

63 *Id.*

64 Jonathan R. Macey and Maureen O'Hara, Regulation and Scholarship: Constant Companions or Occasional Bedfellows?, 26 Yale J. on Reg. 89, 108 (2009).

65 Larry E. Ribstein, Congress' Insider Trading, Ideoblog, Mar. 29. 2006.

66 Manne, *supra* note 60, at 184.

67 H.R. 5015, 109th Cong., 2d Sess. (2006).

68 H.R. 2341, 110th Cong., 1st Sess. (2007).

69 H.R. 682, 111th Cong., 1st Sess. (2009).

70 Press Release, *supra* note 13.

71 H.R. 682, 111th Cong., 1st Sess. § 2(a) (2009).

72 The Act does amend the Congressional ethics rules to ban tipping with respect to pending or prospective legislative action. H.R. 682, 111th Cong., 1st Sess. § 3 (2009).

73 Macey & O'Hara, *supra* note 64, at 109.

74 H.R. 682, 111th Cong., 1st Sess. § 4 (2009).

75 *Id.*

76 *See* 15 U.S.C. § 78p(a) (setting out disclosure rules for specified corporate insiders).

77 The modern federal insider trading prohibition is generally regarded to have originated with *SEC v. Texas Gulf Sulphur Co.*, 401 F.2d 833 (2d Cir. 1968), cert. denied, 394 U.S. 976 (1969).

The Financial Reform Plan: What It Means for Insurance Companies

*By Laura Kotelman**

On June 17, 2009, the U.S. Treasury Department released the Obama Administration's framework for financial regulatory reform. As part of the larger effort to strengthen the regulation of the financial services market, the Financial Reform Plan proposes certain reforms applicable to the insurance industry, including establishment of an Office of National Insurance (the "ONI") within the Treasury Department, modernization of insurance regulation in accordance with six fundamental principles, special treatment of large systemic insurance conglomerates (known as "Tier 1 FHCs"), and the creation of a Consumer Financial Protection Agency (the "CFPA").

I. Office of National Insurance

As proposed, the ONI would monitor all aspects of the insurance industry, primarily by gathering information and identifying the emergence of potential market problems or gaps in existing regulation that could contribute to a financial crisis. The proposal does not provide the ONI with any expressly enumerated enforcement powers; however, it is contemplated that the ONI will be empowered to work with other nations to better represent U.S. interests and increase international cooperation on insurance regulation. It will have the authority to enter into international agreements; carry out the federal government's existing responsibilities under the Terrorism Risk Insurance Act; recommend to the Federal Reserve Board any insurance conglomerates that should be regulated as Tier 1 FHCs (described in more detail below); and consult with the Treasury Department in connection with the orderly resolution of a failing Tier 1 FHC with insurance subsidiaries. Other than the ONI's role in negotiating international agreements, it is not clear from the Financial Reform Plan how the administration intends to implement these principles of reform, whether through federal statute, which could include an optional federal charter for insurance companies, uniform collective action at the state level, or some other implementation procedure.

The Financial Reform Plan enumerates six principles that the Treasury Department will support for modernizing the regulation of insurance:

- Effective systemic risk regulation with respect to insurance, which could be addressed in large measure by the adoption of the proposed Tier 1 FHC regulations;
- Strong capital standards and an appropriate match between capital allocation and liabilities, including risk management related to liquidity and duration risk;
- Meaningful and consistent consumer protection for insurance products and practices to address any gaps that exist under the current regulatory system;
- Increased national uniformity through an optional federal charter or effective action by the states. The Treasury Department described the current insurance regulatory regime as "highly fragmented, inconsistent and inefficient" with marked variances in consumer protections among the states;
- Regulation of insurance companies and affiliates on a consolidated basis, including non-insurance affiliates. The Treasury Department indicated that any new regulations should address the gaps in current insurance holding company regulations; and
- Coordination among international regulatory authorities.

II. Regulation of Tier 1 FHCs

The ONI would be responsible for recommending to the Federal Reserve Board any insurance conglomerates that the ONI believes should be regulated by the Federal Reserve Board as a financial institution whose failure could pose a systemic risk to the financial system (known as "Tier 1 FHCs"), regardless of whether the insurance conglomerate owned a bank or was regulated as a bank holding company. In its analysis, the ONI would consider, among other factors: the impact the insurance conglomerate's failure would have on the financial system and the economy; the insurance conglomerate's combination of size, leverage, and degree of reliance on short-term funding; and the insurance conglomerate's importance as a source of credit for households, businesses, and state and local governments and as a source of liquidity for the financial system.

The Financial Services Oversight Council that would be created to advise the Federal Reserve Board provides for no state role in this agency. The regulators who would constitute this panel would include the Secretary of the Treasury, who would serve as chairman; the Chairman of the Federal Reserve; the Director of the combined Office of the Comptroller of the Currency/Office of Thrift Supervision; the Director of the Consumer Financial Protection Agency; the Chairman of the Securities and Exchange Commission; the Chairman of the Commodities Future Trading Commission; the Chairman of the Federal Deposit Insurance Corporation; and the Director of the Federal Housing Finance Agency.

Being subject to the Tier 1 FHC regulations, in addition to other insurance regulations, could have a significant impact on large insurance companies. Once an institution is designated as a Tier 1 FHC, the Federal Reserve Board's supervisory authority would extend to the parent company and all subsidiaries, U.S. and foreign, including otherwise regulated subsidiaries. Although existing state insurance regulators would remain the primary regulator for an insurance company, the Federal Reserve Board would have the authority to provide its own level of oversight. Tier 1 FHCs could be subject to, among other provisions, increased capital and liquidity requirements and risk management standards (including application of severe stress scenarios). Tier 1 FHCs would be subject to the restrictions of the Bank Holding Company Act on nonfinancial activities,

** Laura Kotelman is a member of the Financial Services & E-Commerce Practice Group.*

which could affect insurance companies that are affiliated with non-financial entities. Such insurance conglomerates may be required to divest themselves of any such non-financial subsidiaries within five years.

In addition, insurance conglomerates that are Tier 1 FHCs could be subject to a proposed special resolution regime providing for the quick and orderly resolution of failing Tier 1 FHCs. The resolution regime would apply to a particular Tier 1 FHC only when activated by the Treasury Department (in consultation with the ONI for insurance specific matters), and, unless activated, the resolution of a Tier 1 FHC would be governed by the U.S. Bankruptcy Code or applicable state insurance insolvency provisions, as the case may be. The Treasury Department would be given the power to decide resolution alternatives, which may include, among other features, appointing a conservator or receiver or stabilizing the failing Tier 1 FHC by providing loans, purchasing assets, guaranteeing the liabilities, or making equity investments with respect to the Tier 1 FHC. The Financial Reform Plan does not address the interplay of this special resolution regime with the traditional state-based insurance insolvency regime which would otherwise apply to insurance company subsidiaries.

III. Consumer Financial Protection Agency

The Financial Reform Plan proposes to create a new federal agency, the CFPA, with jurisdiction over credit, savings, payment, and other consumer financial products and services, other than investment products already regulated by the SEC or the CFTC. It is not clear from the Financial Reform Plan whether the CFPA is intended to have jurisdiction over insurance products. Although almost all examples noted in the Financial Reform Plan are banking products, the proposed general description of the jurisdiction of the CFPA would seemingly encompass many types of insurance products, as the CFPA is expressly intended to have jurisdiction over nonbanking entities. A clear goal of the CFPA is to consolidate federal consumer protection regulation, but the proposal explicitly permits concurrent state consumer protection regulation without federal preemption.

IV. The Future of Insurance Regulation

Under the proposal, the only role for federal authority in insurance would be the creation of an Office of National Insurance within the Treasury Department. However, a Treasury white paper detailing the reform plan states that the administration would support "increased national uniformity through either a federal charter or effective action by the states." Therefore, insurance companies could be subject to federal regulation through an optional federal charter regime, the regulation of insurance holding companies, the oversight of the CFPA, or international agreements arranged by the ONI.

Certain large insurance companies would be subject to unprecedented regulation by the Federal Reserve Board due to their status as Tier 1 FHCs. Insurance conglomerates that are failing Tier 1 FHCs may be subject to a special resolution process controlled by the Treasury Department. It is not clear how the special resolution process would affect the current state-based insurance insolvency system which would normally apply to insurance company subsidiaries. It is not clear whether the CFPA will have regulatory power over insurance products. The Financial Reform Plan does not clearly spell out if the roles of the federal government and state governments will be complementary or exclusive. The ONI within the Treasury would have the authority to recommend to the Fed any insurance companies that the ONI believes should be supervised as Tier 1 financial holding companies. However, the ONI does not appear to have enforcement powers, as compared to certain other federal regulatory agencies such as the SEC.

The National Association of Insurance Commissioners (NAIC), the organization of insurance regulators from the 50 states, the District of Columbia, and the five U.S. territories initially praised the proposal despite some federal regulatory encroachment. According to the NAIC, the proposal basically leaves the state insurance regulatory system intact, reserving the consumer protections and financial solvency oversight for the state-based insurance regulatory system. Nevertheless, state regulators will have to work with Congress and the administration to underscore the benefits of the current state-based insurance regulatory system. While, according to the NAIC, the plan addresses systemic risk and other regulatory gaps, consumer protection will remain priority one for state insurance officials. Who is better capable of adapting as insurance markets evolve will have to be answered by the companies they regulate.

Free Speech & Election Law

A Cold Breeze in California: *ProtectMarriage* Reveals the Chilling Effect of Campaign Finance Disclosure on Ballot Measure Issue Advocacy

*By Stephen R. Klein**

On November 4, 2008, the election of President Obama coincided with the passage of Proposition 8, a ballot measure which banned gay marriage in California through amendment of the state's constitution.[1] In the days leading up to and following the passage of the proposition, public access to the names and pertinent information of individual donors supporting the bill led to some interesting results:

> [W]hen it was discovered that Scott Eckern, director of the nonprofit California Musical Theater in Sacramento, had given $1,000 to Yes on 8, the theater was deluged with criticism from prominent artists. Mr. Eckern was forced to resign. Richard Raddon, the director of the L.A. Film Festival, donated $1,500 to Yes on 8. A threatened boycott and picketing of the next festival forced him to resign. Alan Stock, the chief executive of the Cinemark theater chain, gave $9,999. Cinemark is facing a boycott, and so is the gay-friendly Sundance Film Festival because it uses a Cinemark theater to screen some of its films.[2]

More disturbingly, "[s]ome donors to groups supporting the measure... received death threats and envelopes containing powdery white substance...."[3] Many of these threats were possible only because the names and ZIP codes of donors and the amounts of their respective donations are made publicly available and posted on the internet pursuant to California law.[4] However, unlike previous elections, many names were widely circulated elsewhere on the internet and led to the emergence of websites such as eightmaps.com.[5] This website combines the donor list with Google Maps and gives any visitor to the site an aerial view of the donor's home.[6]

In the midst of this fallout, some pro-Prop 8 committees and donors have sued in the Eastern District of California to enjoin the enforcement of semiannual reporting requirements, to enjoin any criminal or civil actions for failure to comply with reporting requirements, and to enjoin the publishing of reports or statements filed previously.[7] The court denied a preliminary injunction and concluded that "in this case... no serious First Amendment questions are raised."[8]

This article argues to the contrary: although a state government may have an interest in disseminating donor information behind some campaigns for or against ballot measures, the Ninth Circuit's interpretation of the "informational interest" from *Buckley v. Valeo* was not a concern in Proposition 8, which implicated a purely social issue. Thus, in light of the use of donor information to abridge free speech, this articulation of the informational interest does not survive strict scrutiny: as applied, California's disclosure law indirectly infringes upon First Amendment rights by facilitating the suppression of political speech.

I. Getman and ProtectMarriage: Ballot Measure Disclosure in the Ninth Circuit

On January 30, 2009, Judge Morrison England, Jr., denied a preliminary injunction in the *ProtectMarriage* case. The Ninth Circuit's current stance (and, as a result, the stance of the Eastern District of California) on compelled disclosure for money spent on direct democratic lawmaking is well-intentioned, but, in light of Proposition 8 and other social-issue ballot measures, provides a tool for chilling political speech. Furthermore, such disclosure is not supported by *Buckley v. Valeo* and its progeny.[9]

The committees bringing the *ProtectMarriage* case include ProtectMarriage.com, the National Organization for Marriage, and John Doe #1, who also represents a class of pro-Proposition 8 donors.[10] The plaintiffs filed a number of anonymous declarations from John Does, nine of which the court describes in its denial for preliminary injunction.[11] Many of these declarations include claims that the John Does will be reluctant to make similar types of donations in the future.[12] The plaintiffs claim that "'California's threshold for compelled disclosure of contributors is not narrowly tailored to serve a compelling government interest in violation of the First Amendment to the United States Constitution.'"[13] For purposes of a preliminary injunction, Judge England rejects this argument.[14]

A. Precedent (Or Lack Thereof)

"Plaintiffs concede... that California has a compelling justification for requiring disclosure of Plaintiffs contributors."[15] However, after stating that this concession "gives short shrift to both the nature and magnitude of the State's actual interest,"[16] Judge England determines that "the Supreme Court has repeatedly emphasized the importance of disclosure as it relates to the passage of initiatives."[17] Rather than address–much less name–these Supreme Court cases, Judge England supports his assertion with a citation to a Slip Opinion from the remand of *California Pro-Life Council v. Getman*.[18] On remand, in "*Getman II*" Judge Frank Damrell, Jr., stated that "the Supreme Court repeatedly has recognized the importance of expenditure and contribution disclosure in the ballot measure context."[19] He cited three cases: *First National Bank of Boston v. Bellotti*,[20] *Citizens Against Rent Control v. City of Berkeley*,[21] and *Buckley v. American Constitutional Law Foundation*.[22] This was a shorter repetition of the Ninth Circuit's decision in *Getman*.[23]

Although Judge Damrell's assertion regarding these cases is not entirely inaccurate, he neglected to mention that these cases are, at best, persuasive authority: *Bellotti* overturned restrictions

** Stephen R. Klein is a member of the Executive Committee of the Free Speech and Election Law Group.* *The author thanks Professor John H. Dudley and Jason C. Miller for their comments.*

on corporate advertising in a public issue election.[24] The Court merely stated in its reasoning that "[the people] may consider, in making their judgment [on how to vote], the source and credibility of the advocate."[25] In a footnote the Court stated that "[i]dentification of the source of advertising may be required as a means of disclosure,"[26] but the Court discussed only corporate sponsorship, not individual contributors, and the extent of disclosure was not before the Court.

In *Citizens Against Rent Control*, the Court overturned a law prohibiting contributions greater than $250 to committees formed to support or oppose ballot measures.[27] The Court stated that "[t]he integrity of the political system will be adequately protected if contributors are identified in a public filing revealing the amounts contributed; if it is thought wise, *legislation can outlaw anonymous contributions*."[28] However, the issue of anonymous contributions was not before the Court, nor was a regulatory scheme that would disclose contributions for issue advocacy. Finally, in *American Constitutional Law Foundation*, the Court upheld "[d]isclosure of the names of [ballot] initiative sponsors, and of the amounts they have spent gathering support for their initiatives" as a substantial state interest.[29] The Court addressed the informational interest of money spent to "get a measure on the ballot," but did not address disclosure of donors behind political speech once initiatives have been placed on a ballot.[30]

The Ninth Circuit and the Eastern District of California declare that disclosure of issue advocacy is a compelling state interest, but offer no specific precedent. Although some dicta in *Bellotti* and *Citizens Against Rent Control* is quite strong, there is otherwise little to support an informational interest in ballot measures. Perhaps aware of this, both the Ninth Circuit in *Getman* and Judge England in *ProtectMarriage* articulate an informational interest for ballot measure campaigns and contend this interest is in step with *Buckley v. Valeo*.

B. The "Informational" Interest

The *ProtectMarriage* case cites *Buckley*'s three categories of disclosure,[31] and recognizes that "unlike the election before the *Buckley* court, which concerned candidates, the instant case bears on a recent ballot-initiative measure."[32] Judge England continues to rely on the Ninth Circuit's *Getman* precedent and reiterates that the "informational interest," the first category of disclosure in *Buckley*,[33] provides a compelling state interest:

> The influx of money [into ballot measures]... produces a cacophony of political communications through which California voters must pick out meaningful and accurate messages. Given the complexity of the issues and the unwillingness of much of the electorate to independently study the propriety of individual ballot measures, ... being able to evaluate who is doing the talking is of great importance.[34]

Judge England then articulates numerous reasons for this informational interest, but throughout his analysis he fails to recognize that these concerns are not raised in the present case.

1. Understanding the Policy Content of a Ballot Measure

Judge England begins with ballot measures themselves:

> While the ballot pamphlet sent to voters by the state contains the text and a summary of ballot measure initiatives, many voters do not have the time or ability to study the full text and make an informed decision. Since voters might not understand in detail the policy content of a particular measure, they often base their decisions to vote for or against it on cognitive cues such as the names of individuals supporting or opposing a measure....[35]

Leaving aside not-so-subtle-hints of a governmental interest in basing disclosure on the lowest common denominator of citizenry, the policy content of Proposition 8 required very little effort to understand.[36] A vote of "yes" supported constitutionally prohibiting gay marriage; a vote of "no" supported the California Supreme Court's ruling in *In re Marriage Cases*.[37] Moreover, disclosure can just as easily detract from discovering the detail of policy content because it shifts focus to the advocates over the advocacy.[38] Rather than promote the discussion of issues, disclosure allows opposing parties to obfuscate issues with accusations of ulterior motives.[39] Assuming this prong of the Ninth Circuit's informational interest is valid to begin with, it was not a concern in the Proposition 8 campaigns.

2. Citizen-Legislators

Judge England continues to describe the informational interest by again quoting the Court of Appeals:

> [V]oters act as legislators in the ballot-measure context, and interest groups and individuals advocating a measure's defeat or passage act as lobbyists.... Californians, as lawmakers, have an interest in knowing who is lobbying for their vote, just as members of Congress may require lobbyists to disclose who is paying for the lobbyists' services and how much.[40]

In *Getman*, the Ninth Circuit drew this principle from *United States v. Harriss*,[41] which upheld the Lobbying Act.[42] The Supreme Court reasoned that without disclosure to Congress of contributions made to lobbyists, "the voice of the people may all too easily be drowned out by the voice of special interest groups seeking favored treatment while masquerading as proponents of the public weal."[43]

Harriss was decided before *Buckley*, and *Buckley* cited the case three times.[44] In pertinent part, the *Harriss* case was used as support not for the informational interest but for the second disclosure interest, corruption or the appearance of corruption: "A public armed with information about a candidate's most generous supporters is better able to detect any post-election special favors that may be given in return."[45] Treating citizens as legislators with loose reference to *Harriss* does not withstand this classification. Money paid to lobbyists is (or appears to be) property used in exchange for preferential treatment. With this money, lobbyists are paid to persuade members of Congress to vote on certain issues. Disclosure provides members of Congress with information as to where this persuasion comes from. More importantly, this serves *the electorate* by ensuring them that their respective votes can be entrusted with their legislator.

When the citizen is the legislator, their vote is not entrusted to anyone else, and there is no danger of indirect corruption or the appearance of corruption.

This prong, then, while apparently arising from different authority, is largely the same as the first prong: understanding the policy implications of a measure by understanding who the advocates are. Thus, the same criticisms of that prong in the previous section serve to dispel this prong of the informational interest in light of Proposition 8.[46] Judge England also includes this statement from the Ninth Circuit: "'While we would hope that California voters will independently consider the policy ramifications of their vote, and not render a decision based upon a thirty-second sound bite they hear the day before the election, we are not that idealistic nor that naïve.'"[47] Again, the Ninth Circuit's distrust of the electorate's independent consideration causes the court to recognize a compelling state interest in disclosing information voters will consider *instead* of the actual issues behind each ballot proposition.

3. Accurate Identification of Advocate

The final prong of the Ninth Circuit's informational interest is as follows[48]:

> Knowing which interested parties back or oppose a ballot measure is critical, especially when one considers that ballot-measure language is typically confusing, and the long-term policy ramifications of the ballot measure are often unknown. At least by knowing who supports or opposes a given initiative, voters will have a pretty good idea of who stands to benefit from the legislation.[49]

Furthermore, "because groups supporting and opposing ballot measures frequently give themselves ambiguous or misleading names, reliance on the group, without disclosure of its source of funds, can be a trap for unwary voters."[50] Once again, it is notable that although many ballot measures in California are long and potentially confusing to the average voter,[51] Proposition 8 was not one of those measures.[52] Judge England points to special interest groups; he cites favorably the Ninth Circuit's recent *Randolph* decision, which discussed in its reasoning how disclosure allowed a reporter to discover that an effort promoting the passage of Proposition 188 in 1994 (that would have overturned a workplace smoking ban) was heavily financed by Big Tobacco and not—as was claimed—small businesses.[53] *Getman* provides further support: "At least by knowing who backs or opposes a given initiative, voters will have a pretty good idea of who stands to benefit from the legislation."[54] The opinion then cites Proposition 199 in 1996: disclosure revealed that the measure, alleged to assist mobile home park residents with rent, was actually backed by mobile home park owners who wanted to eliminate local rent control.[55]

However, there were no economic special interests behind Proposition 8. The ballot measure was entirely a social issue, and any interests that stood to gain economically by passage or defeat of the proposition were not the concern of voters. Although the Ninth Circuit is admirably against groups that attempt to mask their agenda by claiming to be a grassroots movement when in reality they are not, these same groups can and will find ways to obfuscate regardless of disclosure.[56] Disclosure did nothing to reveal ulterior motives in the Proposition 8 campaigns. Disclosure did reveal backing by the Church of Jesus Christ of Latter-Day Saints as well as other Christian groups both within California and out-of-state,[57] but this served no purpose in revealing a hidden agenda or deception.

Although there are undoubtedly examples of "ambiguous or misleading names" for committees in ballot proposition campaigns, if "Protect Marriage" (the lead organization for Proposition 8) and "Equality for All" (the lead organization against Proposition 8) meet this definition, then one would be hard pressed to find a committee name that does not. As in most campaigns there was heated debate in the months leading up to the passage of Proposition 8 that often sank below the level of mature discourse, but this could not (and should not) be prevented by disclosure laws.

C. The "Informational" Interest Distinguished

Judge England fails to recognize that even if the government does have an informational interest in disclosing donations for ballot measure issue advocacy, none of those interests were implicated in the Proposition 8 campaigns. The policy content of Proposition 8 was clear,[58] citizen-legislators always control their own vote, and committees were not deceptively titled.[59] While none of the prongs of the Ninth Circuit's informational interest are implicated in ballot measures like Proposition 8, the unintended consequence of disclosure—people using the information to send death threats—deters free speech.

It is interesting to note the treatment Judge England gives to the actions taken against pro-Proposition 8 donors. Judge England casually notes that "[o]nly random acts of violence directed at a very small segment of supporters of the initiative are alleged."[60] He references the Declaration of Sarah E. Troupis and quotes an e-mail she received: "If I had a gun I would have gunned you down along with each and every other supporter..."[61] But because this was an isolated incident, Judge England dismisses the gravity of the situation. He rightly notes that other hardships, such as a boycott of one's business, are rightful exercises of the First Amendment rights of others.[62] This consideration and the apparent impregnability of the informational interest allow Judge England to gloss over an important argument by the plaintiffs: not only boycotts, but in some instances death threats, were made possible only because of governmental disclosure.[63]

The oppression faced by many who did not join the *ProtectMarriage* suit is well-documented.[64] Although there have been threats, it does not appear that anyone has actually acted on these threats. But this does not overcome the immeasurable impact of the message sent by some proponents of gay marriage: if you oppose gay marriage, *be afraid*. In light of disclosure serving no governmental interest in the case of Proposition 8, even the slightest impact on free speech through disclosure is enough cause to re-examine ballot measure disclosure.

II. Disclosure in Issue Advocacy: A Narrower, Economic Interest

The interests articulated by the Ninth Circuit in *Getman* and reiterated in *ProtectMarriage* were not implicated in the Proposition 8 fallout. The question is, then, whether Proposition

8 and pure social-issue ballot measures should be carved out of the informational interest or whether *Buckley* leaves no room at all for disclosure in ballot measure advocacy.

A. Buckley's *Informational Interest*

In *ProtectMarriage*, Judge England acknowledges that *Buckley* only discussed elections involving candidates.[65] Although the Ninth Circuit ruled that ballot measure disclosure can fit into the first informational interest discussed in *Buckley*, this is, at best, a stretch:

> First, disclosure provides the electorate with information "as to where political campaign money comes from and how it is spent by the candidate" in order to aid the voters in evaluating those who seek federal office. It allows voters to place each candidate in the political spectrum more precisely than is often possible solely on the basis of party labels and campaign speeches. The sources of a candidate's financial support also alert the voter to the interests to which a candidate is most likely to be responsive and thus facilitate predictions of future performance in office.[66]

Because this interest is separate from the interest of preventing corruption or the appearance of corruption,[67] clear-cut statements by the Supreme Court in cases such as *Bellotti* ("The risk of corruption perceived in cases involving candidate elections... simply is not present in a popular vote on a public issue."[68]) do not foreclose disclosure for issue advocacy.

The best support for the Ninth Circuit's position comes from this sentence: "[Disclosure] allows voters to place each candidate in the *political spectrum* more precisely than is often possible solely on the basis of party labels and campaign speeches."[69] This check on honesty, for the Ninth Circuit, easily translates to ballot measures: "the same considerations apply just as forcefully, if not more so, for voter-decided ballot measures."[70] As discussed in the previous section, disclosure allegedly serves to provide the electorate with a better understanding of the policy in a ballot measure by showing who supports it.[71] But while a candidate or an officeholder can make promises to act one way and then act in an entirely different manner, a law is a black letter document. Perhaps it will be enforced in an unexpected manner, but this has more bearing on candidacy disclosure than on a ballot measure. "California's... need to educate its electorate"[72] is high-minded, but it amounts to *protecting* its electorate from the First Amendment, "which was designed 'to secure "the widest possible dissemination of information from diverse and antagonistic sources,"'" and "'to assure unfettered interchange of ideas for the bringing about of political and social changes desired by the people.'"[73]

Although the informational interest is distinct from the interest of preventing corruption or the appearance of corruption, the Ninth Circuit appears to not consider it as wholly independent: "At least by knowing who backs or opposes a given initiative, voters will have a pretty good idea of who stands to *benefit* from the legislation."[74] The Ninth Circuit does not explicitly state that this benefit must be monetary, but their support for this statement points to a solely economic interest.[75] The separate example cited in *ProtectMarriage* is also to an economic interest.[76] "Benefit" is to say, then, that supporters are not looking to vindicate a political issue for what they believe is for the good of society as a whole, but are instead seeking economic gain. In such a situation, money is not spent solely for political communication, but in search of (perhaps less corrupt) quid pro quo.[77] When this use predominates over speech, as discussed in the corruption section of *Buckley*, "the integrity of our system of representative democracy is undermined."[78]

B. Carve It Out or Can It

Economic interests were not the driving force for many—if any—donors on either side of Proposition 8: gay marriage is a social issue. Although money "'produces a cacophony of political communications through which California voters must pick out meaningful and accurate messages,'"[79] this is the objective of the First Amendment, not a problem to be solved.[80] The Ninth Circuit's extension of *Buckley*'s informational interest may have merit for disclosure in ballot measures that will primarily benefit and/or deprive different segments of the population *economically*, but as applied to the Proposition 8 fallout it serves no legitimate governmental interest.

Disclosure did not further understanding of Proposition 8, prevent confusion of "citizen-legislators," or expose large interest groups masquerading as something different.[81] Instead, disclosure provided uncivil proponents of gay marriage with the means to scare supporters of traditional marriage from supporting their view politically should the issue ever arise again in the ballot context. Given this result, the Ninth Circuit's articulation needs work. The informational interest for disclosure should be narrowly tailored to exclude predominantly social-issue ballot measures such as Proposition 8. Given that campaign finance law has already given rise to numerous vague standards that put judges in the position of "know[ing] [a violation] when [they] see it,"[82] the Proposition 8 fallout provides further evidence of the wisdom behind the Framers' use of the word "abridge" in the First Amendment.[83]

Conclusion

Given the chilling effect on the speech of pro-Proposition 8 donors and the potential for future campaigns of intimidation facilitated by disclosure laws relating to ballot propositions, the Ninth Circuit should reconsider the *Getman* precedent if the *ProtectMarriage* case ends in the same manner as Judge England's denial of a preliminary injunction. If the Ninth Circuit refuses to do so, the Supreme Court should grant certification and narrow the informational interest, perhaps going even so far as to restrict it to donations made to candidates or candidate-based elections. Advocacy surrounding ballot proposition campaigns is wholly protected by the First Amendment, which plainly states that "Congress shall make no law... abridging the freedom of speech."[84] In the context of issue advocacy, money is spent only as a tool of speech, and this speech is protected whether it is truthful or dishonest, clear or misleading. The California government's desire to have a better-informed electorate is admirable, but its disclosure law has provided a means for opposing parties to intimidate and silence opinions different from their own. At the same time, this campaign implicated none of the prongs of the Ninth Circuit's purported informational interest. In the Proposition 8 fallout the Ninth

Circuit fails to provide a governmental interest that withstands scrutiny for disclosure's "deterrent effect on the exercise of First Amendment rights [that] arises, not through direct government action, but indirectly as an unintended but inevitable result of the government's conduct in requiring disclosure."[85] In the context of Proposition 8 and other ballot measures involving purely social issues, campaign finance disclosure for issue advocacy is unconstitutional.

Endnotes

1 "Thirty states have now passed bans on same-sex marriage, raising serious questions about the country's support for the rights of gays and lesbians—something that Obama has advocated repeatedly on the campaign trail and in some of his most famous speeches." Christi Parsons, *Gays Find Victory Falls Short; Measures Temper Joy of Obama Win*, Chi. Trib., Nov. 7, 2008, at C4.

2 John R. Lott Jr. and Bradley Smith, *Donor Disclosure Has Its Downsides*, Wall Street J., Dec. 26, 2008, at A13.

3 Brad Stone, *Disclosure, Mangified On the Web*, N.Y. Times, Feb. 8, 2009, at A3.

4 "To implement the Legislature's intent, the Secretary of State... shall do all of the following: ... (d) Make all the data filed available on the Internet in an easily understood format that provides the greatest public access. The data shall be made available free of charge and as soon as possible after receipt.... The data made available on the Internet *shall not contain the street name and building number of the persons*... listed on the electronically filed forms or any bank account number required to be disclosed pursuant to this title." Cal. Code Regs. tit. 2, § 18550 (2006) (emphasis added); *see also* Cal Gov't Code § 84602 (2006).

5 eightmaps, http://www.eightmaps.com (last accessed Apr. 27, 2009).

6 "The site takes the names and ZIP codes of people who donated to the ballot measure... and overlays the data on a Google map. Visitors can see markers indicating a contributor's name, approximate location, amount donated and, if the donor listed it, employer. That is often enough information for interested parties to find the rest–like an e-mail or home address." Stone, *supra* note 3.

7 ProtectMarriage.com v. Bowen, 599 F. Supp. 2d 1197, 1199 (E.D. Cal. 2009).

8 *Id.* at 1226.

9 *See generally* Buckley v. Valeo, 424 U.S. 1 (1976).

10 *ProtectMarriage.com*, 599 F. Supp. 2d at 1199.

11 *See id.* at 1201–04.

12 *Id.*

13 *Id.* at 1204.

14 *Id.* at 1226.

15 *Id.* at 1207.

16 *Id.*

17 *Id.*

18 California Pro-Life Council, Inc. v. Getman ("*Getman II*"), No. 00-1968, slip op. (E.D.Cal. Feb. 22, 2005).

19 *Getman II*, No. 00-1968, slip op. at 15.

20 435 U.S. 765 (1978).

21 454 U.S. 290 (1981).

22 525 U.S. 182 (1999).

23 *See* California Pro-Life Council v. Getman, 328 F.3d 1088, 1102–04 (9th Cir. 2003).

24 *Bellotti*, 435 U.S. at 795.

25 *Id.* at 791–92.

26 *Id.* at 792 n.32.

27 *Citizens Against Rent Control*, 454 U.S. at 299.

28 *Id.* at 299–300 (emphasis added).

29 *American Constitutional Law Foundation*, 525 U.S. at 202–03.

30 *Id.* at 203. These interests are distinct: Like lobbying, paying people to collect signatures to have a piece of legislation put on a ballot is to spend money as much (if not more so) as currency than as a means of political speech. The Court cites *Bellotti* and reiterates that "ballot initiatives do not involve the risk of '*quid pro quo*' corruption" during the actual balloting stage, but does not make this distinction, which would be wise. *Id.* (citing *Bellotti*, 435 U.S. at 790). *See infra* Part II.B.

31 "First, disclosure provides the electorate with information 'as to where political campaign money comes from and how it is spent by the candidate' in order to aid the voters in evaluating those who seek federal office.... Second, disclosure requirements deter actual corruption and avoid the appearance of corruption by exposing large contributions and expenditures to the light of publicity.... Third... recordkeeping, reporting, and disclosure requirements are an essential means of gathering the data necessary to detect violations of... contribution limitations." *Buckley*, 424 U.S. at 66–68, (citing H.R. Rep. No. 92-564, p. 4 (1971)).

32 *ProtectMarriage.com*, 599 F. Supp. 2d at 1207.

33 *See supra* note 31. *But see infra* part II.B.

34 *ProtectMarriage.com*, 599 F. Supp. 2d at 1208 (citing *Getman*, 328 F.3d at 1105).

35 *ProtectMarriage.com*, 599 F. Supp. 2d at 1208 (citing *Getman II*, No. 00-198, slip op. at 17).

36 Proposition 8 read as follows:

> This initiative measure is submitted to the people in accordance with the provisions of Article II, Section 8, of the California Constitution. This initiative measure expressly amends the California Constitution by adding a section thereto; therefore, new provisions proposed to be added are printed in *italic type* to indicate that they are new.
>
> SECTION 1. Title
>
> This measure shall be known and may be cited as the "California Marriage Protection Act."
>
> SECTION 2. Section 7.5 is added to Article I of the California Constitution, to read:
>
> *SEC. 7.5. Only marriage between a man and a woman is valid or recognized in California.*

Available at http://voterguide.sos.ca.gov/past/2008/general/text-proposed-laws/text-of-proposed-laws.pdf#prop8.

37 183 P.3d 384 (2008).

38 *Compare* "Hello, I'm No On Prop 8," available at http://www.youtube.com/watch?v=b9T7ux8M4Go, *with* "Yes on Proposition 8 (Prop 8) protect marriage," available at http://www.youtube.com/watch?v=Y4nqtDrJI7A.

39 An example of this arose in the Proposition 8 campaign: the Courage Campaign Issues Committee released an advertisement claiming that members of the Mormon Church donated more than $20 million to the passage of Proposition 8. In the commercial, two Mormons invade a married lesbian couple's home, confiscate their wedding rings and tear up their marriage license. The commercial ends with the Mormons leaving the home and one asking, "What should we ban next?" "'Home Invasion': Vote No on Prop 8", available at http://www.youtube.com/watch?v=q28UwAyzUkE.

40 *ProtectMarriage.com*, 599 F. Supp. 2d at 1208 (citing *Getman*, 328 F.3d at 1106).

41 347 U.S. 612 (1954).

42 *Getman*, 328 F.3d at 1106.

43 *Harriss*, 347 U.S. at 625.

44 *Buckley*, 424 U.S. at 67, 77, 78.

45 *Id.* at 67.

46 *See supra* notes 35–39 and accompanying text.

47 *ProtectMarriage.com*, 599 F. Supp. 2d at 1209 (citing *Getman*, 328 F.3d at 1106).

48 I use the term "prong" loosely: neither of the *Getman* opinions, *Randolph* or *ProtectMarriage* attempt to separate these components of the informational interest. Each of these prongs overlaps and ties into the overall alleged interest of disclosing who's behind a message.

49 *ProtectMarriage.com*, 599 F. Supp. 2d at 1209 (citing *Getman*, 328 F.3d at 1105–06).

50 *Getman II*, No. 00-198, slip op. at 17.

51 A good example is Proposition 7, which was on the same ballot as Proposition 8. *See* Proposition 7 – Title and Summary – Voter Information Guide 2008, http://voterguide.sos.ca.gov/past/2008/general/title-sum/prop7-title-sum.htm (last visited Apr. 27, 2009).

52 *See supra* note 36.

53 *ProtectMarriage.com*, 599 F. Supp. 2d at 1209 n. 5 (citing *Randolph*, 507 F.3d at 1179 n. 8).

54 *Getman*, 328 F.3d at 1106.

55 *Id.* at 1106 n. 24.

56 Even if an organization does not obfuscate, it may have to spend a lot of time and money proving it. *See* Tony Semerad, *Leaked Memos: Gay Rights Group Make New Charges Over LDS Prop 8 Role*, Salt Lake Trib., Mar. 19, 2009 ("In new charges filed Thursday with the California Fair Political Practices Commission, the Los Angeles-based Californians Against Hate accuses the church of creating the National Organization for Marriage in California as early as summer 2007 as a front group for its agenda, while failing to report the costs as required by California law.").

57 "As many as 1,025 people and businesses in Utah donated $3.8 million to Proposition 8 efforts, with 70 percent going to campaigns supporting the measure." *Id.*

58 *See supra* note 36.

59 One could argue that it was only because the laws were in place that these did not occur. In such an instance, I re-assert my argument that in non-economically based ballot measures, disclosure provides, at best, material to obfuscate the issue and, at worst, to oppress donors.

60 *Id.* at 1217.

61 *ProtectMarriage.com*, 599 F. Sup. 2d at 1200; *see also* Second Amended Complaint at 35, ProtectMarriage.com v. Bowen, 599 F. Supp. 2d 1197 (E.D. Cal. Jan. 22, 2009).

62 *Id.* at 1218–20.

63 Contrast the Proposition 8 fallout with part of the boycott at issue in *NAACP v. Claiborne Hardware Co.*, 458 U.S. 886, 903–04 (1982):

> One form of "discipline" of black persons who violated the boycott appears to have been employed with some regularity. Individuals stood outside of boycotted stores and identified those who traded with the merchants. Some of these "store watchers" were members of a group known as the "Black Hats" or the "Deacons." The names of persons who violated the boycott were read at meetings of the Claiborne County NAACP and published in a mimeographed paper entitled the "Black Times." As stated by the chancellor, those persons "were branded as traitors to the black cause, called demeaning names, and socially ostracized for merely trading with whites."

It's one thing to engage in such tough tactics; it's quite another when the government assists.

64 *See, e.g.*, Stone, *supra* note 3; Drake Bennet, *Time for a Muzzle—The Online World of Lies and Rumor Grows Ever More Vicious. Is it Time to Rethink Free Speech?*, Boston Globe, Feb. 15, 2009, at C1 ("A number of... Proposition 8 supporters have since reported threatening e-mails and phone calls."); Eric Fonnesbeck, *Mormon Church Unfairly Attacked During Prop. 8 Campaign*, Contra Costa Times (California), Feb. 3, 2009 ("Numerous physical attacks on members of the Mormon church have been reported in California and many other areas."). *But see* Thomas Elias, *Prop. 8 Backers Seeking the End of Openness in Politics*, San Gabriel Valley Trib. (California), Feb. 13, 2009 ("[A]ll that's been reliably documented are a few examples of gays and their supporters staying away from businesses whose owners were Proposition 8 donors. No physical harm whatsoever.").

65 *See supra* note 32.

66 *Buckley*, 424 U.S. at 66–67 (citation omitted).

67 *See supra* note 31.

68 *Bellotti*, 435 U.S. at 790.

69 *Buckley*, 424 U.S. at 67 (emphasis added).

70 *Getman*, 328 F.3d at 1105.

71 *See supra* Part I.B.

72 *ProtectMarriage.com*, 599 F. Supp. 2d at 1219.

73 *Buckley*, 424 U.S. at 48–49 (citing New York Times Co. v. Sullivan, 376 U.S. 254, 266 (1964) (citing Associated Press v. United States, 326 U.S. 1, 20 (1945))).

74 *Getman*, 328 F.3d at 1106 (emphasis added).

75 *See id.* at 1106 n.24.

76 *ProtectMarriage.com*, 599 F. Supp. 2d at 1209 n.5.

77 "[Latin 'something for something'] An action or thing that is exchanged for another action or thing of more or less equal value; a substitute." Black's Law Dictionary 1282 (8th ed. 2004).

78 *Buckley*, 424 U.S. at 26–27.

79 *ProtectMarriage.com*, 599 F. Supp. 2d at 1208 (citing *Getman*, 328 F.3d at 1105).

80 *See supra* note 73 and accompanying text.

81 *See supra* Part I.B–C.

82 This is a paraphrase of Justice Stewart's short concurrence in *Jacobellis v. State of Ohio*, 378 U.S. 184, 197 (Stewart, J. concurring) (1964), wherein he states

> I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description (hard-core pornography); and perhaps I could never succeed in intelligibly doing so. But I know it when I see it, and the motion picture involved in this case is not that.

83 *See* U.S. Const. amend. I; Black's Law Dictionary 6 (8th ed. 2004) ("abridge, *vb.* 1. To reduce or diminish <abridge one's civil liberties>").

84 U.S. Const., amend. I. Note that this is the second time throughout this article that I quote the relevant text of the First Amendment. While *ProtectMarriage* quotes the Amendment at its outset, and notes its application to the states, the opinion which follows—like many campaign finance decisions—loses sight of the word "abridge" entirely. *ProtectMarriage.com*, 599 F. Supp. 2d at 1205.

85 *Buckley*, 424 U.S. at 65.

Attack Ballot Issue Disclosure Root and Branch:

Comment on *A Cold Breeze in California:* ProtectMarriage *Reveals the Chilling Effect of Campaign Finance Disclosure on Ballot Issue Advocacy*

*By Steve Simpson**

For years, the lower federal and many state courts have given short shrift to the First Amendment rights of those who wish to contribute money to groups that advocate the passage or defeat of ballot measures. Twenty-four states allow legislation to be passed in this manner, and in every one, the law requires groups advocating the passage or defeat of ballot measures to disclose the names, addresses, and often employers of their contributors.[1] This not only chills the participation of potential contributors, as Stephen Klein ably demonstrates; it can be an enormous burden on ballot issue groups as well.[2] Many states treat them like political committees, requiring them to file registration statements, appoint treasurers, and track and report not only contributions but also all expenditures.[3]

For the most part, lower courts have ignored these burdens on speech and association and have concluded that the same government interests that support candidate disclosure laws apply to ballot issue disclosure laws as well.[4] Admittedly, the legal landscape in the Supreme Court is not great for opponents of ballot issue disclosure laws. The Court has approved of the idea of ballot issue disclosure in dicta in three cases.[5] But neither is the law exactly bad for those asserting their First Amendment rights in this context. The Court has made clear in past cases that the interests served by candidate campaign finance laws do not apply to ballot issues;[6] it has upheld the right of anonymous speech[7] and the right of association against disclosure laws and efforts to require groups to disclose membership lists;[8] and it has noted the significant burdens that political committee regulations impose on voluntary groups.[9] By and large, the lower courts, especially those in the Ninth Circuit, have navigated around these precedents and have upheld disclosure laws in the ballot issue context as they have in the candidate context.

Stephen Klein does a yeoman's job of criticizing the latest example of poor judicial reasoning in this context in *ProtectMarriage.com v. Bowen*. He recognizes the lawyer's dilemma in these cases: how to convince the court that all disclosure laws are not created equal, and that those imposed on ballot issue committees pose a greater threat to freedom of speech and are supported by a far less convincing justification than disclosure laws in the candidate context. Unfortunately, Klein's proposed solution, well-meaning though it is, will not convince courts to uphold rights to anonymous speech and association and will end up doing more harm than good.

Klein proposes a distinction between ballot measures that raise purely "social issues" and those that implicate economic interests. According to Klein, while a compelling interest in disclosure might exist in the latter case, there is no such interest where purely "social" issues are concerned. The reason, as Klein sees it, is that groups with a social agenda, unlike those with economic interests at stake, have no pecuniary motives and thus no incentive to hide their agendas.

If this sounds a bit circular, that's because it is. Certainly, many groups and individuals have an economic stake in the outcome of ballot issues, but it is not clear why they have any greater or lesser reason to hide their identities or have "hidden" agendas than groups with a social agenda. Would it not benefit a campaign against gay marriage to cast itself as a grassroots campaign rather than one backed and funded by the "Religious Right"? Certainly no less so than it would benefit a campaign against smoking bans to cast it as one backed by small business rather than "Big Tobacco."

This circularity is not Klein's fault, however. At its root, the entire argument for disclosure in the ballot issue context is one big circular argument that begins with the premise that anyone who wishes to conceal their or their supporters' identities is doing something wrong. Many courts rely on a variant of Justice Brandeis's famous dictum "Sunlight is said to be the best of disinfectants."[10] But what, precisely, is disclosure intended to "disinfect" in this context? According to proponents, the laws are intended to prevent people from having "hidden agendas." But this is ultimately no different from saying that we want to know who supports or opposes ballot issues simply because we want to know.

If we take the right to privacy and anonymous speech seriously—as the Supreme Court has done in past cases—then we must recognize that the "agendas" or motivations of those who wish to remain anonymous is their business, not ours. Keeping one's views private is, after all, the reason for speaking anonymously.[11] If disclosure is justified by the desire to expose "hidden agendas," then the argument for disclosure is simply that privacy and anonymity themselves are illicit, because the purpose of those rights is to keep agendas, views, motivations—whatever one wishes to call them—private.

Thus, the problem with Klein's argument is that he accepts the premise of disclosure in part, but then tries to carve out a special exemption for a certain category of speech. Again, this is understandable given the sorry state of the law on ballot issue disclosure in the Ninth Circuit. Klein is describing a strategy for an as-applied constitutional challenge, in which fine distinctions often win the day, and lawyers must take the bad precedent as it comes and do with it what they can.

But Klein's approach must ultimately fail for two reasons. First, the distinction between social and economic issues is simply untenable. Those speaking out on social issues are just as likely to have, or be seen as having, hidden agendas as those speaking out about issues that affect their pecuniary interests. And it is not at all clear how we are to define social versus economic issues. Is immigration a social or an economic issue? What about global warming and other environmental issues that affect the economic interests of virtually everyone in the nation? Moreover, Klein's approach would, in effect, create a

* *Senior Attorney, Institute for Justice*

content-based distinction within First Amendment law itself, which would be an approach akin to burning the village in order to save it.

Second, and more importantly, one cannot defeat disclosure laws by accepting them as valid at their very core, as Klein does. Disclosure laws will never be defeated unless we can convince courts that they serve no legitimate purpose in the ballot issue context. Judges have upheld disclosure laws largely because they believe, as many Americans do, that disclosure is just a good idea regardless of the context. Klein does a good job of shooting down many of the arguments that the Ninth Circuit has embraced, but he ultimately accepts the central premise of disclosure: that it is improper to hide one's identity or those of one's supporters in certain contexts. Having accepted that premise, he is left to hope that the courts will leave just a bit of privacy and anonymity for those who promise only to speak about issues in which they have no economic interests.

Admittedly, opposing disclosure in principle, even if only in the ballot issue context, is not an easy row to hoe. One often finds oneself on the side of those accused of outright deception and lying to the public about their agendas. On closer inspection, however, the alleged abuses of anonymity are either largely overblown or simply irrelevant to a proper understanding of the First Amendment.

Take what proponents of disclosure seem to view as their silver bullet—the alleged efforts of "Big Business" to hide their support of or opposition to ballot measures. The Ninth Circuit relied as evidence of the importance of disclosure on the alleged "revelation" that California Proposition 188—which would have overturned smoking bans—was financed in large part by tobacco companies, rather than small businesses as was claimed.[12] But, in fact, Prop. 188 was indeed supported by many small businesses, no doubt because they believed that smoking bans increased costs and lost them business.[13] It was also supported by tobacco companies, but that is hardly a revelation. Is there anyone in California who could not have figured out for themselves that tobacco companies oppose smoking bans and support their repeal?

Likewise, in another case, the Ninth Circuit claimed disclosure revealed that Proposition 199, which was alleged to assist mobile home park residents with rent, was really a rent control measure supported by park owners.[14] But Proposition 199 in fact did both—it sought to repeal rent control *and* it helped mobile home park residents with rent. This was crystal clear from the language of the measure itself, and it was even revealed in some of the supporters' campaign literature.[15]

The claim that advocates in these campaigns were engaged in deception is reminiscent of the claims during every campaign season that each side's opponent is "lying" by taking a different view of the issues. Thus, if small business backs a measure that is also backed by tobacco companies, according to the proponents of disclosure it is deceptive to characterize it as anything but a law that serves the interests of Big Tobacco. And if landlords don't emphasize the aspects of a measure that its opponents believe are most relevant, they are not disclosing the whole truth.

A cardinal principle of the First Amendment is that the speaker gets to choose the content of his message, not the government or the speaker's critics.[16] Debates will often be heated and contentious; at times, speakers may even make wild and unfounded claims. But outside of narrow contexts like libel law and commercial fraud, the remedy for speech you don't like—even allegedly false speech you don't like—is more speech.[17]

Those principles ought to apply with even greater force in the context of debates over ballot issues, for the simple reason that the language of a ballot issue is there for all to read and understand. Ballot issues cannot have hidden agendas. True, the proponents and opponents of a ballot issue themselves can have hidden agendas, but the motivations or agendas of speakers in the ballot issue context cannot be a reason to impose disclosure obligations on them.[18] The desire to discover the thinking behind someone's support for or opposition to a ballot issue is simply a rejection of their right to anonymity and privacy. Again, the whole point of speaking anonymously is to sever the connection between one's views on a particular topic and one's identity, as well as one's other views, motivations, and "agendas."[19] Anonymity is just another aspect of one's message that one gets to decide for oneself.[20]

Moreover, the impulse to reveal hidden agendas has no limiting principle. Why, in other words, stop with those who contribute money to ballot issue committees? It is arguably far more important to understand the possible hidden agendas of the media and the various interest groups and think tanks that are constantly cajoling members of the public to think one thing or another on important policy questions. And the disclosure of a bare contribution conveys only one's support for a particular viewpoint. If we truly wish to reveal hidden agendas and uncover information that voters might find useful, why settle for the disclosure of only the identities, addresses, and employers of contributors? Requiring them to disclose their religious, political, and other group affiliations would reveal much more about the possible agendas of the groups to which they contribute. And while we are at it, why not require everyone to disclose which way they vote on issues? Disclosure already accomplishes that for contributors to ballot issue committees anyway, and keeping a database of everyone's voting history would be a wonderful way to assess their possible agendas in future elections.

Certainly, the language of ballot issues can be complicated at times, and it is possible that some voters might be able to use contributor disclosure as a "cue" that helps them understand the issues involved. But if voters are really interested in following the recommendations of others, loads of groups and individuals—from the news media, to interest groups, to politicians, to scholars—stand ready during each election to educate voters about all aspects of the measures on the ballot.[21]

Ultimately, the argument for disclosure boils down to the extraordinary claim that voters are unable or unwilling to understand a ballot initiative by reading the language and considering public information about it, but they can be counted on to divine its meaning by sifting through the disclosure rolls to see who has given money to the groups on each side. According to the district court in *Protectmarriage.com*, it is "naïve" to think that voters will actually take the time to understand a ballot issue,[22] so, in effect, we must force contributors to become

unwilling endorsers of the measures they support. The true path to voter education, in other words, is not to encourage voters to understand the issues themselves, but to encourage them to understand what their neighbors think.

In fact, if there is anything naïve about the prevailing view of disclosure laws, it is the view that disclosure is benign and costless. Dick Carpenter, Jeff Milyo, and John Ross illustrate in this issue of Engage the regulatory burdens of disclosure and its impact on rights to privacy.[23] Many people have expressed concerns about having their positions on issues revealed, about identity theft, and about the possible repercussions for their jobs, their businesses, their union memberships, and the like.[24] Evidence from the *Protectmarriage.com* case and a case now pending in Washington state[25] shows that they have good reason to be concerned.

Even short of being used for outright intimidation and harassment, disclosure laws are very effective political tools for each side of a campaign. Denver-based political consultant Floyd Ciruli testified in a challenge to Colorado's disclosure laws that they are regularly used by campaigns to keep track of and even gain an advantage over their opponents.[26] Robert Stern, general counsel of the California-based Center for Governmental Studies agrees. In Stern's view, many people want disclosure laws in order to be able to keep track of the activities of politically unpopular groups.[27]

This is no doubt true. As the debates over health care have shown, it is always more effective to characterize one's opponent as a mouthpiece for big business or some other special interest. But it is not clear why the state has a compelling interest in arming campaigns with the ability to use each side's contributors as a weapon in this battle.

In *McIntyre v. Ohio Bd. of Elections*, the Supreme Court struck down a state law requiring the disclosure of the authors of political writings, holding that the law violated the right to anonymous speech. In rejecting the claim that disclosure was necessary to allow the public to evaluate the message, the Court stated,

> Of course, the identity of the source is helpful in evaluating ideas. But the best test of truth is the power of the thought to get itself accepted in the competition of the market.... People are intelligent enough to evaluate the source of an anonymous writing. They can see it is anonymous. They know it is anonymous. They can evaluate its anonymity along with its message, as long as they are permitted, as they must be, to read that message. And then, once they have done so, it is for them to decide what is responsible, what is valuable, and what is truth.[28]

This very common-sense point will likely not shake the faith of disclosure's most ardent supporters. But convincing the rest of the public and the courts to think twice about disclosure laws will take more than fine distinctions among types of political speech. Stephen Klein has done a good job advancing some clear thinking in this context, but to defeat the arguments for disclosure once and for all, opponents will need to attack disclosure root and branch.

Endnotes

1 Dick M. Carpenter, III, Institute for Justice, *Disclosure Costs: Unintended Consequences of Campaign Finance Reform* 1 (March 2007).

2 Carpenter, *supra* note 1, at 1-2; Jeffrey Milyo, Institute for Justice, *Campaign Finance Red Tape: Strangling Free Speech & Political Debate* (October 2007).

3 Dick M. Carpenter, III, Jeffrey Milyo and John Ross, Institute for Justice, *Politics for Professionals Only: Ballot Measures, Campaign Finance "Reform" and the First Amendment,* 10 ENGAGE (forthcoming 2009) (manuscript at 5, on file with authors).

4 California Pro-Life Council, Inc. v. Randolph, 507 F.3d 1172 (9th Cir. 2007); California Pro-Life Council, Inc. v. Getman, 328 F.3d 1088 (9th Cir. 2003); ProtectMarriage.com v. Bowen, 599 F. Supp. 2d 1197 (E.D. Cal. 2009); Richey v. Tyson, 120 F. Supp. 2d 1298 (S.D. Ala. 2000); Volle v. Webster, 69 F. Supp. 2d 171 (D.Me. 1999).

5 Buckley v. American Constitutional Law Foundation, 525 U.S. 182, 202-03 (1999); Citizens Against Rent Control v. City of Berkeley, 454 U.S. 290, 298-300 (1981); First National Bank of Boston v. Bellotti, 435 U.S. 765, 791-92 & n. 32 (1978).

6 *Citizens Against Rent Control*, 454 U.S. at 296-97; *Bellotti*, 435 U.S. at 789-92.

7 McIntyre v. Ohio Elections Board, 514 U.S. 334, 342 (1995).

8 NAACP v. Alabama, 371 U.S. 415 (1963).

9 FEC v. Wisconsin Right to Life, Inc., 551 U.S. 449, 464 (2007); FEC v. Massachusetts Citizens for Life, 479 U.S. 238, 255-56 (1986).

10 Buckley v. Valeo, 424 U.S. 1, 67 (1976); *see also Getman*, 328 F.3d at 1106 n. 24 (stating that disclosure "prevents the wolf from masquerading in sheep's clothing").

11 *McIntyre*, 514 U.S. at 342-43.

12 *Randolph*, 507 F.3d at 1179 n. 8.

13 Alan Liddle, *Voters Shoot Down Prop. 188, Uphold Smoking Bans in California*, NATION'S RESTAURANT NEWS, November 21, 1994, at 3.

14 *Getman*, 328 F.3d at 1106 n. 24.

15 Affidavit of Stephen K. Hopcraft as Expert Witness in Support of Defendants' Motion for Summary Judgment, Exhibit C (filed in proceedings of *Randolph*, 507 F.3d 1172 (9th Cir. 2007)) (on file at Institute for Justice).

16 *Wisconsin Right to Life*, 551 U.S. at 477 n. 9; Hurley v. Irish-American Gay, Lesbian and Bisexual Group of Boston, 515 U.S. 557, 581 (1995); Cohen v. California, 403 U.S. 15, 21 (1971).

17 Texas v. Johnson, 491 U.S. 397, 419 (1989); Meese v. Keene, 481 U.S. 465, 481 (1987); Brown v. Hartlage, 456 U.S. 45, 61 (1982); Linmark Associates, Inc. v. Willingboro Township, 431 U.S. 85, 97 (1977); Whitney v. California, 274 U.S. 357, 377 (1927) ("the remedy to be applied is more speech, not enforced silence."); *see also* New York Times v. Sullivan, 376 U.S. 254 (1964).

18 *Cf.* Wisconsin Right to Life, 551 U.S. at 468 (rejecting speaker's intent as proper test of what constitutes express advocacy).

19 *McIntyre*, 514 U.S. at 342-43.

20 *Id.* at 342.

21 Carpenter, Milyo and Ross, *supra* note 3. (manuscript at 5, on file with authors).

22 ProtectMarriage.com v. Bowen, 599 F. Supp. 2d at 1209.

23 Carpenter, Milyo and Ross, *supra* note 3.

24 *Id.* (manuscript at 5, on file with authors); Carpenter, *supra* note 1, at 7.

25 Protect Marriage Washington v. Reed, No. 0:09-cv-05456-BHS (W.D. Wash. filed July 28, 2009).

26 Deposition of Floyd Ciruli at 37:19–39:1 (October 4, 2007) (filed in proceedings of Sampson v. Dennis, No. 06-cv-01858-RPM-MJW (D. Colo. 2007) (on file at the Institute for Justice).

27 Deposition of Robert Stern at 36:4–37:11 (September 26, 2007) (filed in proceedings of Sampson v. Dennis, No. 06-cv-01858-RPM-MJW (D. Colo. 2007) (on file at the Institute for Justice)).

28 *McIntyre*, 514 U.S. at 348 n. 11 (internal quotations omitted).

Rebuttal to Steve Simpson's Response to *A Cold Breeze in California: ProtectMarriage Reveals the Chilling Effect of Campaign Finance Disclosure on Ballot Issue Advocacy*

By Stephen R. Klein

I have had the opportunity to consider First Amendment associational privacy and anonymity in greater detail since writing the article appearing above in this edition of *Engage*.[1] Steve Simpson's observation that my argument takes for granted a governmental interest in ballot measure disclosure where there is plainly none is aptly put. Despite my best intentions, I treated the First Amendment in light of judicial precedent, and, using such a backwards paradigm, called for a visit to the proverbial free speech woodshed.

Nevertheless, while I agree that there is no governmental interest in ballot measure campaign disclosure, this maxim has had little effect in practice. Although the Ninth Circuit is the only Court that has described the so-called "informational interest" in detail,[2] First Amendment challenges against similar concoctions have also failed in Alabama,[3] Maine,[4] Utah,[5] and Colorado.[6] Free speech finally scored a win recently in Wisconsin,[7] and this will hopefully amount to more than but a moment of clarity. But it is up against a large body of careless precedent.

Furthermore, Simpson's assertion that "neither is the law exactly bad for those asserting their First Amendment rights in this context" seems overly optimistic. Though Simpson acknowledges that "lower courts... have navigated around [Supreme Court] precedents," he does not acknowledge that the Supreme Court itself has provided part of the map, and not merely in the *Bellotti/Citizens Against Rent Control/ ACLF* line of dicta.[8] *McConnell v. FEC* also contains ample expansions of *Buckley*, complete with implicit assertions that the government has an interest in restricting political groups from "misleading" names.[9] Even *McIntyre v. Ohio Elections Comm'n*, the quintessential case affirming the First Amendment right to anonymous speech, contains dictum that squelches anonymity in the face of campaign finance law:

> Disclosure of an expenditure and its use, without more, reveals far less information. It may be information that a person prefers to keep secret, and undoubtedly it often gives away something about the spender's political views. Nonetheless, even though money may "talk," its speech is less specific, less personal, and less provocative than a handbill—and as a result, when money supports an unpopular viewpoint it is less likely to precipitate retaliation.[10]

So, despite recent progress in First Amendment campaign finance actions, working to narrow the informational interest may be more effective (albeit far slower and more frustrating) than a root-and-branch attack.

Though Simpson correctly argues that differentiating between economic issues and social issues is unworkable in other contexts,[11] in the *ProtectMarriage* case the distinction would work. I did not argue that a group may have more or less interest in hiding their agenda if their interest is guided by economic or social principles, but rather that government only has an interest in disclosing donors who may appear to be "buying" a law that will enrich them. Again I acknowledge that this argument draws from case law rather than the First Amendment, but the argument would force the Ninth Circuit and/or the Supreme Court to confront the spurious reasoning that superimposes *Buckley* onto ballot measure disclosure and offers a solution that works in the context of *ProtectMarriage*: although there is a powerful gun lobby, tobacco lobby, and other lobbies in the United States looking to protect their industries, the "marriage lobby" is not out to protect marriage parlors or religious service fees. The Proposition 8 campaign was unquestionably driven by morality and morality alone, a social issue distinguishable from any hint of money used as quid pro quo. Simpson argues that this solution would do more harm than good in the long run, but it would vindicate the rights of those who contributed to Proposition 8 and would force courts to at least consider disclosure in future cases rather than sweep aside all arguments with faithful recitations of *Getman*.[12]

Simpson illustrates numerous other social issues, such as gun control, that have economic components, and correctly argues that groups advocating positions in related ballot measures should have no less First Amendment protection than the Proposition 8 donors. But by narrowing the "informational interest" for disclosure with the distinction of social and economic issues, the interest will become a far easier target in future challenges by such organizations. In other hotly contested areas of campaign finance law, such as the "functional equivalent of express advocacy," it is only through a series of as-applied challenges that judges have come to recognize the burdens the law places on political speech, and to finally "err on the side of protecting political speech rather than suppressing it."[13]

The First Amendment's victory over ballot measure disclosure in Wisconsin will, I hope, become a pattern, but, in the meantime, advocates of free speech should—in addition to root-and-branch arguments—work to clarify shoddy precedent to the greatest extent possible. This can lead to exposing the oppressive nature of campaign finance laws. Either way, Simpson and I share the ultimate end of freeing citizenry to engage in constitutionally guaranteed political speech.

Endnotes

1 The author recently co-authored, with attorney Benjamin Barr, an *amicus curiae* brief for certiorari by the United States Supreme Court in the case *Independence Institute v. Coffman*, 209 P.3d 1130 (Colo. Ct. App. 2009). The brief is available at http://www.campaignfreedom.org/news_center/detail/ccp-files-friend-of-the-court-brief-in-colorado-free-speech-case. Institute for Justice is lead counsel on the case.

2 California Pro-Life Council, Inc. v. Getman, 328 F.3d 1088, 1100–04

(9th Cir. 2003).

3 Richey v. Tyson, 120 F. Supp. 2d 1298, 1310 (S.D. Ala. 2000).

4 Volle v. Webster, 69 F. Supp. 2d 171, 173–74 (D. Maine 1999).

5 Nat'l Right to Work Legal Defense Fund and Educ. Found., Inc. v. Herbert, 581 F. Supp. 2d 1132, 1145 (D. Utah 2008).

6 *Coffman*, 209 P.3d at 1142–43.

7 Swaffer v. Cane, 610 F. Supp. 2d 962, 968 (E.D. Wis. 2009) ("The government's interest in keeping the public informed of where and how the teetotalers of Whitewater are spending their money to rally support against a liquor referendum is not commensurate with the government's interest in knowing which candidates for public office those same teetotalers financially support.").

8 *See* Simpson Response, *supra*, 74 n.5 and accompanying text.

9 McConnell v. FEC, 540 U.S. 93, 128 (2003) (citations omitted) ("Because FECA's disclosure requirements did not apply to so-called issue ads, sponsors of such ads often used misleading names to conceal their identity. 'Citizens for Better Medicare,' for instance, was not a grassroots organization of citizens, as its name might suggest, but was instead a platform for an association of drug manufacturers. And 'Republicans for Clean Air,' which ran ads in the 2000 Republican Presidential primary, was actually an organization consisting of just two individuals- brothers who together spent $25 million on ads supporting their favored candidate."); *id.* at 196 ("'Curiously, Plaintiffs want to preserve the ability to run these advertisements while hiding behind dubious and misleading names like: "The Coalition-Americans Working for Real Change" (funded by business organizations opposed to organized labor), "Citizens for Better Medicare" (funded by the pharmaceutical industry), "Republicans for Clean Air" (funded by brothers Charles and Sam Wyly).... Given these tactics, Plaintiffs never satisfactorily answer the question of how "uninhibited, robust, and wide-open" speech can occur when organizations hide themselves from the scrutiny of the voting public.... Plaintiffs' argument for striking down BCRA's disclosure provisions does not reinforce the precious First Amendment values that Plaintiffs argue are trampled by BCRA, but ignores the competing First Amendment interests of individual citizens seeking to make informed choices in the political marketplace.'") (citing McConnell v. FEC, 251 F. Supp. 2d 176, 237 (D.D.C. 2003) (internal citations omitted)).

10 McIntyre v. Ohio Elections Comm'n, 514 U.S. 334, 355 (1995).

11 *See* Simpson Response, *supra*, at 74.

12 It is still mind-boggling that after the argument was conceded by counsel, Judge England would go to the trouble of reciting *Getman* and, despite his effort, not realize that not one alleged governmental interest was implicated in the case at hand. *See* ProtectMarriage.com v. Bowen, 599 F.Supp.2d 1197, 1207–11 (E.D. Cal. 2009).

13 FEC v. Wisconsin Right to Life, Inc., 551 U.S. 449, 457 (2007).

Politics for Professionals Only: Ballot Measures, Campaign Finance "Reform," and the First Amendment

*By Dick M. Carpenter, Ph.D., Jeffrey Milyo, Ph.D., and John K. Ross**

When Scott Eckern donated money to an election committee, little did he know that it would cost him his job of 25 years. Eckern had worked successfully as the artistic director of the California Musical Theatre. In the heat of California's Proposition 8 battle over defining marriage, Eckern donated $1,000 to the "Yes on 8" committee. Consistent with California's campaign finance disclosure laws, Eckern's name, occupation, and employer were posted on a state website. Opponents of Proposition 8 saw his name on the list and called for a boycott of the Theatre, causing a public furor. "To protect the organization and to help the healing in the local theater-going and creative community," Eckern resigned his position.[1]

Eckern's case was not an isolated one as both sides of the heated issue used campaign-finance disclosures to intimidate opponents. A Proposition 8 opposition group used disclosure lists to publish a so-called "Dishonor Roll" of donors to the Yes on 8 campaign.[2] Geoff Kors, a member of the No on 8 campaign committee, said the Yes on 8 campaign sent "blackmail" letters to opponents of the measure demanding equal contributions.[3] According to a lawsuit filed on behalf of Yes on 8, those who gave money to support the ballot measure received menacing phone calls (including death threats), e-mails, and postcards. Another donor had a window broken, one had a flier distributed around his hometown calling him a bigot, and others received envelopes containing suspicious white powder.[4]

The lawsuit was filed to challenge the constitutionality of California laws that require campaign contributors to disclose personal information. But California is not alone. Citizens in all 24 states that allow ballot issues (also called initiatives or propositions) face the same scenario. That is, when citizens join together to speak out on issues (even by something as simple as donating to a campaign), they run the risk of finding themselves mired in the murky world of campaign finance regulation.

Campaign finance regulation was originally meant to prevent corruption and the appearance of corruption in politics—specifically, the trading of political favors for campaign contributions. But in the context of ballot measures, where the people are voting on proposed laws directly, there is no one to corrupt. Nonetheless, the kinds of campaign finance regulations intended for politicians have seeped into the realm of ballot issues.

** Dick Carpenter is Director of Strategic Research at the Institute for Justice and an associate professor at the University of Colorado, Colorado Springs.*

Jeffrey Milyo is the Middlebush Professor of Social Science at the University of Missouri, the Hanna Family Scholar in the Center for Applied Economics at the University of Kansas School of Business, and a senior fellow at the Cato Institute.

John Ross is the Research Associate at the Institute for Justice.

These regulations force citizens to register with the government and disclose the names, addresses, and even employers of supporters, simply because they choose to exercise their constitutional rights of association and free speech by joining together and speaking out about political issues. Those disclosures are then made public, typically on state websites.

The forms required to comply with campaign finance regulations are at least as complicated as tax forms, and the sanctions for even small clerical errors can sideline a group and expose them to legal liability in the midst of a campaign. Ostensibly, such strict requirements keep special interests at bay. In reality, however, these rules keep politics an insider's game.

I. Parker North

In 2006, several residents of Parker North, Colorado, a suburb outside Denver, caught wind of a possible ballot issue being put to their neighbors that would have annexed their neighborhood into the nearby town of Parker. These residents opposed the plan on several grounds and did what any citizen in a democracy should do—they spoke out about it. They met informally with other neighbors, wrote letters to the editor, set up an online forum for discussion and debate with proponents, distributed flyers, and put up signs—the essence of grassroots activism.[5]

The two main proponents of annexation—a lawyer and another resident—soon realized that this civic participation was hurting their efforts. So they sued six prominent critics of annexation for violating campaign finance laws. The complaint threatened "investigation, scrutinization," and fines for anyone involved with the matter. Further, the proponents attempted to subpoena the names and addresses of "all persons [who] sold, gifted, or transferred signs, banners or any campaign information" and "all communications amongst the [neighbors] or anyone else" concerning the annexation.

Under Colorado law, when two or more people join together and spend more than $200 on political activities related to a ballot issue, they must register as an "issue committee." Furthermore, any Colorado citizen may bring a private suit to enforce campaign finance laws. While intended as a way to enhance enforcement of campaign finance laws, this merely grants one side of a political issue a method to use government authority to bully political opponents.

Unaware of these laws and unconvinced that their informal and ad hoc activism warranted state attention, the opponents of annexation had failed to comply with the laws' numerous Byzantine mandates. Among them: designate formal officers—they had been meeting on porches and in kitchens and pitching in where they could—open a separate bank account, itemize all monetary and non-monetary transactions of more than $20, and provide the names and employers of supporters to the state for publication on a web site maintained by the Colorado Secretary of State.

Faced with the possibility of fines, the neighbors quickly moved to comply, only to find that the state's 92-page handbook

was confusing and not even authoritative—it was to be used for "reference and training purposes only." Worse, questions directed to the very state employees in charge of compliance did not provide answers to basic questions.[6] For instance, Becky Cornwell, who became the registered agent for the "issue committee," discovered that another resident, who was not involved with the already loosely-affiliated neighbors named in the lawsuit, had begun to sell anti-annexation t-shirts. Did she have to track his activities and file them as contributions or expenditures, or not at all? The neighbors were forced to face these problems, all in order to disclose a dozen or so donors who contributed $2,240 in monetary and non-monetary contributions over a twelve-month period—and most of that was for legal advice.

Even the Secretary of State describes the host of regulations it oversees as "often complex and unclear." State employees could not answer Becky's questions and had one piece of advice: Hire a lawyer.

Once registered, the neighbors asked the proponents to drop their suit but were refused unless they abandoned their advocacy and, among other things, "removed from sight all signs and campaign materials." Unwilling to cave to such opportunistic bullying after months of effort, the neighbors did indeed hire a lawyer to defend themselves. Soon after, the Institute for Justice and the Parker North Six filed a separate lawsuit in federal court challenging the regulations as a violation of the First Amendment.

In February 2007, annexation was soundly rejected 351-21 at the ballot box, but the experience of the Parker North neighbors showed how the very act of being dragged into court deters the political activity the First Amendment was enacted to protect: "[W]e had no clue about these laws or how to navigate through the process," said Karen Sampson, a neighbor named in the complaint. "We spent more defending ourselves than fighting annexation."[7]

Further, regulators possess broad discretion to punish activists for transgressions as harmless as clerical errors. In California, for instance, a grassroots group consisting mainly of two activists—Steve Cicero and Russ Howard—and calling themselves Californians Against Corruption started a petition to recall the president of the state Senate. They were hit with $808,000 in fines for failing to disclose the occupations of 93 donors. State regulators explicitly cited as an aggravating factor the fact that Howard had criticized campaign finance laws to a journalist, saying, "The little guy can't [participate] in politics without running afoul of technical violations."

Indeed, prosecutors cared more about the paperwork than any disclosure. According to Howard, "They have a copy of every check we ever received, and the vast majority have those addresses on the check." "My life should be ruined because I didn't fill out the proper forms in triplicate?" In 2003, a full nine years after the ultimately unsuccessful petition, the California Supreme Court refused to hear their appeals and the two are on the hook for almost $1.1 million plus interest—more than they could ever hope to pay off.

Howard, a former stockbroker, lost his job because of his work on the petition. "[T]he law sets up prerequisites for the exercise of First Amendment rights that are extremely complicated and enforced as selectively as they want," he said.

> I'd have to be an attorney or an accountant to be able to wade my way through [the law]. If I had had to stay up until five in the morning filling out those forms, there never would have been a… recall. We were a grassroots organization. We all had other jobs. We're not like the big parties. [Complying properly] would have taken a huge percentage of our resources.[8]

II. Campaign Finance Reformers

Campaign finance "reformers" argue that any abuse of the system, if indeed they acknowledge it, is insubstantial compared to the benefits society gleans from mandatory disclosure. These benefits allegedly consist of a better-informed electorate and an institutional safeguard against corruption.[9]

It is not clear how the names and addresses of people who donate $20 to an issue campaign are particularly valuable information to anyone. Nor, as we shall see, is it clear that many citizens make any use of this information.

Likewise, "reformers" fail to show how a ballot measure can be corrupted. Unlike a politician, an unchanging text cannot exchange favors. To see donations to an issue campaign as corruption, one must change the definition of "corruption" to mean any attempt to influence the outcome of an election. (And, in particular, to influence that outcome in a direction that one does not like.)

In a free country, citizens appeal to one another in hopes of enacting an array of often mutually exclusive policies. This competition is an essential part of a vibrant, healthy society, and certainly some participants will be more influential than others. Thus, the claims that disclosure unmasks "undue influence" of the political process emanating from campaign finance supporters ring hollow. After all, who is to decide what constitutes "undue influence?"

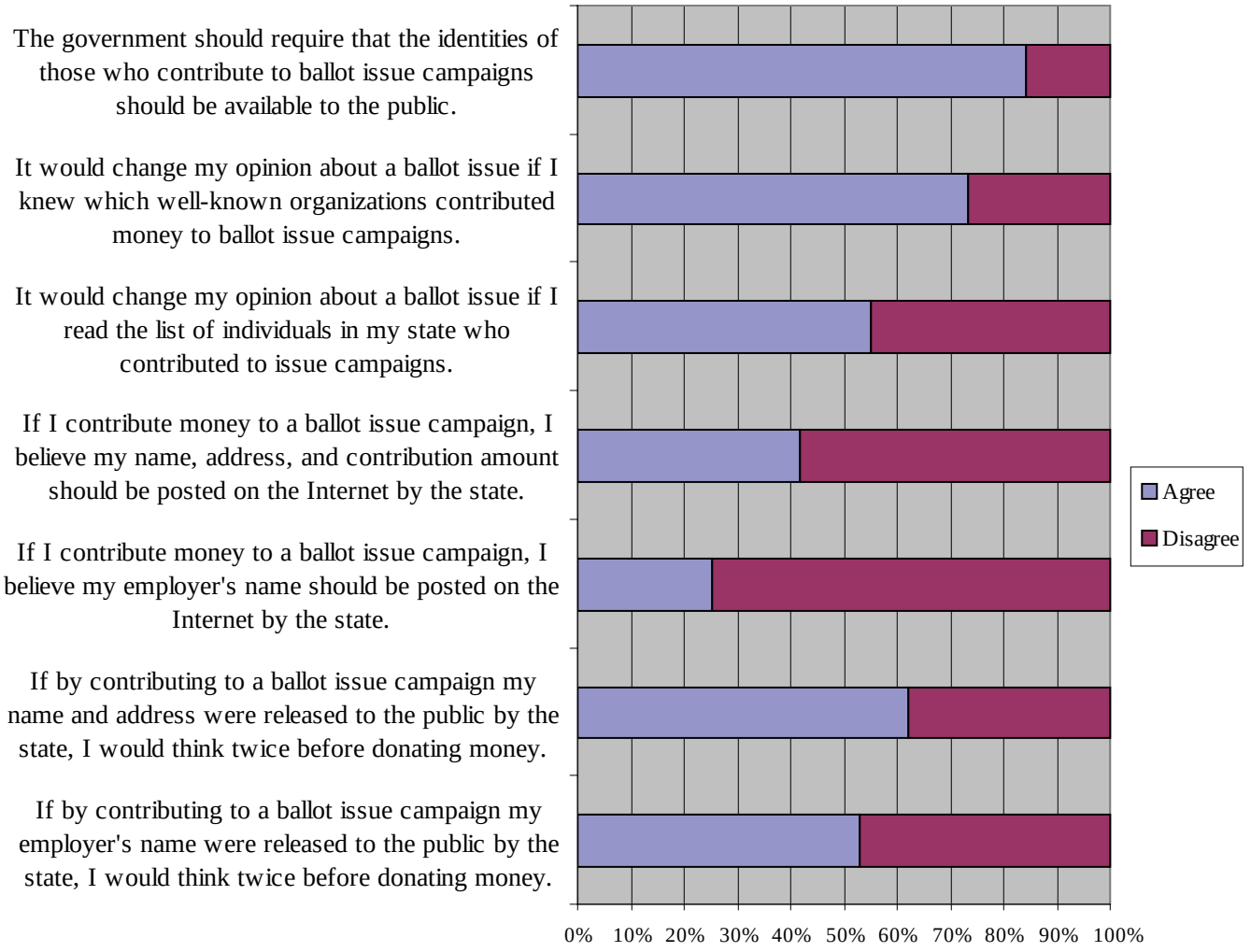
Mandatory disclosure simply allows established and moneyed interests—with professional political experts, accountants, and lawyers who will not get tripped up by reporting requirements—to continue exerting their influence while silencing small ad-hoc groups with little experience running campaigns. It is curious then, that reformers insist that campaign finance regulations *prevent* entrenched interests from subverting the public's will.

III. Disclosure Costs

In spite of the hazards they pose to ordinary citizens, campaign finance rules do enjoy broad support, at least in the abstract. In the months before the November 2006 elections, we (specifically Carpenter) polled voters in six states where citizens vote on ballot issues and found more than 82 percent approve of mandatory disclosure.[10] Moreover, more than 70 percent say information yielded from disclosure about organizations is influential and valuable and more than 50 percent likewise said the same thing about individuals.

However, once we asked voters about whether their own political activities should trigger disclosure, the tables turned.

Figure 1: Support for Mandatory Disclosure for Others, for Respondents



Some 56 percent disagreed that their own information should be publicized—and that grew to 71 percent when that disclosure included their employer. When asked why they did not want their information released, 54 percent cited a desire to remain anonymous. Others expressed concern for their personal safety and a fear of repercussions (particularly when employer information is involved) and harassment.

These fears are not unfounded: the NAACP famously fought attempts to turn its donor lists over to the government for this very reason.[11] Supporters of Californians Against Corruption, who had their home addresses disclosed on a state website, testified to receiving threatening calls and letters, and Russ Howard says some received swastikas in the mail. A local newspaper columnist even printed donors' names and contribution amounts—a step not taken for other ballot measures.[12]

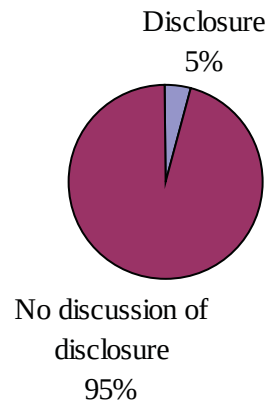
Others we polled said requiring disclosure violated their right to a private vote—as revealing their donations to the public would reveal their electoral choice—or worried about identity theft. These answers clearly indicate that mandatory disclosure foments reluctance to speak or associate in the political arena—where such rights are arguably the most important.

Nevertheless, campaign finance supporters assert that the information garnered from mandatory disclosure is important to the decision-making capabilities of an informed electorate. But the vast majority (nearly two-thirds) of respondents in our poll did not know where to access that information and never actively seek it out. Indeed, about 75 percent of those polled could not name any specific funders of or contributors to issue campaigns in their state.

Journalists and watchdog groups often protest that *they* use the information gleaned from disclosure, which is useful, if not necessary, to their investigations.[13] We examined that claim by analyzing news stories, editorials, letters to the editor, state-produced information, reports from think tanks and non-profits, and campaign-generated materials available to voters about the issues on the November 2006 ballot in Colorado.

We found that Colorado voters enjoy a wealth of information and opinions from a broad range of sources on ballot measures—our sample *undercounted* these points of information, as hardcopy campaign materials, paid

Figure 2: News Stories, Opinion Pieces, Campaign-Materials and State-Produced Information that Utilized Disclosure-Related Data



advertisements, and electronic media, such as radio and television reports and commentary, were not available.

Of the 1,078 points of information we found, only 4.8 percent included any discussion of campaign finance disclosure-related data. The other 95 percent of sources focused on the ballot issues, predicting the effects of the issues' passage or defeat, and generally discussed their merits and demerits without referring to any information generated by disclosure. Voters seeking information free from opinion could easily find it—our sample included many news stories and state ballot summaries on "what it does." Likewise, the views of proponents and opponents were clearly and numerously represented.

As mentioned above, some two-thirds of respondents report they never seek out information resulting from disclosure. It seems the same is true for an even higher percentage of journalists writing about those issues. There is, moreover, absolutely no data confirming—or even suggesting—that this paltry coverage of disclosure-related data made an impact on voters.

IV. Red Tape

Citizens like the neighbors in Parker North who wish to do more than simply donate to a political cause face additional hardships in the name of transparency. In each of the 24 states that put ballot issues to voters, citizens who wish to join together to support or oppose an issue must register as a political committee, track expenditures, and report all contributions. Failure to follow these rules can result in substantial penalties.

Table 1: Selected Tasks for Neighbors United

	Percentage of Participants Completing Task Correctly		
Task	**California**	**Colorado**	**Missouri**
Register as political committee	25%	72%	82%
Statement declaring position on ballot issue	36%	n.a.	n.a.
Reporting initial funds on hand	44%	67%	52%
Record $2,000 check contribution	60%	72%	80%
Record Anonymous $15 cash contribution	69%	51%	77%
Record Illegal Anonymous $1,000 Contribution	2%	3%	8%
Record Non-Monetary Contribution of $8 in refreshments	30%	36%	24%
Record Non-Monetary Contribution of $40 in supplies	18%	46%	26%
Record Non-Monetary Contribution of $500 in t-shirts	0%	6%	14%
Report expenditure of $1,500 for a newspaper advertisement	49%	89%	72%
(No miscellaneous clerical errors on all tasks)	5%	6%	2%

And the rules are especially challenging for ad-hoc, amateur activists who, like the neighbors in Parker North, may not even know to abide by them. To test that hypothesis, we (specifically Milyo) placed 255 experimental subjects—mostly graduate students—in the position of the Parker North neighbors.[14] We gave them a hypothetical campaign issue and asked them to fill out the appropriate paperwork to register a ballot committee, Neighbors United, and comply with reporting requirements of three different, representative states (California, Colorado and Missouri).

Of the 255 participants, not a single one correctly completed each of the 20 tasks on the campaign finance disclosure forms. The participant with the highest score correctly completed only 80 percent of the tasks. The mean score was just 41 percent correct. Had this been a real world exercise, every single participant could have been liable for violating campaign finance laws.

The trouble started early: like the Parker North neighbors, 93 percent of participants had no idea they needed to register as a political committee to speak out in the first place. So without the explicit instructions provided, participants would have done even worse.

While reporting simple contributions proved difficult, subjects had even more trouble with non-monetary contributions—the t-shirts, posters, flyers and other supplies that are typical of a grassroots campaign. Even informed of the fair market value of the objects to be itemized—not always readily available in the real world—participants could only report a gift of $8 in refreshments correctly 30 percent of the time in California, 36 percent in Colorado, and 24 percent in Missouri. Another scenario in which a contributor spent $500 on t-shirts and then donated them to the group was the most formidable. No one in the California group reported this transaction correctly, and only 6 percent in the Colorado and 14 percent in the Missouri group succeeded.

Subjects were also directed to aggregate multiple donations from an individual donor in two separate tasks. The highest score on either task from any state was only 7 percent (California). Participants simply made minor errors in arithmetic that threw off the sum total. This illustrates how fines that are levied per violation can compound.

Participants were given the opportunity to comment in writing on the experience with the disclosure forms and instructions, and 94 of the 255 did so. Of those, 90 out of 94

Table 2: State Disclosure Laws for Ballot Issues – Minimum Thresholds That Trigger Selected Disclosure Requirements

		Contributors		
	Register as Committee	**Name and Address**	**Employer or Occupation**	**Itemize Committee Expenditures**
Alaska	$500	No minimum	$250	$100
Arizona	500	$25	25	No minimum
Arkansas	500	200	n.a.	100
California	1000	No minimum	100	100
Colorado	200	20	100	20
Florida	500	No minimum	100	No minimum
Idaho	500	50	n.a.	25
Illinois	3,000	150	500	150
Maine	1,500	50	50	No minimum
Massachusetts	No minimum	50	200	50
Michigan	500	No minimum	100	50
Mississippi	200	200	200	200
Missouri	500	100	100	100
Montana	No minimum	35	35	No minimum
Nebraska	5,000	250	n.a.	250
Nevada	No minimum	100	n.a.	100
North Dakota	No minimum	100	n.a.	100
Ohio	No minimum	No minimum	100	25
Oklahoma	500	50	50	50
Oregon	No minimum	100	100	100
South Dakota	500	100	n.a.	n.a.
Utah	750	50	50	50
Washington	No minimum	25	100	50
Wyoming	No minimum	No minimum	n.a.	No minimum

expressed frustration with the forms:

Seriously, a person needs a lawyer to do this correctly.

Worse than the IRS!

Good Lord! I would never volunteer to do this for any committee.

These forms make me feel stupid!

Another participant had this to say:

I serve as the Treasurer of a political coordination committee/ political action committee formed within the last year. Even with that limited experience I found this exercise to be complicated and mentally challenging.… The burdensome paper work and fines imposed for errors in reporting proved to be a hurdle that prevented the formation of our PAC (that is affiliated with the non-profit I work for) for a number of years. That being said, in politics it is important to know the major contributors of our elected officials and hold contributors and recipients accountable to the degree possible.

That is, even a political treasurer sympathetic to mandatory disclosure (though notably for contributions to elected officials and not ballot initiatives) failed to comply with the law.

V. Conclusion

These findings point to a serious disconnect between intentions and consequences. Rather than abetting a clean and transparent initiative process, the campaign finance laws that regulate speech about ballot issues discourage political participation. They allow political opponents to drown out speech and grant regulators an enormous amount of power to penalize transgressors—a power that is wielded selectively, if not capriciously.[15]

For grassroots activists, who are often newcomers to the realities of participating in politics, even the threat of prosecution for campaign finance violations is a daunting prospect that distracts from the task at hand. Political insiders know this—even if campaign finance proponents do not—and like the pro-annexation litigants in Parker North, too frequently abuse the law to shut down opposition.

The issue is not likely to disappear. As we found, people perceive that campaign finance regulations apply only to politicians, powerful interest groups, and the wealthy. As new grassroots efforts composed of amateur activists emerge—as they will each election season—many of them will invariably fail to comply, whether they are unaware of the law or simply could not figure out the forms. But in a country that values the First Amendment, speech about issues on the ballot should not be burdened with useless regulation and endless red tape.

Endnotes

1 Crowder, M. (2008, November 13). Theater Official Resigns. *Sacramento Bee,* p. A1.

2 Rojas, A. (2009, January 9). Prop. 8 Harassment Suit Filed. *Sacramento Bee,* p. A3.

3 *Id.*

4 ProtectMarriage.com v. Bowen, No. 09-58 (E. D. Cal. Filed Jan. 9, 2009).

5 PLAINTIFFS' MEMORANDUM OF LAW IN SUPPORT OF THEIR MOTION FOR SUMMARY JUDGEMENT at 1-19, Sampson v. Coffman, No. 06-CV-01858-RPM-MJW (D. Colo. Nov. 30, 2007).

6 Cornwell, B. (2008, April 13) Opposition to Annexation Riles Neighbors, Brings Lawsuit Designed to Shut Down Debate. *Rocky Mountain News*, p. 5NEWS.

7 Cardona, F. (2008, April 21) Judge to Rule on Free Speech; A Parker North Group That Fought Annexation Has Filed a Lawsuit. *Denver Post*, p. A13.

8 *Fair Political Practices Commission v. Californians Against Corruption*, 109 Cal. App. 4th 269 (Cal Ct. App. 2003); Doherty, B. (1996) Disclosure Flaw: The Perils of Campaign-Finance Disclosure Laws. Retrieved July 14, 2008 from http://www.reason.com/news/show/29856.html; Doherty, B. (2001) How Campaign Finance Law Hurts Participation in Politics. Retrieved July 14, 2008 from http://www.reason.com/news/show/32341.html.

9 Boehmke, F. J. (2005). Sources of Variation in the Frequency of Statewide Initiatives: The Role of Interest Group Populations. *Political Research Quarterly, 58*(4), 565-575; Broder, D. (2000). *Democracy Derailed: Initiative Campaigns and the Power of Money*. New York: Harcourt; Davis-Denny, G. (2005). Divergent Disclosure: The Value of Uniformity in State Campaign Finance Disclosure Laws. *Election Law Journal, 4*(4), 282-294; Edwards, D. (1995). A Stepping Stone to Reform: $100 Campaign Contribution Limits. *Social Policy 26*(1), 9-11; Garrett, E., & D Smith, D. (2005). Veiled Political Actors and Campaign Disclosure Laws in Direct Democracy, *Election Law Journal*, *4*(4): 295-328; Gerber, E. (1999). *The Populist Paradox: Interest Group Influence and the Promise of Direct Democracy*. Princeton, NJ: Princeton University Press; Gerber, E. R., & Lupia, A. (1999). Voter Competence in Direct Legislation Elections. In S. Elkin & K. Soltan (Eds.), *Citizen Competence* (pp. 147-160). University Park, PA: Pennsylvania State University; Hoesly, C. (2005). Reforming Direct Democracy: Lessons from Oregon. *California Law Review, 93*, 1191-1248; McBride, A. (1995). Don't Give Up on Washington. *Social Policy, 26*(1), 15-17; Meier, K. J., & Holbrook, T. M. (1992). I Seen My Opportunities and I Took 'Em: Political Corruption in the American States. *Journal of Politics, 54*(1), 135-155; Seidman, H., & Gilmore, R. (1986). *Politics, Position, and Power*. New York: Oxford University Press; Thomas, T. E. (1990). Has Business Captured the California Initiative Agenda. *California Management Review, 33*(1), 131-147; "The Campaign Finance Reform Blind Spot: Ballot Measure Disclosure," Ballot Initiative and Strategy Center, at Ballot Initiative and Strategy Center. (n. d.). Retrieved July 14, 2008 from http://www.ballot.org/vertical/Sites/%7B26C6ABED-7A22-4B17-A84A-CB72F7D15E3F%7D/uploads/%7BA8911D38-14D3-438F-AE43-B78BBADBE500%7D.PDF.

10 For further details, see Carpenter, D. M. (2009). Mandatory Disclosure for Ballot Initiative Campaigns. *Independent Review, 13*(4), 567-583; Carpenter, D. M. (2007). *Disclosure Costs: Unintended Consequences of Campaign Finance Reform.* Arlington, VA: Institute for Justice.

11 NAACP v. Alabama, 377 U.S. 288 (1964).

12 Doherty, 1996; Doherty, 2001.

13 Deposition of Fred Brown at 12: 4-13, 16:17-17:2, 17:18-19, Sampson v. Coffman, No. 06-CV-01858-RPM-MJW (D. Colo. Oct. 11, 2007) (on file with the Institute for Justice).

14 For further details, *see* Milyo, J. (2007). *Campaign Finance Red Tape: Strangling Free Speech and Political Debate.* Arlington, VA: Institute for Justice.

15 Lucas, S. S. (2000). Overly Complex and Unduly Burdensome. Retrieved December 23, 2006, from http://www.fppc.ca.gov/pdf/McPherson.pdf.

Intellectual Property

Qui-tam-osaurus, the Statutory Dinosaur: Evolution or Extinction for the Qui Tam Patent False Marking Statute?

*by Trevor K. Copeland and Laura A. Lydigsen**

Get rich quick! Sue for fun and profit! Sound like a hoax? Only time will tell. Several enterprising attorneys and other private parties are giving it a try by exploiting an arcane provision of the Patent Act known as the false marking statute.[1] In this article, we briefly examine the history of the false marking statute and qui tam laws, the recent explosion in false marking actions (including several suits based on marking with expired patent numbers), and some constitutional problems posed by these actions.

The legal basis for patents in American law is enshrined in the Constitution.[2] Patents protect the patent owner's exclusive right to make, use, sell, or offer for sale a patented product. What would prevent a dishonest purveyor of goods from falsely claiming to have a patent? Such false marking might deceive the public (including potential competitors) into believing an unpatented product was patented. In cases where the false marker used a real patent number of a competitor, the false marking could directly hurt that competitor. In 1842, Congress addressed these concerns by enlisting the help of the very public likely to be duped by such fraudulent tactics.[3] Rather than saddling the government or the patent-holder with the responsibility and expense of policing such fraud, Congress adopted a patent law qui tam statute[4] that essentially deputized any person who found false marking and empowered him to sue the wrongdoer. The motive for doing so was a bounty of sorts—one half of the damages awarded in a civil action against the false marker.[5]

I. The Marking Statutes

The false marking statute works hand-in-hand with the marking statute, 35 U.S.C. § 287(a).[6] The marking statute creates an incentive for patent-holders to mark their patented products. Under section 287, a patentee cannot recover damages for past infringement unless the patentee marks its patented product or otherwise notifies the accused infringer.

While section 287 incentivizes patent-holders to mark, section 292 establishes a penalty for marking improperly. Section 292 establishes a penalty for falsely marking an "unpatented article" as patented.[7] Notably, section 292(b) provides that "[a]ny person may sue for the penalty, in which event one-half shall go to the person suing and the other to the use of the United States."[8] The statute allows a person to recover up to $500 for every "offense" and arguably makes the false marking statute one of the few remaining qui tam statutes.

* *Trevor K. Copeland and Laura A. Lydigsen are associates at the Chicago firm of Brinks Hofer Gilson & Lione. The views expressed in the article are the authors' and not those of their law firm or any of its past, present, or future clients.*

II. Historical Context

A. Evolution of section 292, 1842-Present

Congress first enacted both the marking and false marking statutes in 1842.[9] Although the available legislative history is silent as to Congress's motivation for enacting the marking statutes, history suggests some possible influencing factors. First, in the mid-1800s the United States Patent and Trademark Office ("USPTO") lacked complete records of all issued patents due to an 1836 fire which burned the Patent Office, including 10,000 patents. Of the lost patents, only 2,845 were "reconstructed" by contacting the patent owners.[10] Second, the public had limited access to copies of issued patents in the 1800s. At the time, the USPTO published descriptions of issued patents in the Franklin Institute of the State of Pennsylvania.[11] As such, it would have been difficult to determine whether a particular article was patented unless the patent owner itself advertised the fact.

In contrast to the limited resources available in the mid-1800s, the USPTO now retains detailed public records of all issued patents. Free copies of United States patents may be obtained on a number of websites. There are also reliable free resources available to determine when a patent expires. For example, free services such as www.patentcalculator.com will determine the expiration date of an issued patent after a user enters data, most of which can be obtained from the front page of that issued patent.

Despite these dramatic changes to information accessibility, the current false marking statute is remarkably similar to its antiquated predecessor. As originally enacted in 1842, the statute established a "penalty of not less than one hundred dollars," with one half payable to the United States and the "other half to any person or persons who shall sue for the same."[12] Similarly, the modern statute prohibits marking an "unpatented article" and establishes an award with one half payable to "[a]ny person" who brings suit for false marking.[13]

In contrast, the marking statute has undergone a series of changes over the years. When it was enacted in 1842, the marking statute provided that a patentee who "neglect[ed]" to mark a patented article was "liable to the same penalty" as applicable for false marking.[14] Thus, the original marking statute established an affirmative duty to mark. However, the current marking statute has replaced the monetary penalty with an affirmative incentive: a patent owner may obtain monetary damages in a patent infringement action only if a patent owner marks its products or otherwise provides actual notice. Requirements for compliance with the marking statute have also evolved between 1842 and the present. In 1842, the patent owner was required to mark with "the date of the patent." This made sense then, because the issue date determined the expiration date of the patent, and many

patents still in force in 1842 were not numbered. In contrast, current section 287 requires only "the word 'patent' or the abbreviation 'pat.', together with the number of the patent."[15] The patent number allows anyone to gather the information necessary to determine whether it is expired, and evaluate the scope of the patent it claims. Because utility patent term is no longer dependent upon issue date, the 1842 marking requirement would not work today.

B. Qui tam statutes, then and now

Qui tam actions originated in England and were prevalent in early America around the time of the framing of the Constitution.[16] A qui tam action allows a private prosecutor (called a qui tam relator) to bring suit on behalf of the government and to share in any recovery.[17] Early informer statutes allowed an informer to retain a portion of the "bounty" received as a result of bringing suit.[18] Such statutes provided a supplemental means of law enforcement during the early republic.[19]

Informer statutes were subject to abuse, particularly obsolete statutes.[20] Plaintiffs might bring vexatious suits based on obsolete statutes, or the statutes could be rendered ineffective because a wrongdoer's friend could bring suit and settle for nominal damages or allow the wrongdoer to win.[21] To curb abuses, American legislators imposed strict limits on qui tam statutes.[22] Over time, these statutes gradually died out.[23] Only three American qui tam statutes have survived: the false marking statute, the False Claims Act,[24] and one Indian protection statute.[25, 26]

Of these, until the recent spate of false marking suits, only the False Claims Act (FCA) was commonly litigated.[27] Just as with other qui tam statutes, the FCA was commonly abused in the years after it was enacted.[28] In response to the abuse, Congress added a number of procedural safeguards to the FCA requiring that: (*i*) the relator deliver the complaint and any supporting evidence to the Government;[29] (*ii*) the Government has 60 days to intervene;[30] and (*iii*) the relator's recovery may be reduced if the Government opts to intervene.[31]

III. Federal Circuit Decisions Addressing False Marking

Only two precedential Federal Circuit decisions provide substantive analysis of the false marking statute—*Arcadia Machine & Tool, Inc. v. Sturm, Ruger & Co.* and *Clontech Laboratories, Inc. v. Invitrogen Corp*.[32] Both decisions focus on the requirement that the defendants have an intent to deceive the public.[33] Neither decision addresses the issue of plaintiff standing. Likewise, neither addresses whether marking with an expired patent comprises false marking under section 292.

IV. False Marking Plaintiffs: A New Kind of Patent Troll?

Several individuals recently have sought to exploit the false marking statute, many alleging that marking products with the numbers of expired patents constitutes false marking. Although this concept has been discussed among law students, professors, and patent practitioners for years, until recently, false marking claims usually were brought by accused patent infringers as counterclaims or declaratory judgment claims in a larger dispute where the accused infringer alleged that the patent owner's marked product was not actually covered by its patent.[34] In contrast, many of the recent false marking actions involve marking with expired patents and were brought by individuals who have no interest other than the statutory bounty.

Pursuit of qui tam false marking claims for marking with an expired patent as the primary, if not sole, basis for lawsuits appears to have its genesis in some loose wording from a footnote in a 2006 district court decision.[35] That case, which is devoid of any statutory construction analysis, relied as its sole authority on an equivocal statement from a patent treatise that "a strong case can be made" for false marking based on an expired patent.[36] The court was likely influenced by the egregious conduct of the defendant, which included: (*i*) beginning to mark its products with the expired patent more than one year after the patent expired; (*ii*) marking products with several other patents that did not cover the products; and (*iii*) sending letters to customers threatening suit based on its patents—regardless of whether they actually covered the product in question. In the face of such blatant misconduct, the court summarily adjudged the defendant's marking a product with the number of an expired patent to be false marking.[37] Relatively soon thereafter, several lawsuits were filed alleging false marking when the patentee continued marking products after the subject patent expired.

Matthew Pequignot was one of the earliest plaintiffs, filing three lawsuits exclusively on the basis of alleged patent false marking.[38] Pequignot, a patent attorney, filed a pro se complaint alleging that Solo Cup Corporation falsely marked its coffee cup lids with expired patent numbers. Pequignot also sued Gillette alleging false marking with expired patent numbers as well as false marking with patents not corresponding to marked products.[39] He also sued Arrow Fastener Co., a case which is reported to have settled for an undisclosed sum.[40] As of this writing, the Gillette court is considering opposing motions on whether false marking is fraud such that any complaint alleging it must meet the particularized pleading requirements of Fed. R. Civ. P. 9(b), as was recently held by one California district court.[41]

Pequignot's suit against Solo Cup was dismissed when the court granted Solo Cup's motion for summary judgment based upon its lack of intent, a decision that Pequignot is appealing.[42] Pequignot previously had survived two motions to dismiss: (1) a motion to dismiss for failure to state a claim based principally on the argument that marking with an expired patent is not actionable under section 292; and (2) a motion to dismiss for lack of jurisdiction, in which Solo Cup alleged that Pequignot lacked Article III standing to bring suit because he had not suffered any injury.[43] Both motions were denied.[44]

Another patent attorney, James Harrington, filed several false marking lawsuits throughout 2008. In *Harrington v. New Products Marketing, Inc.*, Harrington alleged false marking of a string reel device marked with two commonly-assigned patent numbers when the two patents had been distinguished from one another on a structural basis during prosecution.[45] This lawsuit was dropped before any answer was filed. In a second lawsuit filed against Monsanto, Pioneer Hi-Bred, Asgrow, and other seed-selling companies, Harrington alleged false marking

of several seed products on behalf of a group of famers.[46] A third Harrington-led lawsuit alleges that CIBA Vision Corp. falsely marked a contact lens disinfectant product because the claims of its patents are allegedly directed to methods and an apparatus rather than to a cleaning product.[47] The judge in the CIBA Vision case denied a motion to dismiss based upon a challenge to the constitutionality of the false marking statute.

Still another patent attorney, Paul Hletko, through a newly formed company (Heathcote Holdings) sued the maker of Mentadent® toothpaste, Arm & Hammer CleanShower®, and Nair for Men® hair remover, alleging false marking.[48] As of this writing, the court is considering the defendant's motion to dismiss, which alleges insufficient particularity of pleadings under Fed. R. Civ. P. 9 and insufficiency of fact pleading, and threatens to later raise constitutional arguments if needed.

Jennifer Brinkmeier sued Graco Children's Products, Inc., alleging "patent misuse" but appears to be alleging false marking.[49] Ms. Brinkmeier alleged that Graco committed false marking by placing patent numbers of expired patents on its website and variety of "play yard" child-care products. Like Gillette, Graco has alleged inadequacy of the complaint under Fed. R. Civ. P. 9(b). Graco has also argued that the complaint fails to allege intent to deceive because Graco marked with conditional language ("protected by one or more of the following patents").[50]

In *Brule Research Associates Team, LLC v. A.O. Smith Corp.*, Brule's 92-page complaint cites 33 expired patents and four patents allegedly not covering any marked product, one or more of which is alleged to have been marked on one or more of the hundreds of models of water heaters that defendant makes.[51] In a motion to dismiss, the defendant has challenged, inter alia, the plaintiff's Article III standing. The U.S. government has sought to intervene in the case to defend the constitutionality of the false marking statute.

Raymond Stauffer, an attorney, sued Brooks Bros. alleging that its placing numbers of expired patents on bow ties was false marking.[52] The court determined that Stauffer had suffered no injury in fact as required by Article III and dismissed the false marking claims.[53] Stauffer appealed in early July 2009.

Public Patent Foundation, Inc. (PUBPAT), a New York not-for-profit claiming to protect freedom in the patent system,[54] has filed several false marking lawsuits, alleging that the defendant is violating § 292 by marking with the numbers of expired patents. Its targets to date include Cumberland Packaging Corp. (maker of Sweet-n-Low®);[55] Iovate Health Science Research, Inc. (maker of Hydroxycut® and Xenadrine® weight-loss products);[56] McNeil-PPC, Inc. (maker of Tylenol®);[57] and GlaxoSmithKline Consumer Healthcare, L.P.,[58] and it filed an amicus brief siding with Pequignot in the Solo Cup case.

V. Constitutional Issues: Article III Standing and Article II's Take Care Clause

These plaintiffs each may be alleged to suffer from some deficiency with regard to Article III standing.[59] In particular, an Article III jurisprudence has evolved that poses difficulties for many qui tam plaintiffs.

The history and evolution of the standing doctrine serve an interesting example of the "law of unintended consequences" —particularly as applied to the false marking statute. This problem is not new, as recognized by its discussion in several academic papers,[60] but it has become a very real and practical problem in view of the recent wave of patent false marking litigation. The requirement that a plaintiff must present a case or controversy to be heard by a federal court is as old as the Republic.[61] Courts throughout the 19th century considered whether parties had a right to be heard in federal courts, but did not dwell or expound upon a doctrine of standing with the specificity that developed throughout the 20th century.[62]

During the first half of the 20th century, U.S. courts developed a rigorous framework for determining whether a plaintiff has standing under Article III to bring a lawsuit. The framework developed with virtually no consideration for the vanishing—but not yet extinct—qui tam statutes. In spite of being highly valued by the government for its fraud-deterrent value, the FCA and other qui tam statutes seem virtually to have been ignored during the evolution of modern standing doctrine. This difficulty has yet to be fully addressed by the Supreme Court, even in its most recent pronouncement on qui tam plaintiff standing in *Vermont Agency of Natural Resources v. United States* ex rel. *Stevens*, where the Court attempted to reconcile the qui tam provisions of the FCA with its 20th century standing jurisprudence.[63] In *Vermont Agency*, the Court held that every plaintiff (including a qui tam plaintiff) must meet an irreducible constitutional minimum of standing under Article III of the Constitution: (1) "he must demonstrate an injury in fact—a harm that is both concrete and actual or imminent, not conjectural or hypothetical"; (2) "he must establish causation—a fairly traceable connection between the alleged injury in fact and the alleged conduct of the defendant"; and (3) "he must demonstrate redressability—a substantial likelihood that the requested relief will remedy the alleged injury in fact."[64] The Court concluded that an FCA qui tam plaintiff has Article III standing as a partial assignee of the government's interest, based upon explicit FCA statutory provisions.

There are some differences between the FCA and the false marking statute, which make the *Vermont Agency* court's assignment rationale less applicable to the false marking statute. As noted above, the FCA was amended to curb abuses, with rigorous procedural safeguards in place. Congress reformed the law to require a rigid schedule that mandated notice to the government, which could opt in or opt out of the case (with a reserved right to enter at a later date). Several courts have relied on these stringent requirements, wherein the government maintains a degree of control over the lawsuit, in upholding the constitutionality of the FCA.[65] These courts highlight a second constitutionality issue with regard to the FCA and the False Marking statute, namely the "Take Care" clause of Article II, which requires the Executive Branch to see that the law is enforced. This requirement does not allow for absolute delegation of that power to another party.[66]

However, a contrast between the FCA and false marking statutes is clear: the FCA has a comprehensive framework that requires the Executive Branch to be notified and that provides it with power to control the litigation and terms of settlement, if any. In contrast, there is not even provision for

effective notice of a false marking suit to the law enforcers of the Executive Branch (e.g. the Attorney General's office), much less provision for any governmental control.[67]

For plaintiffs currently pursuing section 292 as a primary cause of action, the standing issue arises from a tension between (*A*) the statute's provision that "any person may sue", and (*B*) an apparent lack of the qui tam plaintiff's injury in fact.[68] As such, there are three theories upon which a qui tam plaintiff may rely to pursue a lawsuit: (1) a plaintiff is a relator, effectively an assignee of the government, having a right to bring the lawsuit on behalf of himself and the government under assignment granted by the "any person" language of the statute; (2) because the false marking statute is a qui tam statute that pre-dates modern standing doctrine, its provision that "any person may sue" prevails over later-developed standing requirements without undermining the constitutionality of the statute as applied;[69] or (3) the plaintiff, as a member of the public, has suffered injury by the harm of false marking to the public interest.

The first argument faces difficulties: the false marking statute lacks any language assigning relator status. The court in *Solo Cup* read in such a provision.[70] However, under a plain reading of the statute, even language stating that "any person may sue" and requiring that person to provide one-half of any penalty to the United States does not provide for offer, acceptance, consideration, and a meeting of the minds, or any other contract basics that would be considered binding on parties in any other circumstance (notwithstanding the Court's reliance upon language of qui tam statutes that became defunct before or essentially outside of the early-to-mid-20th century evolution of Article III constitutional standing jurisprudence).[71]

Nonetheless, in some limited circumstances, an injury suffered by the United States may be assigned to an individual such that the individual is allowed to bring an action even though she has not personally suffered any injury.[72] The FCA provides just that, as it expressly (*i*) authorizes an individual to bring an action *on behalf of* the United States, (*ii*) authorizes an individual to bring an action *in the name of* the United States, and (*iii*) allows an individual to proceed on behalf of the United States—all only after first giving the United States notice of the lawsuit.[73] The current leading case, *Vermont Agency,* held that a qui tam plaintiff/relator has standing only as a partial assignee of the government's interest/injury and negated various other methodologies by which lower courts had previously granted standing in qui tam lawsuits.[74]

The second argument presents a conflict with Supreme Court precedent requiring injury in fact for Article III standing. A court may not ignore clear Supreme Court precedent requiring that a plaintiff must have an actual injury.[75] It is clear that a plaintiff must satisfy the Article III standing requirement in order to bring a lawsuit—regardless of the basis.[76]

With regard to the third argument, the Supreme Court has generally held that a single member of the public will not have standing to bring suit for a generalized but highly dilute harm.[77] A plaintiff cannot rely upon an injury to another, including the United States, to provide standing—even as a citizen or taxpayer,[78] except under very narrow circumstances where a clearly defined constitutional right is at issue.[79] None of these circumstances applies to a false marking plaintiff.

VI. Expired Patents

The recent flood of false marking qui tam suits have raised a second difficult question: what is the effect of marking with expired patents?[80] The statute provides that marking with an "unpatented" article is required for false marking. However, section 292 provides little guidance as to what "unpatented" means. Applying principles of statutory construction, arguments can be made both for and against section 292 covering marking with expired patents.

The only court to squarely address this issue held that "unpatented" articles under section 292 include articles covered by expired patents based on the ordinary meaning of "unpatented" and public policy.[81] The *Pequignot* court asserted that black-letter patent law indicated that articles covered by expired patents are unpatented because they are in the public domain.[82] The court further reasoned that court decisions using the word "unpatented" indicated that the ordinary meaning of the term encompassed articles covered by expired patents.[83] The court found that the doctrine of double-patenting supported finding articles covered by expired patents "unpatented" for purposes of section 292.[84] The court also held that public policy supported construing section 292 to cover marking with expired patents because marking with a patent number was analogous to a "no trespassing sign" and that "[t]he public could no longer assume the status of the intellectual property by the simple presence of a 'Patent No. XXX' marking."[85] The court said that marking with an expired patent would "force [potential inventors and consumers] to look up *every* patent marking to discern whether the patent was valid or expired, possibly leading some to shy away from using that article."[86]

Although the *Pequignot* court presents some strong reasons for finding that marking with an expired patent is false marking, there are also several compelling arguments for holding that marking with expired patents is not false marking.[87] First, other provisions of the Patent Act suggest that articles covered by expired patents are not "unpatented" for purposes of the Patent Act. The third clause of section 292(a) suggests that articles covered by expired patents were not contemplated by the false marking statute. The third clause provides a specific remedy for false marking relative to a patent application that is no longer pending, in effect, "expired."[88] In contrast, there is no parallel provision in the second clause providing a penalty for false marking relative to a patent that is expired. In addition, Section 271(a) of the Patent Act provides that making, using, selling, or offering to sell "any patented invention during the term of the patent" comprises infringement of that patent. The term "patented" is modified by the explanatory phrase "during the term of the patent." This suggests that the term "patented" alone refers to inventions both during the term of the patent and after expiration of the patent term. If the terms "patented" or "unpatented" alone were sufficient to indicate whether an article was covered by an expired patent, the phrase "during the term of the patent" would be superfluous.

Second, public policy also provides particularly compelling

grounds for finding that "unpatented" articles under section 292 exclude articles covered by expired patents. Marking with an expired patent allows anyone to quickly determine the patent expiration date. Indeed, even in the 1800s when obtaining copies of issued patents was not as simple as it is today, at least one court recognized that marking with an expired patent was probably not harmful to the public.[89]

Third, the penal nature of the false marking statute suggests that it should be strictly construed.[90] A strict construction would require construing the statute to exclude activities that are not clearly covered by the statutory language.[91]

VII. Conclusions

The proliferation of false marking qui tam suits in the past two years might suggest that the false marking statute is undergoing a resurrection of sorts. Instead of being relegated to a counterclaim in infringement actions as it has been for most of the 20th century, the false marking statute has re-emerged as its own animal, providing the sole basis for law suits seeking millions of dollars in damages. However, it is unlikely that this trend will continue in its present form. In light of the Article III constitutional challenges to plaintiff standing and Article II complications described herein, the false marking statute almost certainly must evolve or be rendered extinct.

So far, the courts that have dealt with these qui tam false marking suits to completion have unanimously issued judgments favorable to the defendants. As noted above, some courts have found lack of intent to deceive.[92] At least one has based its decision on the lack of Article III standing.[93] Other courts may yet decide that marking with an expired patent is not actionable marking or that qui tam suits brought under section 292 may not be permissible under the take care clause of Article II.

The reason for the unanimous outcomes in favor of the qui tam defendants may ultimately be rooted in the equities. Most of the recent qui tam plaintiffs have sought massive damages judgments; for instance, the plaintiff in the *Pequignot* case sought $500 for every disposable cup lid manufactured by Solo Cup Company. The plaintiffs in nearly all of the recent qui tam false marking actions cannot make any argument of actual injury. Indeed, the fact that most of these plaintiffs have focused on marking with expired patents (with no evidence of any threats of patent assertion) speaks volumes. These individuals can determine in minutes whether a patent is expired and then seek massive judgments, saddling companies with either undergoing the expense of discovery or settling to the benefit of these opportunists. This does not seem like the outcome that was intended in 1842 when the false marking statute was enacted. Indeed, the false marking qui tam suits seem similar to the types of abusive suits based on arcane statutes that contributed to the extinction of qui tam in England and the near end of qui tam in the United States.

So should the qui tam provisions of section 292 be repealed and rendered extinct?[94] Or just modified? The authors believe that legislative action may resolve the problems with qui tam false marking suits without eliminating false marking altogether and that legislatively-directed evolution is better suited to address the issues raised here than piecemeal and contradictory court decisions that will potentially (*i*) take too long to offer a resolution at a nationwide level; and (*ii*) have a negative impact upon predictability for patent-holders and the way they mark their products. With regard to standing and the take care clause, a framework modeled on the False Claim Act's provisions for providing the government with notice and the right to intervene in qui tam action would likely resolve these constitutional problems. It also may discourage opportunistic qui tam relators from bringing unmeritorious actions.[95] Legislative modification to clarify whether false marking encompasses marking with expired patents would also prove beneficial.[96] Congress should evaluate whether marking with expired patents is actually harmful to the public given today's free patent resources, and particularly to competition, and determine whether to exclude this from the statute.

This article has dealt only in summary fashion with a few issues regarding the false marking statute, but these and other issues (e.g., counting false marking offenses, determining what comprises false marking with respect to method claims, res judicata effect of a false marking verdict on less than all of a litany of patents marked on a product, etc.) necessitate that this qui tam dinosaur evolve or be rendered extinct – one hopes without damaging free markets and the patent system.

Endnotes

1 35 U.S.C. § 292.

2 Art. I, § 8, cl. 8 provides, "The Congress shall have power... [t]o promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries…."

3 5 Stat. 544-45 (1842).

4 Qui tam is short for *qui tam pro domino rege quam pro se ipso in hac parte sequitur*, which means "who pursues this action on our Lord the King's behalf as well as his own." *See* Vt. Agency of Natural Res. v. United States *ex rel.* Stevens, 529 U.S. 765, 768-69 n.1 (2000); 3 W. Blackstone, Commentaries on the Laws of England 160 (1st ed. 1768).

5 5 Stat. 544-45 (1842).

6 The marking statute provides:

> Patentees, and persons making, offering for sale, or selling within the United States any patented article for or under them, or importing any patented article into the United States, may give notice to the public that the same is patented, either by fixing thereon the word "patent" or the abbreviation "pat.," together with the number of the patent, or when, from the character of the article, this cannot be done, by fixing to it, or to the package wherein one or more of them is contained, a label containing a like notice. In the event of failure so to mark, no damages shall be recovered by the patentee in any action for infringement, except on proof that the infringer was notified of the infringement and continued to infringe thereafter, in which event damages may be recovered only for infringement occurring after such notice. Filing of an action for infringement shall constitute such notice.

35 U.S.C. § 287(a).

7 The false marking statute provides:

> (a) Whoever, without the consent of the patentee, marks upon, or affixes to, or uses in advertising in connection with anything made, used, offered for sale, or sold by such person within the United States, or imported by the person into the United States, the name or any imitation of the name of the patentee, the patent number, or the words "patent," "patentee," or the like, with the intent of counterfeiting or

imitating the mark of the patentee, or of deceiving the public and inducing them to believe that the thing was made, offered for sale, sold, or imported into the United States by or with the consent of the patentee; or Whoever marks upon, or affixes to, or uses in advertising in connection with any unpatented article the word "patent" or any word or number importing the same is patented, for the purpose of deceiving the public; or Whoever marks upon, or affixes to, or uses in advertising in connection with any article the words "patent applied for," "patent pending," or any word importing that an application for patent has been made, when no application for patent has been made, or if made, is not pending, for the purpose of deceiving the public - Shall be fined not more than $500 for every such offense.

(b) Any person may sue for the penalty, in which event one-half shall go to the person suing and the other to the use of the United States.

35 U.S.C. § 292.

8 *Id.*

9 5 Stat. 544-45 (1842).

10 December 15 Marks the 165th Anniversary of the Great Patent Office Fire of 1836, http://www.uspto.gov/web/offices/com/speeches/01-60.htm (last visited Aug. 19, 2009).

11 H.R. 102, 30th Cong. § 12 (1st Sess. 1848).

12 As originally enacted in 1842, the statute prohibited marking an "unpatented article." 5 Stat. 544 (1842). The original statute allowed an individual to bring suit for false marking in the courts. *Id.*

13 35 U.S.C. § 292.

14 S. 220, 27th Cong. § 6 (2d Sess. 1842).

15 35 U.S.C. § 287.

16 *Vt. Agency*, 529 U.S. at 776; Note, History and Development of Qui Tam, 1972 Wash. U. L.Q. 81, 94-97 (1973).

17 *See* Wash. U. L.Q., *supra* note 16, at 87.

18 *Vt. Agency*, 529 U.S. at 774.

19 *See* Wash. U. L.Q., *supra* note 16, at 101.

20 *Vt. Agency*, 529 U.S. at 775. *See* Wash. U. L.Q., *supra* note 16, at 97. It should be noted that there is a technical distinction between "informer statutes"—where a governmental authority would still prosecute the offense upon notice by an informer in exchange for a bounty (with the bounty subject to successful prosecution)—and "qui tam statutes"—where the informer would prosecute the suit and then split the proceeds with the government, but the concerns and abuses overlap such that the distinction is immaterial for the present discussion.

21 *See* Wash U. L.Q., *supra* note 16, at 89.

22 Legislators employed several tactics to curb qui tam abuse: (*i*) requiring the informer to pay costs if it did not prevail; (*ii*) limiting the preclusive effect of informer suits; (*iii*) enacting strict statutes of limitations and venue statutes; (*iv*) imposing fines on wrong-doing informers; and (*v*) giving the State complete control over penal actions. *See id.* at 97.

23 *See id.* at 99.

24 31 U.S.C. §§ 3729-3733.

25 25 U.S.C. § 201.

26 The *Vermont Agency* court listed four existing American qui tam statutes as of 2000. *Vt. Agency*, 529 U.S. at 768-69 n.1. One of those statutes, 25 U.S.C. § 81, has since been repealed by amendment. Indian Tribal Economic Development and Contract Employment Act of 2000, Public Law 106-179, 114 Stat. 46-47 (codified as amended at 25 U.S.C. § 81). England repealed its last informer statute in 1951. *Vt. Agency*, 529 U.S. at 776.

27 Evan Caminker, The Constitutionality of Qui Tam Actions, 99 Yale L. J. 341, 342 n.5 (1989).

28 *E.g.,* United States *ex rel.* Marcus v. Hess, 317 U.S. 537, 545-47 (1943).

29 31 U.S.C. § 3730(b)(2).

30 *Id.* § 3730(b)(2), (4).

31 *Id.* § 3730(d)(1)-(2). A review of the statutory history and contemporary case law shows that reforms in the 1940s weakened the FCA significantly, but reforms in the mid-1980s added teeth back to it, while providing for strong governmental involvement. However, detailed treatment is outside the scope of the present article. *See, e.g.,* Sullivan, M.A., A "False Claims Act" Is Finally Enacted in Georgia: What Georgia Lawyers Should Know About the "State False Medicaid Claims Act," 13 Georgia Bar Journal 12 (2007).

32 Clontech Labs., Inc. v. Invitrogen Corp., 406 F.3d 1347, 1351 (Fed. Cir. 2005); Arcadia Mach. & Tool, Inc. v. Sturm, Ruger & Co., 786 F.2d 1124 (Fed. Cir. 1986). The false marking statute is part of the substantive patent law, and the United States Court of Appeals for the Federal Circuit, therefore, has exclusive appellate jurisdiction. 28 U.S.C. § 1295(a)(1).

33 In *Arcadia*, the defendant affixed labels that provided, "This Ruger firearm is manufactured under one or more of the following U.S. Patents, or under one or more Patents Pending," and listed a number of patents. 786 F.2d at 1125. The Federal Circuit affirmed the district court's grant of summary judgment of no false marking because the plaintiff "had totally failed, after at least nine months of discovery, to produce any evidence of intent to deceive the public." *Id.*

In *Clontech Laboratories, Inc. v. Invitrogen Corp.*, the district court held that Invitrogen intentionally falsely marked its molecular biology products based on experiments that put Invitrogen on notice that its products were not actually covered by its patents. 406 F.3d at 1351. The Federal Circuit's decision clarifies that the requisite intent for false marking must be shown by a preponderance of the evidence under objective standards. *Id.* at 1352. This requires that the false marking plaintiff show that the accused false marker did not have a reasonable belief that its articles were properly marked. *Id.* at 1352-53. An inference of intent to deceive may arise if the plaintiff demonstrates that the defendant had knowledge of an actual misrepresentation. *Id.* The defendant cannot overcome the inference of intent to deceive by a claim that it honestly believed its products were properly marked. *Id.* at 1353 n.3. Based on the foregoing objective standards, the Federal Circuit held that the *Clontech* district court erred in finding intent to deceive. *Id.* at 1355.

34 *E.g.,* Brose v. Sears Roebuck & Co., 455 F.2d 763, 765 (5th Cir. 1972); FMC Corp v. Control Solutions, Inc., 786 F. Supp. 1287, 1296 (N.D. Ohio 1991); Smith Welding Equip. Corp. v. Pearl, 21 F.R.D. 196 (W.D. Pa. 1956). In some instances, false marking lawsuits also arise between business competitors as a cause of action directly related to the parties' competing interests. *See, e.g., Clontech*, 406 F.3d at 1351; D.P. Wagner Mfg., Inc. v. Pro Patch Sys., Inc., 434 F. Supp. 2d 445 (S.D. Tex. 2006).

35 *D.P. Wagner*, 434 F. Supp. 2d 445.

36 *Id.* at 452 n.3 (citing Chisum on Patents § 20.03[7][c][vii], at 20-657).

37 *Id.*

38 Pequignot v. Solo Cup Co., No. 1:07-cv-897 (E.D. Va.); Pequignot v. Gillette Co., No. 2:08-cv-222 (E.D. Tex.); Pequignot v. Arrow Fastener Co., No. 2:08-cv-353 (E.D. Tex.).

39 Pequignot initially sued both Solo Cup and Gillette in the Eastern District of Virginia. He later obtained counsel, dropped the Gillette case in Virginia, and filed suit against Gillette in the Eastern District of Texas. Interestingly, Pequignot's counsel in the Solo Cup, Gillette, and Arrow Fastener cases now represents several other plaintiffs in different actions across the country, including Brule Research Associates and Heathcote Holdings Corp.

40 http://www.bloomberg.com/apps/news?pid=20670001&sid=aB9JR.4T3mAM (last visited Aug. 19, 2009). However, as the stipulated dismissal was without prejudice, its exact status is unclear.

41 Juniper Networks v. Shipley, No. C 09-0696, 2009 WL 1381873, at *4 (N.D. Cal. May 14, 2009).

42 Minute Order, Pequignot v. Solo Cup. Co., No. 1:07-cv-897, (E.D. Va. Jul. 2, 2009); Pequignot v. Solo Cup. Co., No. 1:07-cv-897, slip op. (E.D. Va. Aug. 25, 2009), *appeal docketed*, No. 2009-1547 (Fed. Cir. Sept. 10, 2009). In the spirit of full disclosure, it should be noted that the authors served as counsel for Solo Cup early in the Pequignot litigation, but are no longer so engaged, as Solo Cup followed the lead attorney in that case to another law firm in July 2008.

43 This motion alleged that, if Pequignot had standing, the application of the false marking statute to him would be suspect under the Take Care clause

of the U.S. Constitution, Art. II, discussed below. Based upon this, the Court sua sponte demanded briefing from the United States on the topic, and the U.S. remained engaged as a third party.

44 Pequignot v. Solo Cup. Co., No. 1:07-cv-897, 2009 WL 874488 (E.D. Va. Mar. 27, 2009) (denying motion to dismiss for lack of jurisdiction based on lack of standing); Pequignot v. Solo Cup Co., 540 F. Supp. 2d 649 (E.D. Va. Mar. 24, 2008) (denying motion to dismiss under 12(b)(6)).

45 Complaint at 4-5, Harrington v. New Products Marketing Corp., No. 3:08-cv-225 (W.D.N.C. filed May 13, 2008).

46 N.C. Farmers' Assistance Fund v. Monsanto, No. 08-cv-409 (M.D.N.C., filed June 17, 2008). One of the defendants, Pioneer, is challenging this suit under the U.S. Const., Art. II.

47 Complaint at 3-4, Harrington v. CIBA Vision Corp., 3:08-cv-251 (W.D.N.C. filed June 3, 2008).

48 Heathcote Holdings Corp., Inc., v. Church & Dwight Co., Inc., No. 2:08-cv-00349 (E.D. Tex. filed Sept. 12, 2008).

49 Brinkmeier v. Graco Children's Prods. Inc., No. 09-cv-00262 (D. Del. filed Apr. 20, 2009).

50 In *Arcadia*, the CAFC affirmed that a patentee's use of conditional marking language (this product is covered by "one or more" of the following patents...) is neither per se untrue nor deceptive. 786 F.2d at 1125.

51 No. 1:08-cv-977 (E.D. Va. filed Sept. 19, 2008); No. 2:08-cv-1116 (E.D. Wisc. filed Dec. 23, 2008). The suit was originally filed in the Eastern District of Virginia in September 2008, and assigned to the same judge as the Solo Cup case. However, the defendant prevailed on a venue-transfer motion, and it is on hold in the Eastern District of Wisconsin, pending a decision in the *Solo Cup* appeal. Similar to some so-called "non-practicing-entities/patent trolls," according to public records, the plaintiff is a limited liability company that was formed about three weeks before the lawsuit was filed, with its sole member being Wisconsin attorney Kathleen Eisenmann and its registered agent being Virginia attorney F. Joseph Brinig, who is the counsel of record for plaintiff.

52 Stauffer v. Brooks Bros., Inc., 615 F. Supp. 2d 248, 250-51 (S.D.N.Y. 2009).

53 *Id.* at 255-56. The court declined to reach the issues of Article II constitutionality, false marking status of marking with expired patents, and whether Stauffer's complaint met the heightened pleading requirement of Fed. R. Civ. P. 9(b). *Id.* at 251 n.1.

54 About PUBPAT, http://www.pubpat.org/About.htm (last visited Aug. 19, 2009).

55 Public Patent Found., Inc. v. Cumberland Packing Corp., No. 1:09-cv-04360 (S.D.N.Y. filed May 6, 2009) (voluntarily dismissed without prejudice July 6, 2009).

56 Public Patent Found., Inc. v. Iovate Health Sci. Research Inc., 09-cv-04361 (S.D.N.Y. filed May 6, 2009) (voluntarily dismissed without prejudice, July 6, 2009).

57 Public Patent Found., Inc. v. McNeil-PPC, Inc., No. 09-cv-5471 (S.D.N.Y. filed June 16, 2009).

58 Public Patent Found., Inc. v. GlaxoSmithKline Consumer Healthcare, L.P., No. 09-cv-5881 (S.D.N.Y. filed June 26, 2009).

59 One potential exception is the North Carolina farmers of *North Carolina Farmers' Assistance Fund v. Monsanto*, who arguably would pay lower prices for unpatented seed than for patented seed if the patent does not actually cover the seed at issue. *N.C. Farmers' Assistance Fund*, No. 08-cv-409 (M.D.N.C., filed June 17, 2008).

60 *See, e.g.,* O.B. Roberts, Actions Qui Tam Under the Patent Statutes of the United states, 10 Harvard L. Rev. 265 (1896); Wash. U. L.Q. 81, *supra* note 16; T.R. Lee, The Standing of Qui Tam Relators Under the False Claims Act, 57 U. Chi. L. Rev. 543 (1990).

61 U.S. Const., Art. III, § 2 provides, in relevant part, "The judicial power shall extend to all cases, arising in Law and Equity, under this Constitution, the Laws of the United States, ... [and] to Controversies... between Citizens of different States...."

62 *See* Ann Woolhandler & Caleb Nelson, Does History Defeat Standing Doctrine?, 102 Mich. L. Rev. 689 (2004).

63 528 U.S. 765.

64 *Id.* at 771.

65 United States *ex rel.* Stone v. Rockwell Int'l Corp., 282 F.3d 787, 807 (10th Cir. 2002) (qui tam provisions of FCA did not violate Article II when U.S. intervened); Riley v. St. Luke's Episcopal Hosp., 252 F.3d 749, 753 (5th Cir. 2001) ("the Executive retains significant control over litigation pursued under the FCA by a qui tam relator"); United States *ex rel.* Taxpayers Against Fraud v. Gen. Elec. Co., 41 F.3d 1032, 1041 (6th Cir. 1994) (under the FCA "the Executive Branch retains 'sufficient control' over the relator's conduct to insure that the President is able to perform his constitutionally assigned duties"); United States *ex rel.* Kelly v. Boeing Co., 9 F.3d 743, 758 (9th Cir. 1993) (holding that Executive Branch retains "sufficient control" over FCA relators).

66 The President "shall take Care that the Laws be faithfully executed." U.S. Const. Art. II, § 3; *cf.* Morrison v. Olsen, 487 U.S. 654, 696 (1988) (upholding independent prosecutor statute as constitutional only because it gave the Executive Branch "sufficient control over the independent counsel to ensure that the President is able to perform his constitutionally assigned duties"); Printz v. United States, 521 U.S. 898, 922 (1997) (striking provisions of the Brady Act that "effectively transfer[red] this responsibility [to take care that the laws be faithfully executed] to thousands of [state officers] without meaningful Presidential control").

67 Some arguments have been raised by the U.S. government (e.g., when it was invited by the court to intervene in the Solo Cup case) and by parties in other cases that the mandatory notice of a patent case having been filed that must be communicated to the USPTO within one month would suffice for notice to the Executive branch. However, there is no provision at the USPTO for notifying any law enforcement power of a false marking suit, such that the Executive Branch has no opportunity to intervene (unless invited sua sponte by the Court or a party to the suit), and a plaintiff who settled with a defendant within that first month—or at any time—would be hard put to determine where he should send a check with "the government's half of the money." Moreover, alleged false markers may collude with plaintiffs to settle on far different terms than the government would—on behalf of the people it purports to represent—be willing to accept. And, regardless of the fairness or desirability, principles of res judicata and equity would likely bind the government to the terms agreed upon by its erstwhile representative in the lawsuit. The Article II issues raised briefly here are too complex to address fairly in the scope of this article, and so are addressed here only in summary fashion to highlight one of the key Constitutional problems with the false marking statute.

68 A survey by the authors of false marking actions from the last 50 years shows that, in almost every case, the action was part of a larger patent infringement lawsuit or other business dispute between parties who clearly had standing on other grounds.

69 This approach essentially ignores the conflict, as illustrated in the *Solo Cup* decision. *Pequignot*, 2009 WL 874488, at *8. However, courts have seriously considered the "it's been done historically, so it must be all right" argument. *Cf.*, Riley v. St. Luke's Episcopal Hosp., 252 F.3d 749, 753 (5th Cir. 2001).

70 *Pequignot*, 2009 WL 874488, at *8.

71 *See, e.g.,* Federal Government Construction Contracts, 54-57. T.J. Kelleher Jr., T.E. Abernathy IV; and H.J. Bell Jr.; Wiley. 2008. (592 pp.).

72 529 U.S. 765.

73 31 U.S.C. § 3730(b)(2).

74 529 U.S. 765.

75 *Id.* at 771.

76 *Id.* (holding that FCA must meet Article III standing requirements).

77 Lujan v. Defenders of Wildlife, 504 U.S. 555, 562-63 (1992) (holding that the "'injury in fact' test requires more than an injury to a cognizable interest. It requires that the party seeking review be himself among the injured.").

78 Allen v. Wright, 468 U.S. 737, 752 (1984); *Lujan*, 504 U.S. at 573-74.

79 Flast v. Cohen, 392 U.S. 83, 99 (1968); Hein v. Freedom From Religion Found., 551 U.S. 587 (2007).

80 *E.g., Stauffer*, 615 F. Supp. 2d 248; *Pequignot*, 1:07-cv-897 (E.D. Va.); *Pequignot*, No. 2:08-cv-222 (E.D. Tex.); *Brinkmeier*, No. 09-cv-00262 (D. Del.); Brule Research Assocs. Team L.L.C. v. A.O. Smith Corp., No. 08-cv-1116 (E.D. Wisc.); Public Patent, 09-cv-04361 (S.D.N.Y.); Public Patent, No. 1:09-cv-04360 (S.D.N.Y.); Public Patent, No. 09-cv-5471 (S.D.N.Y.); Public Patent, No. 09-cv-5881 (S.D.N.Y.).

81 *Pequignot*, 540 F. Supp. 2d at 652-54. The district court in *D.P. Wagner* does not provide any statutory interpretation analysis. 434 F. Supp. 2d at 452 n.3. Relying on *Pequignot*, the Eastern District of Texas recently noted that "marking with expired patent numbers can be false marking." Hearing Components, Inc. v. Sure, Inc, No. 9:09-cv-104, 2009 WL 815526, at *4 (E.D. Tex. Mar. 26, 2009).

82 *Id.* at 652.

83 *Id.* (quoting Bonito Boats, Inc. v. Thunder Craft Boats, Inc., 489 U.S. 141, 159 (1989) (article "freely exposed to the public for a period in excess of six years... stands in the same stead as an item for which a patent has expired or been denied: it is unpatented and unpatentable"); Prestole Corp v. Tinnerman Prods., Inc. 271 F.2d 146, 155 (6th Cir. 1959) ("[A]fter the expiration of a patent, the invention originally protected thereby becomes, for all purposes, an unpatented device."); Sylvania Indus. Corp. v. Visking Corp., 123 F.2d 947, 950 (4th Cir. 1943) ("Patent No. 1,070,776 expired on August 19, 1930, and since that date [the articles] have been unpatented articles of commerce.")).

84 *Id.* at 653.

85 *Id.* at 654.

86 *Id.*

87 In addition to the arguments discussed herein, analysis of the case law and legislative history of section 292 lends support to interpreting "unpatented" articles in section 292 as excluding articles covered by expired patents. *See* Mem. in Supp. of Solo Cup Co.'s Mot. to Dismiss, No. 1:07-cv-897 (E.D. Va. filed Jan. 29, 2008) (discussing additional arguments supporting construing section 292 as excluding marking with expired patents).

88 The third clause of section 292(a) provides for a penalty against "Who[m]ever marks upon, or affixes to, or uses in advertising in connection with any article the words 'patent applied for,' 'patent pending,' or any word importing that an application for patent has been made, *or if made, is not pending*...." 35 U.S.C. § 292(a) (emphasis added).

89 In *Wilson v. Singer Manufacturing Co.*, 9 Biss. 173 (D.C. Ill. 1879), *aff'd*, 12 Fed. 57 (C.C.D. Ill. 1882), Wilson brought a qui tam action against Singer for marking its sewing machines as "Patented Sept. 10, 1846, May 8, 1849, Nov. 13, 1850, Aug. 4, 1851, Aug. 12, 1851, Apr. 11, 1854, May 30, 1854, Oct. 9, 1855" when its patents had expired by the time of the marking. *Id.* at 223. The district court reasoned that under the law governing patent terms then in effect, the patent terms for all of the marked patents were expired at the time of the marking. *Id.* Because the public was presumed to know the law and knew the issue dates of the patents, the public would not have been deceived. *Id.* In fact, the court reasoned that "[i]t may be valuable information to the public to be told that a machine offered for sale is made in accordance with a patent which has been granted, but which has expired." *Id.* The court of appeals affirmed.

90 *E.g.,* Mayview Corp. v. Rodstein, 620 F.2d 1347, 1359 (9th Cir. 1980); Brose v. Sears, Roebuck & Co., 455 F.2d 763, 765 (5th Cir. 1972); Ansul Co. v. Uniroyal Inc., 306 F. Supp. 541, 566 (S.D.N.Y. 1969).

91 Notably, the *Pequignot* court did not analyze the foregoing arguments in support of construing section 292 to exclude marking with expired patents. These unaddressed issues also highlight some potential weaknesses in the court's opinion.

92 *Pequignot*, No. 1:07-cv-897, slip op. at 20 (E.D. Va. Aug. 25, 2009).

93 Stauffer v. Brooks Bros., Inc., 615 F. Supp. 2d 248 (S.D.N.Y. 2009).

94 At least one influential organization has proposed just this: in its 2009 Board resolutions, The Intellectual Property Owners Association (IPO) resolved that it, "in principle, supports legislation to repeal the qui tam cause of action of Section 292(b) of the patent statute for false marking of products." http://www.ipo.org/AM/Template.cfm?Section=Board_Resolutions_and_Position_Statements&Template=/CM/HTMLDisplay.cfm&ContentID=21672 (last viewed August 18, 2009).

95 One such proposal would be to provide a "government-intervention as right" framework, mirroring the FCA's requirements for notice of the USA, rights of control/intervention, etc., and defining the injury for Article III standing purposes to be the harm done to the public trust in and reliance upon the patent system, which public trust is represented and enforced by the government with or through private relators.

96 One approach would be to clarify the statutory wording to make clear that marking with expired patent numbers is not false marking; provide for searchable public indices (e.g., on the WWW and in the Federal Gazette) that publish patent expirations as they occur; supplementing the existing notice framework that already notices expirations due to nonpayment of maintenance fees; and making certain that the notice function accounts for litigation holdings after exhaustion of appeals, PTA, PTE, re-issues, and re-exams, with a procedure for patentees to contest any wrongful/erroneous notice.

Nashville in Africa: Intellectual Property Law, Creative Industries, and Development

*by Mark Schultz and Alec van Gelder**

Nashville, Tennessee, was once a struggling city in one of the poorest regions of the United States. Early 20th century policymakers pinned the city's economic hopes on industrial development founded on access to raw materials and large, government-funded public works projects. These hopes were never fully realized, but Nashville found success another way—through its creative industries.

Today, Nashville enjoys enviable economic success as "Music City U.S.A." It is home to a multi-billion dollar country music industry, employing tens of thousands of people, and a thriving, diversified economy.

Nashville's ascent serves as an encouraging example of how creative industries can make much from little. Creative industries offer considerable potential as drivers of economic development, since they require relatively low levels of technological, physical, educational, or financial infrastructure.

In this article, we consider how Nashville might serve as a model for cultural and economic development in Africa. Like Appalachia and the Southern U.S. in the early 20th century, many African countries have rich musical traditions and abundant talent. The popular music industry in Africa has vast potential, but there are not yet "African Nashvilles." We consider the barriers to their development and how they might be removed.

Although the presence of a successful creative industry is neither sufficient nor necessary to a thriving economy, it is an example of the type of private, locally-based entrepreneurial effort that poor economies need to foster. We begin by examining how "Music City U.S.A." emerged from circumstances that in some ways are similar to those that persist in parts of Africa today. We then seek to explain Nashville's success, drawing on the literature on creative clusters. We then consider whether modern day Africa could achieve similar success, considering both the great promise and difficult circumstances of the African music industry. We conclude by offering some suggestions of the kinds of reforms that might help establish creative industries and thereby promote economic and cultural development in sub-Saharan Africa.

* *Mark Schultz is an Associate Professor at the Southern Illinois University School of Law in Carbondale, Illinois, U.S.A. Alec van Gelder is Network Director at International Policy Network, London.* *This article has been published in earlier, lengthier versions as a monograph by the International Policy Press in London and as an academic article by the Kentucky Law Journal. The authors wish to thank Steve Margolis, Paul McGreal, Susan Liemer, Julian Morris, the participants at the 2008 Society for Economic Research on Copyright Issues Annual Conference, and Mark Schultz's seminar at DePaul for their helpful comments. Special thanks to our colleagues Mike Andrews and Franklin Cudjoe, who provided insights into local conditions in Africa, and to the global network of think tanks that have sponsored and helped to publicize our work. All opinions expressed herein are those of the authors.*

I. Popular Music as a Grass Roots Economic Development Strategy

The city of Nashville is nicknamed "Music City U.S.A." for good reason.[1] While the United States is blessed with several locales where popular music thrives both creatively and financially, there is no other city or region where music is more central to the local economy and identity.[2] A recent study estimates that music contributes over $6 billion a year to the local economy.[3] According to the same study, 20,000 jobs in the Nashville area are directly related to music production, accounting for over $700 million in annual wages, with Nashville home to "80 record labels, 130 music publishers, more than 180 recording studios, 40 national producers of ad jingles, 27 entertainment publications and some 5,000 working union musicians."[4] The study estimates that 35,000 additional local spill-over jobs exist because of the music industry in fields such as music-related tourism.[5] Since Nashville is a relatively small city, with a population of 1.2 million people, the music industry's benefits to the local economy are particularly significant.[6]

A. The Nashville Story

Nashville's country music success was not part of any grand centrally-orchestrated blueprint. Rather, it resulted from the actions of a few early 20th century music business pioneers—the sort of grass roots entrepreneurs who the more forward-looking development economists would laud.[7] Important lessons can be drawn from their actions and from the conditions that aided their success.

In the early 20th century, Nashville and its surrounding region once faced economic challenges similar to those faced by today's poorest countries. During the first part of the 20th century, the American South was far less developed than other regions of the United States.[8] The South's industries were significantly less advanced, with economic activity focused mainly on agriculture or local industry.[9] "Nashville shared the South's dilemma of being unable to move beyond local resources and local markets."[10] Conditions in the rural regions outside Nashville were particularly daunting. In the early 1930s, per capita incomes in the Tennessee Valley were about 40 percent of the U.S. national average.[11] Subsistence, tenant farming was still common. As one scholar recently noted, the South's socio-economic system was considered so backward that the contemporary mass media "casually referred" to it "as 'feudalism.'"[12] "The low incomes and general economic distress of the South made it impossible to develop educational institutions and public health facilities to the standards attained in most other sections of the United States."[13] In fact, 30% of the population in the Tennessee Valley was infected with malaria in the early 1930s.[14]

Nashville initially pinned its economic hopes on easy access to natural resources, hoping that it would become a

major manufacturing center for steel and textile production.[15] Yet these aspirations were never fully realized: the mere presence of raw materials did not translate into hoped-for manufacturing prowess.[16]

Hopes for Tennessee's development then turned to the sort of grand projects that have been so favored by Western development agencies in Africa. In 1933, the Tennessee Valley Authority ("TVA") was established to build massive infrastructure projects to promote development in the Tennessee region. The TVA built dams, power plants, and other public works, while promoting agricultural and educational programmers. For decades after WWII, the TVA was enthusiastically embraced as the model for development projects around the world—from Aswan Dam in Egypt to the Three Gorges Dam project in China, cement factories in Nigeria, and aluminum smelters in Ghana, governments have seen TVA-style massive public works as a development panacea. Since then, the development community's enthusiasm for the TVA and its descendents has declined precipitously, because, as one scholar observes, many experts now "find these efforts questionable, if not dangerous.... Too often... these vast undertakings hurt the people in their path as the grand, 'high modernist' visions at the core of these programmers ignore people's needs, values, experience, and knowledge."[17]

Compared with the grandiose TVA, Nashville's "Music City" grew from far more modest seeds. The development of the country music business was not the product of detailed plans orchestrated by a centralized authority. Rather, initial events hinged on the private actions, skills, and insights of a handful of individuals: the musical talents of country music pioneers such as Jimmie Rodgers and the Carter Family and the business acumen of entrepreneurs such as Ralph Peer, a record producer and aspiring music publisher. Their success may seem almost accidental, a case of enterprising performers and businesspeople being fortunate enough to find and shape an untapped popular taste for their product. But there is far more to the story than luck.

The initial factor in Nashville's musical success was the unique musical tradition of the rural South and nearby Appalachian region. Music was an essential part of people's daily lives. Folk music was one of the chief forms of family and community entertainment and central to religious worship in the region's many, diverse Protestant religious denominations.[18] A mix of history, geographic isolation, and culture had combined to foster a musical heritage that was both unique and popular.

By the mid-1920s, people were beginning to see that this folk music tradition might be converted into a more commercial art form. Some performers, such as Jimmie Rodgers, were playing engagements throughout the region, and a few, such as "Uncle" Dave Macon, had broken out of local engagements to tour the East Coast.[19] Two high-powered radio stations that reached large parts of the country, WLS in Chicago and WSM in Nashville, began broadcasting weekly live country music shows that quickly proved to be very popular. (WSM's *Grand Ole Opry Radio Show,* is still running today.) Perhaps most consequentially, a New York record producer, Ralph Peer, co-produced what is known as the first commercial country music recording, Fiddlin' John Carson's "That Old Hen Cackled and the Rooster's Goin' to Crow," for OKeh Records.[20] Despite the poor quality of the recording, it quickly sold out, convincing Peer of the potential popularity of country music.[21]

Fortunately for him and the musicians whose careers he launched, Peer was able to secure the financial backing to pursue his insight that there was a market in country music. He left OKeh, formed his own publishing company, and entered into a joint venture with the Victor Talking Machine Company.[22] In the summer of 1927, Victor supplied Peer with $60,000 in financing for an expedition to the American South to record country music.[23] Up to that point, few rural, Southern artists had been recorded, as the then-current analogue recording equipment was bulky and impractical to move, thus necessitating a prohibitively costly trip to New York.[24] With the advent of electronic recording technology in the late '20s, Peer was able, at a significant but feasible cost, to take the equipment closer to the talent. Peer used the money to send out a team of advance scouts to find musicians and ship recording equipment to Bristol, Tennessee, where he set up a temporary recording studio.[25]

Peer's recording session in Bristol established the commercial viability of country music. Johnny Cash called the Bristol Sessions "the single most important event that ever took place in the history of music."[26] Over the course of 15 days, Peer and his crew recorded a total of 76 songs performed by 19 different acts, capturing "an almost perfect representation of early country music: fiddle and banjo tunes, old traditional ballads, gospel music, old popular and vaudeville songs, and rustic comedy."[27] The recordings quickly turned into commercial success, leading to the discovery of both a mass market for the genre and the first country music superstars, Jimmie Rodgers and the Carter Family. "In a 3-month span a year after the Bristol sessions recordings first went on sale, Peer's Southern Music publishing company earned $250,000 in royalties."[28] Later that year, Peer recorded Rodgers singing *Blue Yodel (T for Texas)*, a huge hit selling a million copies.[29]

While the Bristol Sessions were a gold mine for Peer and the most successful performers, they were an important economic opportunity for everybody involved, especially as few other opportunities existed in Nashville. When Peer used stories in the local media to advertise the amounts he was paying performers,[30] the response was overwhelming. Peer "was deluged with long-distance calls from the surrounding mountain region. Groups of singers who had not visited Bristol during their entire lifetime arrived by bus, horse and buggy, trains, or on foot."[31] The royalties that Peer and Victor paid were substantial by the measure of incomes of the day: Peer typically paid a recording fee of $50 per song and a royalty of about 2 ½ cents per record side sold.[32] By comparison, one of the major local industries, coal mining, then paid about 76 cents an hour.[33] Thus, a single recording session paid far more than an average week's wages for a coal miner, and sales of fewer than 6,000 copies of a record would produce royalties equivalent to a month's wages.[34]

Once the Bristol Sessions established the commercial viability of country music, other music business entrepreneurs seized upon the opportunity Peer had uncovered. They observed that Nashville offered the advantage of local talent

and expertise.[35] Pioneers such as Roy Acuff, Harold Rose, and Owen and Harold Bradley set up publishing houses, recording studios, and record labels in Nashville.[36] In their wake, even more music-related businesses flocked to Nashville: other record labels, recording studios, and music publishers; collecting societies; record stores; performance venues; and other essential institutions. By the 1950s, these businesses were so heavily concentrated in one of Nashville's neighborhoods that it became —and remains—known as "Music Row." Those early institutions started a helpful dynamic of competition, cooperation, and shared expertise that continues to this day.

B. The Lessons of Nashville for Less Developed Countries

What country wouldn't envy the economic benefits provided by Nashville's music industry? Indeed, a team of World Bank staffers created the Africa Music Project to promote the African music industry with Nashville partly in mind, saying that their "dream" was that "African countries would create their own Nashvilles."[37] Creating such centers of economic activity is more than just a dream; it can also be a viable development strategy given the right circumstances. Although Nashville's success cannot simply be transplanted wholesale to different countries and times, it does yield compelling lessons.

Just as the early country music recording industry offered an attractive alternative to Tennessee's workers, today's creative industries offer a similarly alluring alternative to workers in poor countries. Most jobs in poor countries are labor-intensive and relatively unproductive; for example, 70 percent of employment in sub-Saharan Africa is agriculture-related, much of it subsistence farming.[38] Moreover, much agricultural work is seasonal and weather-dependent. Work in creative industries such as the recording business is likely to yield a much higher return, as the productivity and profitability advantages are vast. There is thus ample opportunity, in principle, for talented individuals to increase their income by investing some of or all their labor in creative fields such as music.

Recent work on economic development has focused attention on the importance of local centers of excellence and expertise to the economic prospects of a region. As Michael Porter has written in his widely-heralded work on competitive strategy and economic policy, regions and nations are most likely to prosper in the global economy if they foster strong "clusters" of economic activities, such as the country music "cluster" in Nashville.[39] According to Porter, clusters are "geographic concentrations of interconnected companies, specialized suppliers, service providers, firms in related industries, and associated institutions (e.g., universities, standards agencies, and trade associations) in particular fields...."[40]

Clusters of economic activity provide a sustained competitive advantage through a self-reinforcing dynamic of cooperation and competition among related businesses. They exploit economies of scale and build up networks of skills, knowledge, and business relationships. Both employees and ideas circulate among competitors, building up a collective advantage over firms from outside the region. In addition to Nashville, other well-known examples include the microelectronics industry in Silicon Valley, high-performance auto companies in southern Germany, high-end fashion shoes in Italy, and wine industries in South Africa, France, and New Zealand.[41]

Celebrated examples such as Nashville and Hollywood show that clusters of competitive advantage can form around creative industries. However, there is no reason to believe creative clusters are merely the product of wealthy societies. Indeed, Tyler Cowen has documented how specific cultural sectors in poor countries can thrive under globalization, as local creativity benefits from exposure to new technology, wealth, and outside creative influences.[42] Examples he cites include Congolese soukous music, Haitian painting, Jamaican Reggae, Persian textiles, Cuban music, and amate painting from Guerrero, Mexico.[43]

Creative clusters can be particularly powerful drivers of development in poor countries for several reasons. First, they play to local strengths, taking advantage of knowledge, skills, and forms of expression that arise from local culture, and are thus, by definition, largely unique and non-duplicable. Second, for the most part they do not require cutting-edge technology, large capital investments, or a robust infrastructure.[44] Third, although creative work often requires a significant personal investment in training and development, it typically does not require the sort of extensive formal educational system that still remains unavailable to the poor in many less developed countries.

Vibrant creative industries can provide benefits beyond the economic activity they create directly. A thriving creative cluster stimulates investments elsewhere. For example, Nashville's music industry has engendered a substantial tourism industry, with the additional transportation and lodging infrastructure that comes with it.[45] For the music business, Nashville's advantages outweighed any challenges presented by how backward Nashville may have been at that time. Nashville thus became the home of a thriving creative cluster.

Success also provides a large morale boost to people, economies, and cultures. Nashville's success provided such a benefit to a downtrodden region, giving credibility to an obscure, once-dismissed type of music. As Peer wrote of Jimmie Rodgers, the "impetus which he gave to so-called hillbilly music... set in motion the factors which resulted in making this sector of the amusement business into a matter of world-wide importance and a source for a high percentage of our popular hits."[46] This is the same boost that resulted from the actions of artists and entrepreneurs that helped to kick-start the development of the reggae industry in Jamaica.

Entrepreneurship is at the center of the Nashville story. The large government-directed projects of the era, such as the TVA, played little role in the growth of Nashville's creative industries. Instead, the birth of Music City depended far more on the talents and efforts of individuals such as Jimmie Rodgers, Ralph Peer, and the artists and entrepreneurs who followed in their footsteps.

The central role of private action to building creative clusters is both bad news and good news for policymakers. The bad news is that there is little governments can do to ensure success for the creative industries. The good news is that these risks can be placed on the shoulders of private parties rather than resource-strapped governments. Provided they can foresee rewards, artists and creative industries willingly take these risks.

Governments play a lesser—but essential—role in providing the right institutional framework for creative industries to thrive through the enforcement of contracts and institutions, such as copyright.

Copyright was an essential ingredient in Nashville's success. For Peer, it was the very foundation of his entire Bristol enterprise. Peer left his job with OKeh records to found his music publishing business, while convincing Victor to invest a huge sum in the Bristol Sessions. Peer took only $1 per year in salary from Victor, but in exchange he obtained the right to control the copyrights in the compositions he recorded. The property rights created by copyright enabled these transactions, giving Peer and Victor a reason and focus for their entrepreneurial activity.[47]

Peer understood—and helped establish—the entrepreneurial and cultural value of copyright to the young recording industry. He was among the first in the pop recording industry to see the win-win potential created by copyright: "Peer's genius lay in structuring his publishing company based on royalties, making copyrights profitable for the artist as well as himself—the financial model of the modern music industry."[48] He was also the first record executive to encourage his performers to avoid old, copyrighted standards and public domain works in favor of new compositions.[49] The new compositions had the dual virtues of being copyrightable and more culturally relevant (and thus more commercial) than older works.[50] Indeed, once Peer and others showed the value of new, copyrightable material, commercial incentives motivated rural Southern musicians to abandon folk music for royalty-producing works.[51]

The story of Nashville shows just how much a region's culture can aid its economic development, given enough talent, adequate incentives for entrepreneurs, the right laws, and supporting institutions. Although Nashville resides in one of the world's wealthiest countries, it was once a struggling city in the U.S.A.'s most underdeveloped region. Nashville now has a thriving, modern economy. The country music industry played an essential part in Nashville's transformation into a world-class city, sustaining it through difficult times and contributing enormously to its modern economy, cultural identity, and business reputation.

The question, then, is what it will take to create other "Music Cities" in less-developed countries. As the example of Nashville illustrates, certain conditions must be met before local talent and entrepreneurship can result in creative success and economic development.

II. Where are the African Nashvilles?

While Nashville provides an encouraging example of how creativity and entrepreneurship can help lift people out of poverty, one might reasonably ask whether it can be replicated in Africa and other poor parts of the world. Some of the ingredients are certainly present in Africa: rich musical traditions; entrepreneurs who have sought to develop those musical traditions into a profitable business enterprise; and copyright laws in most of Africa. Unfortunately, realizing the dream of creating African Nashvilles has proven challenging.

A. Abundant Talent, Abundant Potential

Nashville is hardly the only, and certainly not the greatest, concentration of musical talent on the globe. For example, a popular guide for adventurous tourists touts Bamako, the capital of Mali, as one of the world's "musical hot-spots": "The West African city's anarchic collection of neighborhoods sprawls from the Niger river in Mali, filled with single-story dwellings and women cooking in their courtyards on charcoal braziers. The place feels like one big village, with music everywhere."[52] The guide continues:

> [N]one of this should be surprising, given Mali's 600-year-old musical tradition. And in the last 15 years, artists like singer Salif Keita and singer-guitarist Ali Farka Touré have shot to international fame, making Mali the centre of West African music and Bamako one of the premier places on the planet to hear it live.[53]

While Mali's musical traditions may be unusual in their richness, many African nations have their own traditions that can be traced back decades if not centuries, from Rai music in Algeria to Highlife and Afrobeat in Nigeria and Ghana; from Mbablax in Senegal to Soukous in Congo and the Democratic Republic of Congo; from Township jive in South Africa to Chimurenga music in Zimbabwe.[54] So great is the influence of African creativity that many of the world's most popular forms of music, from blues and jazz to gospel and reggae, from soul and funk to R&B and hip hop, rely heavily on the rhythms, melodies, and musical traditions of Africa.[55]

The potential of Africa's music industry is widely recognized. As a recent report from the United Nation Conference on Trade and Development (UNCTAD) observed, the creative industries offer some of the best prospects for high growth in least developed countries.[56] One researcher estimated that the Ghanaian music industry alone could generate $53 million a year from foreign sales if local conditions were more amenable to supporting creativity.[57]

In a few spots in Africa, this potential is being realized, at least partly. South Africa's music industry is strong and diverse, with internationally-known musicians such as Hugh Masekala, Miriam Makeba, Ladysmith Black Mambazo, Ray Phiri, and the Soweto Gospel Choir. Elsewhere in Southern Africa, the Zambian music industry is being revived after a near-total collapse in the 1990s.

The Zambian experience illustrates how pioneering entrepreneurs can ignite a creative industry, given the right conditions. First came a new copyright law in the mid-'90s.[58] Then, in 1999, a new Zambian record label, Mondo Music Records, sparked a revival. Much like Ralph Peer showed the way for Nashville's early country music pioneers, Mondo showed the way for other entrepreneurs.[59] "[T]here has been exponential growth in the amount of Zambian music being produced in the last seven years, and also in the consumption and the appreciation of it. Right now, Zambian music dominates... local radio, and [is] also becoming a little bit noticed outside of the country."[60] Mondo's founder, Chisha Folotiya, recognizes the potential value of a creative industry to his country's economy and to creative individuals. He says "we want Zambian music to

contribute towards the economic development of our country. On a small level, as individual artists, retailers, producers, choreographers, dancers who are involved in the music itself, and also at the macroeconomic level, the entire retail sector, and manufacturing sector."[61]

Such bright spots in African popular music show what is possible if the right conditions are created. The possibilities for creative industries are both readily visible and seemingly endless. Such creative clusters present the sort of grass-roots opportunity that is increasingly seen as fundamental to economic development.[62]

B. The Reality

Unfortunately, these hopeful examples remain isolated. Despite widely acknowledged potential, the music industry remains beleaguered in most African nations.

Bamako, Soweto and Lusaka notwithstanding, there are as of yet no Nashvilles in Africa.

African creativity remains an underappreciated and underexploited resource. Rarely do creative sectors contribute more than 1 percent of the relatively low GDPs of any African country.[63] Africa's share of trade in cultural goods constitutes less than 3 percent of the global total.[64] One researcher estimates Africa's share of the world market for sound recordings at a mere 0.4 percent.[65] By comparison, creative industries in wealthy countries employ millions and contribute significantly to national economic production—over 11 percent of GDP in the United States, for instance.[66] European creative industries are the fastest growing industries across the continent and employ over 4.7 million people, according to 2005 figures from UNCTAD.[67]

Despite a few celebrated examples on the world music scene such as King Sunny Ade, Ali Farka Toure, and Youssou N'Dour, African music has yet to become a successful export industry. For example, Senegal is justly celebrated for its illustrious artistic history and for relatively well-known stars such as N'Dour who have been able to transform their musical roots to international success in "outside markets."[68] However, the Africa Music Project estimated that in Senegal only "one dozen of the estimated 30,000 artists enjoy international sales and publicity."[69] Another study attempted to gauge the number of "internationally recognized [music business] celebrities" in other African countries (with "international" connoting success anywhere outside of their home country, rather than global stardom).[70] The relatively successful South African music business boasted a high of 22 percent, but the next most successful country was the Democratic Republic of Congo with only 8 percent.[71]

To the extent that African creators do succeed, their success often fails to produce economic benefits for their home countries. For example, most African music is recorded in either London or Paris—largely depending on where the artist in question originates.[72] Those products are also most often consumed in Western countries as well. The situation is similar for other creative sectors: "Half of a Yellow Sun," a tale of a bloody civil war in the early 1970s written by Chimamanda Adichie, is one of the most popular African books in the first decade in the 21st century. But it has sold just 5,000 copies in the author's native Nigeria, as opposed to at least a quarter of a million in Britain, where Adichie has won critical acclaim.[73]

In stark contrast to the high incomes earned by many musicians in wealthy countries, African musicians are often poorer than their fellow countrymen. The Africa Music Project estimated average income for musicians in Senegal was $600 per year—15 percent lower than the country's GDP per capita.[74] The study further observed that "eighty per cent of Senegalese musicians are either unemployed or underemployed."[75]

C. What's Missing

The failure of Africa to produce healthy creative clusters is disappointing—and may in part explain their lack of economic development. It is certainly not for a lack of entrepreneurial talent. As we noted above, the African Ralph Peers exist, but most are hampered by local policies and practices.

Among the major obstacles to the emergence of successful creative clusters in Africa are:

- Difficulties enforcing copyrights against piracy;
- Government control of copyright collection agencies;
- Irrational, burdensome taxation.

We now consider each of these in turn.

Difficulties enforcing copyright against piracy

Piracy of music is a serious problem in sub-Saharan Africa. Pirated versions of creative works represent *at least* 25 percent of the entire marketplace across Africa.[76] That figure is as high as 90 percent in some West African countries.[77]

Piracy deprives creators and legitimate distributors of sales. As we discuss below, by undermining the potential for downstream revenue, it also prevents creators from securing capital to finance their work—and, indeed, undermines the emergence of a local recording industry.

In spite of nominally strong copyright laws, enforcement is a major problem.[78] According to Nwauche, this results from a combination of: "[i]nadequate funding of enforcement agencies; lack of trained and properly motivated staff [of copyright offices]; stakeholder apathy in the enforcement of their rights; a weak institutional base; poorly trained and paid enforcement (police, customs, and specialized institutions) agents; a cumbersome and tardy judicial system; and unorganized stakeholders."[79]

The lack of enforcement has effectively empowered the pirates, who are better organized and more successful than their victims and probably act in cahoots with the public enforcement agencies. In Senegal, the Africa Music Project observed that "pirates have more means at their disposal than those responsible for policing them." They further observed that "criminals guilty of wide scale commercial piracy are often pardoned because of well-placed connections within local government."[80]

In places where life is already hard, such conditions are not merely unfair—they have tragic consequences. Alhaji Sidiku Buari, President of the Musicians Union of Ghana, has described how all of these unfortunate realities conspire against creators in Ghana: "A musician will do his music and somebody else reframes it and gets all the money in his pocket.... Our musicians have no social security, no insurance, no pension

scheme and most of them die as paupers."[81] Creators' hard work and talent are betrayed by a poor institutional climate. As Orrack Chabaagu, Director of EMI South Africa has observed, "It is unfortunate that after one has gone through thick and thin to produce his music, he does not live to enjoy its results because of piracy."[82]

Piracy also hampers individual musicians from securing capital to finance their creative work or other ventures. Their primary potential assets—their copyrights and the revenue streams that should result from them—are effectively worthless. As a result, musicians are unable to obtain loans from local financial institutions, thus forcing them to pay for instruments, recording time and other business expenses up front and out of their own pockets.[83]

The problem is the same one recognized by Hernando De Soto in his *Mystery of Capital*: without clearly defined, readily enforceable property rights, the poor have no assets with which to secure loans and other capital. They are thus prevented from climbing the economic ladder.[84] The experience of African musicians shows that De Soto's insight regarding physical property applies to IPRs as well.

Ineffective enforcement of copyright has led to a host of other unintended consequences, including pushing the production of African music out of Africa. For example, in the late 1980s most legitimate music businesses in Ghana, including production houses and record factories, were shut down due to competition from piracy. Moreover, "Ghanaian musicians fled to other countries, thereby creating a vacuum in the industry."[85] Zambia's experience was the same, where a once "massive music industry" faded to almost nothing in the early '90s, as local record production shut down because of pirated cassette tapes from abroad and a shortage of capital.[86]

There are many factors that might weigh toward local production, including lower labor costs, local knowledge of and familiarity with local tastes, the existence of some production infrastructure, and convenience. Yet in spite of all of these potential advantages, recording and production in the UK or France is often still the wiser choice, which exacts a further opportunity cost to local economies in lost jobs, lost local income, and lost spending on related goods and services.[87]

The tolerance of rampant piracy also thwarts the development of regional markets for music. Piracy not only deters non-African companies from investing in the development of the African marketplace, but also acts as a barrier to intra-Africa trade as well.[88] Kenyan copyright experts estimate that of all the content emanating from creators across the border in Tanzania, material from just two artists is legitimately sold in Kenya, despite both countries sharing at least two common languages, English and Swahili.[89] The forgone investments represent an important loss for two creative cultures that could potentially feed off each other in order to develop new techniques, styles, and markets. The missed opportunity suppresses the creation of jobs and wealth in desperately impoverished East Africa.

Another consequence of the inability to enforce copyright in Sub-Saharan Africa is the lack of downstream royalty payments, which seriously undermines long-term investment. While there is a place for one-hit-wonders, most artists and record labels are motivated to produce, market, and distribute well-crafted recordings that will sell for years to come. The motivation comes in large part from the ability of copyright holders to earn royalties on the sale of recordings.

The inhospitable environment created by piracy for the music industry in Africa leads to a short-term focus.[90] Outside South Africa, there are few substantial, financially stable record labels.[91] The recording industry tends to be a fragmented, fly-by-night business with irregular distribution. In Ghana the consequence has been a "peculiar and unique" distribution system, whereby many retail outlets sell the work of only one particular record label, making it frustratingly difficult for fans to find recordings.[92]

Recording artists respond to this lack of stability by trying to collect as much revenue as possible in the short-term, exchanging their future (and largely theoretical) royalties for one-off payments from record companies.[93] They then offer very similar work to other recording companies in exchange for further one-off payments. Record companies in turn anticipate that duplicative output and competition with piracy will lead to a brief sales cycle, and thus tend to underpay artists for their work, which further perpetuates the supply of largely unimaginative material. While this vicious cycle of deteriorating quality may seem to be against the interests of all those involved, the institutional environment makes for little alternative.

Government control of and interference in copyright collection agencies

A well-functioning, robust music industry provides many creative and financial opportunities for creators besides selling recordings and earning royalties from sales. Many of these opportunities are facilitated by licensing. For example, composers and musicians often license their work for broadcast on radio and television, for cover versions, for inclusion in movies, television shows, and advertisements, and increasingly, for derivative uses, such as rap songs, re-mixes, and ring-tones. All of this activity expands creative clusters, making them even more significant and beneficial to creators and the local economy.

Unfortunately, most African musicians are not able to enjoy the benefits of such downstream uses of their creations. Some of this failure results from shortcomings in the legal system already discussed, but this is exacerbated by the lack of effective, dependable collective rights organizations (CROs). In most wealthy countries, CROs secure payment to artists for various small uses of their works, such as when records are played in nightclubs or on the radio. The small size of each potential transaction makes it uneconomic for individual parties to pursue agreements or to enforce their rights on a case-by-case basis. CROs remedy this problem by granting blanket licenses in exchange for aggregate subscription fees, which are then allocated to copyright owners in proportion to the use of their works.

Unfortunately, it appears that many CROs are not doing their jobs well in Africa. There are widespread complaints that the amounts bear no relation to the actual frequency of play by radio stations or other public venues.[94] For example, in Senegal, royalties are supposed to be collected by the Bureau Sénégalais

du Droits d'Auteurs (BSDA), but it rarely succeeds in rewarding artists appropriately, as royalty payments are inconsistent at best.[95] In part the problem is that radio stations and other music distributors in Senegal withhold sales information from the BDSA, making it impossible to determine how to allocate royalties.[96] Artists also accuse the BDSA of over-charging for its services.[97]

Similar problems plague musicians in both Ghana and Kenya.[98] Meanwhile, musicians claim that officials of the Copyright Society of Ghana (the Ghanaian CRO) and the government copyright office have corruptly diverted the royalties they do collect.[99]

A significant part of the problem is that many CROs are run by the government or are government-sanctioned monopolies. Such arrangements undermine the effectiveness of CROs by making them less accountable to their members. Such is the case in Senegal, where the Ministry of Culture controls the BDSA, and in Kenya, where the "Music Copyright Society of Kenya" is the only collecting house sanctioned by the Kenya Copyright Board.[100] CROs in Nigeria can only operate with the explicit approval from the government-controlled Nigeria Copyright Commission.[101]

Such restrictions on competition undermine the incentives for collecting agencies to respond to artists' concerns. As the authors of the Africa Music Project observe, "distribution [of royalties], when it takes place, is a political process rather than an objective one."[102]

Government involvement with CROs can also threaten the independence of musicians. Artists in Ghana have accused the Chairman of the Ministry of Culture-controlled Copyrights Office of withholding payments from artists in an attempt to influence the content of their music.[103] Such actions amount to a gross violation of the right to free speech.

Similarly, a station manager for a radio station in Dakar has said government stipends earmarked for local radio stations tend to arrive "only during election time."[104]

Perhaps it is not surprising that African governments should use their power in this way. According to the most recent Freedom House ranking of political freedom, just eight African countries were classified as "free."[105]

Burdensome Taxes

Many governments in Africa impose taxes specifically targeted to musical instruments and other aspects of the music industry.[106] These taxes make a tough business that much tougher.

One particularly troubling target of taxation is live performance. Because of the other difficulties faced by the music business in Africa, musicians often have no option but to scrape a living from live performances. Unfortunately, the very nature of live performances (at least those that have some chance of success) is that they are centrally organized and well publicized. That makes them natural targets for the tax collectors.

One example of such burdensome taxation comes from Ghana, which has recently introduced an arbitrary and confusing tax on tickets sold for live performances. Previously, a 25 percent surcharge was imposed on ticket sales. Now, an indeterminate Value Added Tax is levied.[107] The situation is similar elsewhere: a hefty 25 percent duty and an additional 16 percent Value Added Tax is levied on live performances in Kenya.[108] The results are predictable and are well described by "Dou Dou" Sow, a Senegalese musician: "[T]here are fewer live performances today than in the old days because there isn't enough money [to perform]."[109]

Anecdotal evidence from the Africa Music Project illustrates how complicated tariffs were levied on imported musical instruments, pushing the cost of putting on live performances out of reach for artists struggling to make it from one show to the next.

Various other government-imposed barriers prevent small-scale entrepreneurs with few resources—the epitome of the African artist—from engaging in productive economic activity. Long waiting times and overly-complicated requirements put what should be simple tasks, such as registering property and trading with foreigners, out of reach for poor, struggling artists. These frustrating realities are quantified by Doing Business, a World Bank project that measures the difficulty of engaging in entrepreneurial activities. Doing Business estimates that it takes an *average* of 33 procedures and 780 days to enforce a contract in Senegal. It takes a further 6 procedures and 114 days just to register a property.[110] Clearly these barriers affect the entire economy, but they are nevertheless particularly pernicious to artists, who already scrape by with fewer resources than the average citizen.

III. What to do? Policy Recommendations

Africa's experience bears out Nobel laureate economist Douglass North's truism that "institutions matter." Creative clusters require not only talent and entrepreneurship, but also effective institutions that enable the talented and the entrepreneurial to flourish. In this monograph we have identified a number of institutional failings which, if corrected, would, we believe, lead to the flourishing of creative industries in Africa. In order to ensure strong and credible institutions that support the development of local creative industries, we recommend that governments:

- Enact, implement, and enforce effective copyright laws;
- Reduce intervention in royalty collection;
- Encourage private cooperation; and
- Reduce taxes and regulatory burdens.

A. Enact, Implement and Enforce Effective Copyright Laws

Property rights are generally far more effective when people are able to demonstrate ownership easily and potential licensees and buyers can locate owners quickly and at low cost.[111]

One of the simplest and most effective steps that can be taken to improve conditions for musicians and other creative artists would thus be to establish a simple means of registering ownership of creative works, as well as for recording assignments, licenses, and other transfers. There is much to learn from the experience in wealthy countries, as they often have complicated and inefficient systems for tracking ownership, but are able to afford the sizeable costs of overcoming these defects.[112] People in less-developed countries are less likely to be able to afford the costs of investigating ownership through these archaic systems,

so the best option would be to employ modern information technology to ensure titling systems are quick and easy to use.[113] Registering and tracking ownership through a system that uses this modern technology may perhaps be more costly initially, but will likely reap significant rewards in the cost and time saved for all those involved in creative production.

Aside from employing the optimal system for tracking and registration, copyright laws should also be well equipped to provide creators with the best opportunities to combat piracy. As we have noted, this has a devastating impact on creators in less-developed countries.

Generally speaking, copyright laws must provide for swift, affordable means of immediately halting the distribution of pirated works. Laws should provide for preliminary or "interlocutory" injunctions—in lay terms, this means copyright owners can force the pirates to cease distribution at the start of the case, rather than waiting for a trial and its conclusion. In some cases, it is necessary to provide swift, temporary emergency relief, even if the alleged infringer cannot be found or is absent from court. For example, if an imported shipment of pirated CDs is about to flood the market, the copyright owner should not be required to find the foreign infringer before obtaining any relief. For these remedies to be effective, copyright owners require the cooperation of law enforcement to seize and impound the infringing products. These remedies stop pirates from making money and evading liability while the IP owner suffers lost sales and the cost of litigation while attempting to find infringers and prove liability.

Copyright laws should provide for (and courts should routinely award) financial remedies that make piracy too financially risky. In many cases, remedies that merely deprive the pirate of profits or even of gross revenue are not useful deterrents. In such situations, damages are merely an additional business risk calculated against the likelihood of getting caught and thus are not an effective means of tackling the critical problem.

It should be made clear to the pirate that he would stand to lose significantly more than the potential gains of infringing copyrighted works. Laws thus should provide for payment of the copyright owner's attorneys' fees and either a multiple of actual damages or what are known as "pre-established" or "statutory" damages—significant amounts set by statute that aim to exceed the typical value of an infringement.[114] If sufficiently sizeable, the threat of such compensation should act as a deterrent to potential pirates.

Additionally, copyright laws ought to enable trade associations to combat infringement on behalf of their members. Sometimes, copyright owners cannot afford to pursue legal action. In other instances, a pirate might be infringing the works of many different copyright owners, with each individual infringement too small to pursue. To address such problems in many jurisdictions trade associations are able to act on behalf of their members.[115] However, in some African countries, only registered collecting societies can bring such suits on behalf of members,[116] while in others only the copyright owner can bring suit.[117] These restrictions make it more difficult to tackle piracy. Extending the rights to bring action to a wider range of potential litigants, especially to trade associations (with the consent of the allegedly affected party), would greatly aid the battle against piracy.

Civil enforcement by private parties ought to remain the first and foremost means of defense against piracy. Those with the greatest interest in stopping piracy—the creative industries—are likely to be most effective in monitoring and halting the trade in pirated goods. However, they do require help from law enforcement in doing this job, as they play a limited but essential role in combating copyright piracy. To be safe and effective, the seizure and removal of pirated goods from channels of commerce often requires the backing of police or other law enforcement personnel.

The state's most important direct role in preventing copyright piracy is in investigating and stopping large-scale commercial piracy of CDs and other copyright material. Where criminal gangs conduct piracy, individual copyright owners are often no match (especially where trade associations are barred from acting on behalf of copyright holders). Local creative industries also find it particularly hard to combat piracy where foreign counterfeiters flood the market with pirated copies of local creative works. In such situations, government can help by providing and coordinating credible enforcement.

Crucially, any criminal penalties must be appropriate. More is not always better where criminal penalties are concerned. For example, the Chinese head of the Food and Drug Administration was recently executed on grounds of the continued prevalence of counterfeit medicines in China—and under pressure from the flow of sub-standard goods into the United States.[118] While such actions create a strong impression, they likely are counterproductive. If penalties seem excessive to local sensibilities, then local law enforcement and courts are less likely to apply them, thus making the laws less of a deterrent in reality.[119]

Regardless of the content of IP laws, the courts must be well-prepared to enforce them. As Robert Sherwood says, "until judicial systems in developing and transition countries are upgraded, it will matter little what IPR laws and treaties provide."[120] Judges thus must understand and have the right tools to adjudicate the claims of creators. One way to improve the efficiency of courts would be to allow judges in courts of general jurisdiction to volunteer to take IP cases. This proposal is currently being considered in the United States[121] and has also been advocated for courts in poor countries.[122] The advantage of this proposal is that scarce training resources and experience can be focused on a smaller number of judges.

More generally, all participants in the system—law enforcement, private parties, and judges—would benefit from education. They need both technical training regarding the law and a better understanding of the significance of the creative industries and the importance of their enforcement efforts.

An offshoot of this would be through highly visible and credible anti-piracy campaigns, coordinated by industry and, in some cases, government. In part, the objective of such campaigns is to convince offenders they are likely to be caught.[123] Ideally, they would also change public perceptions of the morality of piracy, change buying habits, and make the business of piracy less socially acceptable.

In addition, the other parts of the legal system need to function well. Much of the effective use of IP is based on contracts, so a well-functioning IP system depends on enforceable contracts. In turn, that enforceability depends on courts being independent from political forces, since politicians are subject to intense lobbying from vested interests. Only when courts are independent are they able to apply the rule of law; that is to say, apply clear, abstract principles of law to all parties in a non-arbitrary, non-discriminatory manner.

B. Remove State Intervention in Royalty Collection

A healthy music industry does not exist in a vacuum. Many other industries are to a degree symbiotic, including radio, mobile telephony, and the movie industry. Unfortunately, in many African countries the licensing market does not work well, if at all, making such symbiosis difficult. Below, we outline some changes that would reduce the barriers to establishing effective licensing systems—beyond the observations already made about copyright law and contracts.

As noted above, in many instances licensing is administered by collective rights organizations (CROs). An alternative to the ineffective and often corrupt CROs that currently exist in Africa would be the establishment of independently audited, private, competing CROs.[124] Such independent, competing CROs exist in other countries, including the United States.[125] Private entities are particularly desirable in less-developed countries, as they relieve resource-starved governments from chores such as rate-setting and royalty collection and place the responsibility on the beneficiaries, who have the greatest incentive to expend resources efficiently to get the job done right. Private CROs are also more likely to remedy many of the complaints of African musicians regarding the ineffectiveness and improper influence of government-run or government-sanctioned CROs.

One problem that many African artists have complained about is poor, over-priced service from state controlled CROs.[126] If CROs were privately controlled and subject to competition, they would have stronger incentives to serve their clients responsibly.

Another problem of which African artists complained is that CROs are influenced by political forces when collecting and paying out royalties.[127] By contrast, privately operated CROs, subject to competition and external auditing, would be less inclined to engage in such corrupt activities. Any CRO found to be corrupt would immediately lose its business to competitors, thereby undermining any benefits from such corruption.

In a deregulated environment, it is possible that pan-African or other international forms of CRO would emerge, providing creators with benefits in terms of wider reach and lower costs.

C. Encourage Private Cooperation

Perhaps more important even than CROs, however, are other cooperative efforts and informal social networks. Economic clusters thrive on the cross-pollination that occurs from business collaboration, shared expertise, and employees switching jobs. In the music business, this kind of networking results in creative innovation, technical refinements in recording, and innovations in financing and marketing.

The area of private capabilities is one where developed world donors, civil society organizations, and specialized agencies such as the World Intellectual Property Organization can provide helpful technical assistance. Each country has its own unique circumstances, but there is much to learn from the experience of successful business models, technical issues like recording, standard form contracts, marketing, combating piracy, and other business issues.

This kind of grass-roots, demand-driven assistance is exactly what is called for by the mandates of various international organizations. For example, article 67 of the agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) calls upon wealthy countries to provide assistance in implementing the agreement.[128] Meanwhile, the general consensus resulting from the "Development Agenda" talks was that World Intellectual Property Organization should increase and improve its technical assistance to less-developed countries. While some of these activities may be profitably focused on government officials—for example sharing best legal practices—private sector actors such as creative businesses and individual artists should receive much of the attention.

However, the most productive assistance is likely to result from helping successful businesspeople share their knowledge and expertise with their counterparts in the less-developed world. There are a number of salutary examples of such programmers. One project used experienced world music industry marketing professionals to train a group of African musical professionals before sending them to pursue opportunities at MIDEM, a large international trade fair for music.[129] Another project assisted the organizers of the "Festival in the Desert" music festival in Mali by bringing in successful festival professionals to "twin" with the local staff members who hold equivalent jobs.[130]

D. Reduce Taxes and Regulatory Burdens on the Music Industry

Governments in poor countries sometimes impose excessive and regressive taxes and regulations on creative and innovative activities. While they are hardly alone among the governments in the world in pursuing such policies, the effects in Africa are especially pernicious. As noted above, artists report high taxes on musical equipment and public performance.[131] Such burdens should be reduced or scrapped altogether.

Lower taxes on these activities directly related to the arts incentivize more artists to create and to promote their work directly to the public. Since innovation and creative work have such tremendous positive ripple effects, it is important not to burden their creation, lest one lose the many downstream benefits.

Generally speaking, reducing the overall burden of taxation will be a significant boon for creators, but it is important to note that a more favorable business environment across Africa will reap massive dividends for the world's most regulated region. In terms of cultural output, a glance at the top of the World Bank's Doing Business rankings identifies all the countries that are currently thought of as cultural hotspots. From vibrant London in the 18th century to Hollywood today, freer economies with credible, stable, and decentralized institutions foster creativity and tend to attract bright talent, regardless of its original location.[132] With brain drain a growing

concern, African economies would do much to stop their loss if talented local creators and innovators were able to secure more benefit from their work at home.

IV. Conclusion

Creative industries are an important opportunity for less-developed countries. In particular, those creative endeavors that tap into existing local talent and skills, such as music industries in Africa, represent low-hanging fruit. While thriving music industries clearly require considerable individual skill and training, they do not require extensive formal education or the development of sophisticated physical infrastructure.

Prioritizing creative clusters requires political will and a commitment to legal and institutional reform. What is required is the right legal environment—no simple task, but one that is more manageable when focused on empowering specific industries. While governments sometimes devote extra resources to enforcement in order to mollify trading partners, we advocate a greater, sustained effort to ensure that the copyright system is supportive of *local* creators.[133] In the case of African music industries, we believe that the relatively modest investment is warranted by the likely economic and social payoff.

Devoting specific resources to the creative industries can provide a strong foundation for spill-over growth in other sectors by providing quick growth and demonstrating that local industry can thrive in less-developed countries without government or donor subsidy. Creative industries are among the fastest-growing economic sectors in *both* rich and some poor countries.[134] Not only will the early growth of creative industries benefit creators and the local economy, it can provide a moral victory, showing that home-grown industries and home-grown culture can compete and thrive.[135] Such proof of the benefits of markets is essential.

However, creative industries are unlikely to prosper in the long run if the rest of the economy is not liberalized. Ultimately, the creative sector alone cannot drive the economy, nor can it flourish without financially healthy local customers and investors. Many of the policies we suggest here—e.g., greater enforcement of property rights, more effective courts, and regulatory and tax reform—would benefit the economy if applied generally across all sectors. And they should be.

Although a thriving popular music industry would help only a small proportion of Africans lift themselves from poverty, its success would be an important moral, economic, and cultural victory for African entrepreneurs. Moreover, the reforms that stimulate its success would also generate stable conditions for other entrepreneurial efforts based on local talents and tastes to follow. They would also hopefully lead to increased support and demand for the rule of law throughout society. That would truly benefit all.

Endnotes

1 For a history chronicling how Nashville became "Music City," *see* M. Kosser, "How Nashville Became Music City U.S.A.; 50 years of music row" (2006).

2 Other U.S. musical hotspots include Seattle, Austin, Memphis, and the state of Georgia. A recent study estimates that Nashville's local music industry contributes twice as much to Nashville's local economy as all those other places *combined* contribute to their own local economies. P. Raines & L. Brown, *The Economic Impact of the Music Industry in the Nashville-Davidson-Murfreesboro MSA* 4, Nashville Area Chamber of Commerce (2006), available at http://www.nashvillechamber.com/president/musicindustryimpactstudy.pdf. While New York and Los Angeles also have very large music industries, they do not define the local identity, as they span more genres and are less central to the local economy.

3 *Id.* at 4.

4 *Id.* at 6, 15-20.

5 *Id.*

6 *Id.* at 6. Nashville has a diversified economy, with the music industry at least the second-most important industry (depending on how one accounts for the impact of music-related tourism and other music-related economic activity), and healthcare, higher education, and the automotive industry also playing a role.

7 *See, e.g.,* W. Easterly, *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good* (2006) (lauding modest projects with measurable, accountable goals that involve locals).

8 H.A. Curtis, "The TVA and the Tennessee Valley—What of the Future?" *Land Economics*, 28(4), at 333 (1952).

9 *Id.*

10 L.M. Kyriakoudes, *The Social Origins Of The Urban South: Race, Gender, And Migration In Nashville And Middle Tennessee, 1890–1930*, at 39 (2003).

11 W.E. Cole, "The Impact of TVA Upon the Southeast," *Social Forces*, 28, at 435, 438 (1950).

12 D. Ekbladh, "'Mr. TVA': Grass-Roots Development, David Lilienthal, and the Rise and Fall of the Tennessee Valley Authority as a Symbol for U.S. Overseas Development, 1933–1973," *Diplomatic History*, 26(3), at 335-374 (2002) (quoting P. Hutchinson, "Revolution by Electricity", *Scribners Magazine*, Oct 1934, at193-200 (1934)).

13 Curtis, *supra* note 8, at 334.

14 U.S. Department of Health and Human Service, Center for Disease Control, *The History of Malaria, an Ancient Disease*, available at http://www.cdc.gov/malaria/history/index.htm (last accessed January 15, 2008).

15 Kyriakoudes, *supra* note 10, at 31, 33, 37.

16 "The hopes that Nashville would emerge as the South's leading iron or textile center were never realized. While the city did develop other industries based upon access to the raw materials and markets in its hinterland— lumber, wheat, and livestock— these resources did not develop the city into a great manufacturing center." *Id.* at 39.

17 Ekbladh, *supra* note 12, at 336. The TVA built a great deal of impressive infrastructure, improved economic conditions in the Tennessee Valley, and unambiguously succeeded in some of its ancillary programmers like eradicating malaria. Over the years, however, it has come to be criticized by some as an inefficient wealth transfer program that may have impeded the growth of the region. *See, e.g.,* W.U. Chandler, "The Myth of the TVA: Conservation and Development in the Tennessee Valley, 1933-80", *Journal of Policy Analysis and Management*, 4(1) (1984). Chandler contended that the TVA would have been more successful if it had focused on projects with more direct, immediate benefits to local people rather than massive projects.

18 K.B. Raitz et al., *Appalachia: A Regional Geography* 138-140 (1984).

19 Macon was a regular on the East Coast vaudeville circuit. *Uncle Dave Macon*, Country Music Hall of Fame and Museum, available at http://www.countrymusichalloffame.com/site/inductees.aspx?cid=142 (adapted from Kingsbury, P. (ed.), *The Encyclopedia of Country Music*, compiled by Country Music Hall of Fame and Museum, Oxford University Press).

20 *See* Peer Music's "Company History," available at http://www.peermusic.com/aboutus/companyhistory.cfm.

21 *Id.*

22 Ralph Peer, *Ralph Peer Remembers Jimmy Rodgers* (c.1953), available at http://www.peermusic.com/news/press_detail.cfm?announcement_id=273&back=press (last visited August 19, 2007) (the article was written by

Peer circa 1953, according the Bluegrass West site, http://www.bluegrasswest.com/ideas/jr-rpeer.htm).

23 J. Maeder, "Bristol Celebrates Birthplace of Country Music," *The Appalachian Voice* (2003), available at http://www.ncpress.net/cgi-local/search.cgi?gffd=774.

24 A number of country musicians did make the trip to New York, experiencing enough early success to indicate the existence of a market. For example, "Uncle" Dave Macon, a successful performer on the East Coast vaudeville circuit, had a number of hits recorded by New York labels in the mid to late '20s. *Uncle Dave Macon*, Country Music Hall of Fame and Museum, http://www.countrymusichalloffame.com/site/inductees.aspx?cid=142 (adapted from Kingsbury, P. (ed.), *The Encyclopedia of Country Music*, compiled by Country Music Hall of Fame and Museum, Oxford University Press). But the country music market fully materialized once the Bristol sessions brought performers together with recording technology, finally giving them a national audience.

25 *Id.*; C.K. Wolfe, "The Legend that Peer Built: Reappraising the Bristol Sessions," *in The Bristol Sessions: Writing About the Big Bang of Country Music*, 21 (T. Olson & C.K. Wolfe eds., 2005).

26 Wolfe, *supra* note 25, at 18.

27 *Id.* at 31.

28 Maeder, *supra* note 23.

29 Available at http://www.alamhof.org/rodgersj.htm

30 Peer, a consummate salesman, persuaded local newspapers to assist him with recruiting acts to record. As he later recalled: "I then appealed to the editor of a local newspaper, explaining to him the great advantages to the community of my enterprise. He thought that I had a good idea and ran a half column on his front page. This worked like dynamite...." Peer, *supra* note 22. Newspapers noted the relatively princely sums earned by other recording artists previously recorded by Victor: "[Ernest] Stoneman [received] $100,and each of his assistants $25 per day. Stoneman... received from the company $3,600 last year as his share of the proceeds on his records." Wolfe, *supra* note 25, at 25.

31 C. Wolfe,, *Liner Notes of The Bristol Sessions*, Country Music Foundation (1987), available at http://www.birthplaceofcountrymusic.org/node/32.

32 *Id.*

33 The 1928 Statistical Abstract of the United States 336, United States Government Printing Office. The .76 number is the general average hourly wage for all jobs in the bituminous coal industry.

34 *See id.* The numbers are based on the assumption that workers worked 48-hour work weeks, as mines operated on such a work week, although the Statistical Abstract notes that many miners worked fewer hours.

35 Nashville was home to the *Grand Ole Opry* Radio show and other important early country music institutions. Kyriakoudes, *supra* note 10, at 18. Country music talent came to Nashville to play the show, which drew record producers, which in turn drew more talent.

36 Acuff and Rose established the future powerhouse of music publishing Acuff-Rose in the 1940s, and the Bradleys founded the first recording studio on what was to become Nashville's famous "Music Row," where they recorded Patsy Cline and many other legends. *See* Kosser, *supra* note 1, at 10-14, 19-26.

37 F. Penna, M. Thorman & J. Finger, "The Africa Music Project," *in Poor People's Knowledge* (M. Thorman & J. Finger eds., 2004). We take inspiration from the Africa Music Project's enthusiasm for "African Nashvilles" and share its promoters' belief in grass-roots efforts. We depart from and elaborate on their analysis, however, in our views of the primary benefits from local music industries and the way to achieve them. While the Africa Music Project sought an export market via the Internet, we see great benefits from the establishment and support—through very specific and defined terms—of local and regional markets. We also map out specific legal and institutional reforms focused on the industry and private capacity building, which we believe should take precedence over particular marketing schemes.

38 *Office of the United States Trade Representative,* "The Benefits of Trade for Developing Countries," Trade Facts, June 2006. http://usmission.ch/Press2006/0701TradeFacts.pdf.

39 M.E. Porter, *The Competitive Advantage of Nations* (1998).

40 M.E. Porter, "Locations, Clusters, and Company Strategy," *in Oxford Handbook of Economic Geography*, 253 (G.L. Clark et al. eds., 2000).

41 M.E. Porter, "Clusters and the New Economics of Competition", *Harvard Business Review*, 76(6) (1998).

42 T. Cowen, *Creative Destruction: How Globalization Is Changing the World's Cultures* (2002).

43 T. Cowen, *Markets and Cultural Voices: Liberty vs. Power in the Lives of Mexican Amate Painters* (2005). The World Intellectual Property Organization has similarly recognized the possibilities of creative clusters as part of development efforts. *See, e.g.,* http://www.wipo.int/ip-development/en/creative_industry/creative_clusters.html.

44 As Keith Maskus has observed, most economies in poor countries are unlikely to reap the rewards of capital- and knowledge-intensive industries immediately. In order for countries to benefit from these types of industries, they need to enhance capacity in three areas to benefit from IPRs: first, they need "strong levels of educational attainment and sizeable endowments of human capital;" second, local firms must invest in research and development; and third, "financial markets [must be] capable of managing the significant risks involved in technology development." K.E. Maskus, "Intellectual Property Rights and Economic Development," *Case Western Journal of International Law*, 32, at 471,497 (2000).

45 *See* Raines & Brown, *supra* note 2.

46 Peer, *supra* note 22.

47 Copyright per se did not play as direct of a role in Victor's financing of the Bristol Sessions, as sound recordings were not then protected by U.S. copyright law, but rather by a patchwork of state laws. Still, Victor's rights in these recordings were effectively protected by law and the difficulty of copying at the time. In 1972, sound recordings were brought under the U.S. Copyright Act in reaction to the growing problem of piracy.

48 Maeder, *supra* note 23.

49 Georgia Music Hall of Fame, "Ralph Peer Biography", available at http://www.georgiamusicstore.com/artist/P219557/.

50 For example, Jimmie Rodgers' first big hit, *Blue Yodel (T for Texas),* was one of six songs that Rodgers recorded during his first two recording sessions with Peer. C. Comber & M. Paris, "Jimmie Rodgers," *in Stars of Country Music*, 127-128 (1975). Of the six, only *Blue Yodel* was completely original. *Id.*

51 Kyriakoudes, *supra* note 10, at 8. The Southern traditionalist intellectual Donald Davidson bemoaned the abandonment of folk music in favor of "*new* songs—quasi folk songs or frankly popular ditties—which will bring [the musician] royalties from phonograph records and music sheets." D. Davidson, "Current Attitudes Toward Folklore" *in Still Rebels, Still Yankees and Other Essays*, 134 (1957).

52 Editors of *Travel and Leisure Magazine*, "100 Greatest Trips", American Express Publishing Corporation (2007).

53 *Id.*

54 We are grateful to Julian Morris for his insights into African music. *See also, e.g.,* T. Cowen, *Creative Destruction* (2004).

55 *See* S.A. Floyd, Jr., *The Power Of Black Music: Interpreting Its History From Africa To The United States* (1995); P.K. Maultsby, "Africanisms in African American Music," *in Africanisms In American Culture*, 326 (J.E. Holloway ed., 2nd ed. 2005).

56 UNCTAD, "Background Paper for Secretary-General's High-Level Panel on the Creative Economy and Industries for Development", TD(XII)/BP/4 (2008), available at http://www.unctad.org/en/docs/tdxiibpd4_en.pdf.

57 GhanaMusic.com, "BUSAC Support For Ghana's Music Industry", June 23, 2007, http://www.ghanamusic.com/modules/news/article.php?storyid=1754 (quoting Professor John Collins, a Lecturer at School of Performing Arts of University of Ghana).

58 Ellyn Hament, Zambia*: A Melting Pot of Musical Styles*, Exploratorium http://www.exploratorium.edu/eclipse/zambia/music.html (2001).

59 Eyre Banning, *Interview with Chisha Folotiya*, Afropop, http://www.

afropop.org/multi/interview/ID/102/Chisha+Folotiya-2006 (2006).

60 *Id.*

61 *Id.*

62 Easterly, *supra* note 7, at 4; *see also* Laurie Garrett, *The Challenge of Global Health*, 86 Foreign Affairs (Jan/Feb 2007). Garrett chronicles the sadly disappointing results of the extraordinary rise in public and private donations directed to public health spending in the developing world in recent years. The fundamental problems of lack of trained health professionals, shortcomings in sanitation and infrastructure, and corruption have frustrated much of this generosity.

63 UNCTAD press release, "Creative Industries Emerge As Key Driver of Economic Growth With Trade Nearly Doubling in Decade", January 14, 2008, available at http://www.unctad.org/Templates/webflyer.asp?docid=9467&intItemID=1528.

64 *Id.*

65 M. Tchebwa, *African Music: New Challenges, New Vocations*, UNESCO (2005).

66 S. Siwek, "Copyright Industries in the US Economy; The 2006 Report," Economists Incorporated, prepared for the International Intellectual Property Alliance, available at http://www.iipa.com/pdf/2006_siwek_full.pdf.

67 UNCTAD press release, "Creative Industries Emerge As Key Driver of Economic Growth With Trade Nearly Doubling in Decade," January 14, 2008, available at http://www.unctad.org/Templates/webflyer.asp?docid=9467&intItemID=1528.

68 The popularity of Youssou N'Dour is undisputable, but it should be known that what many consider to be his breakthrough album, The Lion, released in 1989, featured a popular song he recorded with already hugely popular Englishman Peter Gabriel.

69 Penna, Thorman & Finger, *supra* note 37.

70 R. Letts, *The Promotion of Musical Diversity*, UNESCO (2006), available at http://www.unesco.org/imc/programmes/imc_diversity_report.pdf.

71 *Id.*

72 Creative industries in Paris tend to support Francophone artists, whereas London is more typically suited to Anglophone musicians and filmmakers.

73 *The Economist*, "Bleak Publishing Houses", November 22nd 2007.

74 World Development Indicators, World Bank.

75 Penna, Thorman & Finger, *supra* note 37, at 98.

76 Penna, Thorman & Finger, *supra* note 37.

77 *Id.*

78 E.S. Nwauche, "Intellectual Property Rights, Copyright and Development Policy in a Developing Country: Options For Sub Saharan African Countries," Copyright Workshop at Zimbabwe International Book Fair, presented June 30th, 2003.

79 *Id.*

80 Penna, Thorman & Finger, *supra* note 37.

81 Ghana Billboard, Interview with Alhaji Sidiku Buari, Ghanamusic.com, http://ghanamusic.com/artman/publish/article_1309.shtml (December 4, 2004).

82 Government Urged to Enact Tough Laws to Address Piracy in Music Industry, Ghanamusic.com, http://ghanamusic.com/artman/publish/article_1809.shtml (April 28, 2005).

83 *Id.*

84 Hernando De Soto, *The Mystery of Capital* (2000).

85 N. Norley, "Musicians Bemoan Fall of Industry," *Ghanaian Chronicle* Feb 7, 2007, available at http://allafrica.com/stories/printable/200702070880.html.

86 *Interview with Chisha Folitya,* Afropop, http://www.afropop.org/multi/interview/ID/102/Chisha+Folotiya-2006 (2006). As discussed earlier, the Zambian industry revived at the turn of the century after the efforts of pioneering entrepreneurs and passage of a copyright law.

87 It should also be noted, however, that large expatriate communities, originating from the same countries from which many African creators come, now reside in the UK and France. These communities constitute a growing market demanding "home-grown" music.

88 The Africa Music Project notes that non-African copyright owners are still largely unwilling to license their property to the African marketplace. Penna, Thorman & Finger, *supra* note 37.

89 D. Wabala, "Kenya: Illegal Trade in Pirated Music Gets Bloody As Dealers Scramble for Control," The Nation, March 25, 2007, http://allafrica.com/stories/200703260049.html, accessed March 28, 2007.

90 Penna, Thorman & Finger, *supra* note 37.

91 M. Tchebwa, *African Music: New Challenges, New Vocations*, UNESCO, at 52-53 (2005).

92 "Carlos Sakyi Slams Ghana's Music Distribution System," *The Ghanaian Chronicle,* July 6, 2007, available at http://news.thinkghana.com/entertainment/200707/6540.asp.

93 Penna, Thorman & Finger, *supra* note 37.

94 "Ghanaian musicians are worried!" http://www.myjoyonline.com/tools/print/printnews.asp?contentid=1407, accessed Mar 20, 2007.

95 *Id.*

96 Afrik.com, "Les radios sénégalaises ne paient pas leurs artistes," July 10, 2001, available at http://www.afrik.com/article3021.html.

97 *Id.*

98 *See, e.g.,* GhanaMusic.com, "Carlos Sakyi returns Sidiku's fire," Jan 12, 2008, available at http://www.ghanamusic.com/2008/01/12/carlos-sakyi-returns-sidikus-fire/.

99 The charges are hotly disputed, as Alhaji Buari, Acting Board Chairman of Copyright Society of Ghana, has filed a libel suit against a newspaper that repeated the charges. Ghanamusic.com, "Alhaji Sidiku Buari Sues Weekly Fylla," April 4, 2007, available at http://www.ghanamusic.com/modules/news/article.php?storyid=1368.

100 http://www.mcsk.or.ke/.

101 For a complete description of the Nigerian Copyright Commission's responsibilities, *see* http://www.nigeriafirst.org/cgi-bin/artman/exec/view.cgi?archive=1&num=6616&printer=1.

102 Penna, Thorman & Finger, *supra* note 37.

103 The Chairman of the Copyright Office of Ghana, Alhaji Sidiku Buari, is allegedly withholding more than $6 million that is owed to musicians. He is accused not paying artists and attempting to exert influence over their content. Personal communication on March 3, 2008 with Franklin Cudjoe, Executive Director, IMANI, Accra, Ghana.

104 Penna, Thorman & Finger, *supra* note 37.

105 Freedom House Map of Freedom in the World, available at http://www.freedomhouse.org/template.cfm?page=363&year=2007.

106 Penna, Thorman & Finger, *supra* note 37.

107 *Id.*

108 Personal correspondence with Michael Andrews, Managing Director, Kenya Association of Producers & Videograms.

109 Penna, Thorman & Finger, *supra* note 37.

110 Doing Business Database, World Bank, available at http://www.doingbusiness.org/ExploreEconomies/?economyid=164.

111 De Soto, *supra* note 84.

112 Thus the discussion in the United States about the "orphan works" problem, where users and potential licensees find it impossible to track down owners of many older works.

113 Unfortunately, the current international copyright system created by the Berne Convention is not fully amenable to such formal titling requirements. The U.S. had to give up its registration and notice requirements to join Berne in 1989. Nevertheless, there are ways to encourage them without requiring them.

114 Article 45 of TRIPS requires only compensatory damages and the availability of expenses. It also permits, but does not require, attorney fees and payment of pre-established damages. Copyright laws in several African countries give courts discretion to award "additional damages... as the court may consider appropriate in the circumstances" where the infringement is flagrant. While this approach could be helpful, its utility is limited because it is confined to special cases at the court's discretion. Kenya, South Africa, Zimbabwe, and Nigeria all employ this language in their copyright acts in nearly identical provisions.

115 In many wealthy countries, business rivals pool their efforts to combat the collective problem of piracy through a number of trade associations such as the International Federation of Phonographic Industries (IFPI), International Publishers Association, Recording Industry Association of America (RIAA), Motion Picture Association, Business Software Alliance, and the Industry Trust for IP Awareness. Pirates typically do not specialize in the works of any one business, so it is more effective to pool private resources in order to combat them collectively. Moreover, these organizations develop specialized knowledge and skills in combating piracy that is best housed and shared in one place, rather than scattered among many players.

116 *See, e.g.,* Copyright Law of Nigeria, Sections 15, 15A, available at http://www.wipo.int/clea/docs_new/en/ng/ng001en.html#P203_18719.

117 *See, e.g.,* Copyright Law of Mozambique, Section 59, available at http://www.wipo.int/clea/docs_new/en/mz/mz002en.html#P386_36415.

118 *Associated Press*, "China executes former FDA chief amid product safety crisis," July 10, 2007, available at http://www.cbc.ca/world/story/2007/07/10/china-tainted-products.html.

119 *See* Mark F. Schultz, "Copynorms: Copyright Law and Social Norms," *in Intellectual Property and Information Wealth vol. 1* (P.K. Yu ed., 2007).

120 R.M. Sherwood, "Some Things Cannot Be Legislated," *Cardozo Journal of International and Comparative Law*, 10, at 42 (2002).

121 *See* "An Act to establish a pilot programmed in certain United States district courts to encourage enhancement of expertise in patent cases among district judges," H.R. 34, 110th Congress, available at http://thomas.loc.gov/cgi-bin/bdquery/z?d110:h34.

122 Sherwood, *supra* note 120, at 42.

123 Schultz, *supra* note 119, at 201, 217.

124 We also think that private, competitive CROs would work best in rich countries as well, but that is a discussion for another day.

125 For example, in the United States, composers' rights are administered by three competing, private organizations: The American Society of Composers, Authors and Publishers (ASCAP), Broadcast Music, Inc. (BMI), and SESAC, Inc.

126 *See* Section III(C)(iii).

127 *Id.*

128 Article 67 states:

> In order to facilitate the implementation of this Agreement, developed country Members shall provide, on request and on mutually agreed terms and conditions, technical and financial cooperation in favor of developing and least-developed country Members. Such cooperation shall include assistance in the preparation of laws and regulations on the protection and enforcement of intellectual property rights as well as on the prevention of their abuse, and shall include support regarding the establishment or reinforcement of domestic offices and agencies relevant to these matters, including the training of personnel.

TRIPS, Article 67, available at http://www.wto.org/english/docs_e/legal_e/27-trips_08_e.htm#art67.

129 http://portal.unesco.org/culture/en/ev.php-URL_ID=28581&URL_DO=DO_TOPIC&URL_SECTION=201.html.

130 *Id.*

131 *See* Section III(C)(iv).

132 *See* T. Cowen, *In Praise of Commercial Culture* (1998).

133 TRIPS compliance merely sets minimum standards, largely focused on treatment of foreign IP owners. It does not require a special effort with respect to IPRs—as Article 41(5) of TRIPS makes clear, countries are not required to prioritize IPR enforcement when allocating judicial and enforcement resources. Article 41(5) states:

> It is understood that this Part does not create any obligation to put in place a judicial system for the enforcement of intellectual property rights distinct from that for the enforcement of law in general, nor does it affect the capacity of Members to enforce their law in general. Nothing in this Part creates any obligation with respect to the distribution of resources as between enforcement of intellectual property rights and the enforcement of law in general.

TRIPS, Article 41, available at http://www.wto.org/english/tratop_e/trips_e/t_agm4_e.htm.

134 *Id.*

135 Penna, Thorman & Finger, *supra* note 37, at 97.

International & National Security Law

The Fourth Amendment Goes to War

*By Robert J. Delahunty**

Early one morning last March, I received a phone call from David Barron, who had recently begun working for the Obama Administration as Deputy Assistant Attorney General in the Justice Department's Office of Legal Counsel (OLC). David and I had been colleagues at OLC during the Clinton Administration. I stayed on the OLC staff to work for President Bush, and David moved on to become a member of the Harvard Law School faculty. We remained on friendly terms. David was calling to give me a head's up that later in the day the Justice Department would be releasing ten OLC memoranda from the Bush Administration, two of which I had co-authored.[1] (I was, and am, very grateful to David for his thoughtfulness.)

The media storm that broke out upon the memos' release was brief but (predictably) intense, and the focus was on one of the opinions that I had co-authored. "Quite astounding," opined the head of the ACLU's national security project.[2] "[A] theory of presidential power amounting to virtual dictatorship," wrote Rosa Brooks.[3] The opinion permitted the military to "search anyone's home, wiretap anyone's phone, or arrest and hold any citizen, all without a warrant, and based on the flimsiest suspicion," said Lawrence Rosenthal.[4] Newsweek's Michael Isikoff—the author of the discredited 2005 report that U.S. interrogators at Guantanamo had flushed a Koran down a toilet to offend detainees[5]—characterized the opinion as an effort to "potentially suspend First Amendment freedom-of-the-press rights in order to combat the terror threat."[6] Ironically, only a few days after the opinions were released, the media reported that President Obama was considering the deployment of troops along the Mexican border to control escalating violence there—a report that triggered no angry outbursts from the "civil libertarians" who shortly before had been denouncing an eight-year-old legal opinion from the Bush Administration because it envisaged the domestic use of the military.[7]

The opinion that provoked this outcry was prepared by John Yoo, then a Deputy Assistant Attorney General at OLC, and me. At the time, I was a senior civil servant at OLC. (Previously, I had worked in the Justice Department for fifteen years under four Presidents.) The opinion was written for Alberto Gonzales, then the White House Counsel, and William J. Haynes II, then the General Counsel of the Defense Department.[8] The opinion was dated October 23, 2001—not yet six weeks after the September 11, 2001, attacks on the World Trade Center and the Pentagon and the planned but abortive attack on the Capitol, of September 11, 2001. Stated in full, the propositions in the opinion that attracted the critics' ire were these:

> Fourth, we turn to the question whether the Fourth Amendment would apply to the use of the military domestically against foreign terrorists. Although the situation is novel (at least in the nation's recent experience), we think that the better view is that the Fourth Amendment would not apply in these circumstances. Thus, for example, we do not think that a military commander carrying out a raid on a terrorist cell would be required to demonstrate probable cause or to obtain a warrant.
>
> Fifth, we examine the consequences of assuming that the Fourth Amendment applies to domestic military operations against terrorists. Even if such were the case, we believe that the courts would not generally require a warrant, at least when the action was authorized by the President or other high executive branch official. The Government's compelling interest in protecting the nation from attack and in prosecuting the war effort would outweigh the relevant privacy interests, making the search or seizure reasonable.[9]

Those propositions seemed to me entirely defensible in the circumstances of October 2001, and while we may perceive the risks of terrorism differently eight years later, I believe that those propositions would still hold true if comparable terrorist threats were to emerge again[10]—as they did in the November 26-29, 2008 attacks attributed to the terrorist group Lashkar-e-Taiba on the city of Mumbai in India. The Mumbai attacks killed 172 people, overwhelmed local police contingents (including the Anti-Terrorism Squad), and required the Indian government to deploy army and marine units—including the elite National Security Guard or "Black Cat Commandos"—into the city to conduct search-and-rescue operations and to engage the terrorists in combat.[11]

In what follows, I shall first examine the factual and legal circumstances in which the 10/23/01 OLC Opinion was prepared. In that section, I will explain why the "war paradigm"—rather than the previously-prevailing "law enforcement paradigm"—was the underlying presupposition of the opinion. Second, I will briefly outline a general understanding of how the Bill of Rights has been interpreted and applied in wartime or other similar crisis situations. Third, in light of that background, I will analyze and defend the opinion's conclusions that the Fourth Amendment would not apply to domestic military operations against al Qaeda or other, similar terrorist groups, or that, in the alternative, the courts would apply only a relaxed "reasonableness" test to such operations.

*Associate Professor, University of St. Thomas School of Law, Minneapolis, MN. *I would like to thank Thomas J. Alford, Teresa S. Collett, Michael S. Paulsen, Antonio F. Perez, John Radsan, Jeffrey K. Shapiro, Gregory Sisk, and John C. Yoo for their comments, and Joel Whitlock, Class of 2011, for his help as my Research Assistant in preparing this paper. I alone am responsible for the views expressed.*

To be perfectly clear: the position being defended here is *not* that the Warrant Clause of the Fourth Amendment is somehow "suspended" during wartime, owing perhaps to an "emergency" situation. Rather, the central claim being defended is that the Warrant Clause does not *reach* the relevant military actions *in the first place*.

I.

When the 10/23/01 OLC Opinion was being prepared, the United States seemed plainly to be—in both the material and the legal sense—a nation at war. Writing in 2002, a criminal justice scholar and a retired Lieutenant Colonel succinctly described the situation as it stood in the weeks just after the 9/11 attacks:

> When members of the Al Qaeda terrorist organization attacked the World Trade Center and the Pentagon with hijacked commercial aircraft on September 11, it was the first time since a young America fought pitched battles with British troops during the War of 1812 that aggressors from abroad had engaged targets on contiguous American soil. In short order, the coordinated attack by terrorists became a watershed event in U.S. history, as it led to substantial changes in the fabric of our nation's life. Since September 11, America has been on a war footing, with armed soldiers standing guard at our nation's airports, enhanced security at nuclear power plants and other vulnerable locations, and military jets flying combat air patrols in order to intercept and shoot down hijacked commercial aircraft. The legal climate has also been affected by the events of September 11. Congress has passed, and the President has signed, anti-terrorism legislation that expands police surveillance powers. Additionally, the President has announced that suspected terrorists who are not U.S. citizens may be tried in special military tribunals lacking many of the due process standards of American criminal courts.[12]

As the 10/23/01 OLC Opinion was being written, Air Force fighter jets were patrolling the skies above Washington, D.C. to protect the nation's capital from another attack by hijacked, weaponized civilian aircraft. There had been few objections to the announcement that fighter jets had been dispatched to intercept and shoot down the aircraft hijacked on September 11,[13] and few objections to the subsequent announcement that procedures were being established for the military to shoot down hijacked aircraft in the future. The atmosphere throughout the country was one of apprehension and anxiety. Further attacks were expected, although there was no certainty as to when, where, or how they would occur. Against that backdrop, health officials in Florida announced on October 4, 2001 the first case of pulmonary anthrax in the United States in almost 25 years.[14] Later cases of anthrax exposure soon began to be reported, including several at major media outlets in New York City. In mid-October, an anthrax-laden letter was opened in the Washington, D.C. office of Senator Tom Daschle (D-SD). Congressional buildings were evacuated and federal government mail delivery in Washington, D.C. was virtually halted. An additional anthrax-laden letter addressed to Senator Patrick Leahy (D-VT) was also discovered. Leading political figures and many members of the general public linked the anthrax episodes to the 9/11 attacks, and feared that a bioterrorist war against the United States had begun.[15]

Legally also, the condition of the nation was one of war. On September 18, Congress had enacted Pub. L. No. 107-40, 115 Stat. 224 (2001), entitled "Joint Resolution to authorize the use of United States Armed Forces against those responsible for the recent attacks launched against the United States," which found that the attacks of September 11 had rendered it "both necessary and appropriate that the United States exercise its rights to self-defense and to protect United States citizens both *at home* and abroad" (emphasis added). The operative part of the statute authorized the President, without geographical limitation,

> to 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, in order to prevent any future acts of international terrorism against the United States by such nations, organizations or persons.[16]

Likewise, both the United Nations Security Council[17] and the North Atlantic Treaty Organization[18] had recognized that the United States had a right of self-defense under the international law of armed conflict.

In the eight years since the 10/23/01 OLC Opinion was issued, scholars have debated whether the legal framework created by these governmental and inter-governmental decisions should continue to be applied to the "war on terror."[19] Some scholars have argued that *neither* the war paradigm *nor* the law enforcement paradigm should provide the legal framework for domestic counter-terrorism, even when it requires the use of military force.[20] In this view, "terrorism" may present a hybrid or *sui generis* paradigm, in which the laws of war may partly apply and the rules of the criminal justice system may also partly apply. If such a paradigm could be fully articulated, it might well show that the Fourth Amendment *did* constrain the use of military force in domestic counter-terrorism operations, at least in some circumstances or to some degree.

Such a novel paradigm was not conceptually available in October 2001, however, and may not be so even eight years later. When the 10/23/01 OLC Opinion was being written, the *only* recognized alternative to the war paradigm in assessing the legality of counter-measures to al Qaeda's attacks was the more traditional law enforcement paradigm. Although that approach had generally dominated the United States' responses to al Qaeda before 9/11[21] and still had some supporters after it, it had plainly failed to prevent the attacks and seemed entirely inadequate legally and practically as a response to them.[22] The scale, intensity, lethality, and purposes of the 9/11 attacks and the nature of the responses they required precluded viewing them through the prism of the criminal justice system.

Second and no less important, given the decisions made by Congress, the President, the Security Council, and NATO that the United States was in a state of war and could

lawfully use its armed forces both for prosecuting that war and for defending itself, OLC was unquestionably bound to rely on the war paradigm in providing advice to the Defense Department on domestic deployments of the Armed Forces for military purposes.

How, then, would the Fourth Amendment apply to such deployments?

II.

The Constitution was consciously designed to enable the United States to wage war effectively. The preamble of the Constitution expressly sets forth, among the basic purposes of the Founding, that of "provid[ing] for the common defence." Over the nation's existence, the nation's courts and Presidents have read the Constitution with that purpose in mind. Typical of many Supreme Court pronouncements on the Constitution in wartime is its 1948 decision, *Lichter v. United States*.[23] There the Court said:

> [I]t is of the highest importance that the fundamental purposes of the Constitution be kept in mind and given effect in order that, through the Constitution, the people of the United States may in time of war as in peace bring to the support of those purposes the full force of their united action. In time of crisis nothing could be more tragic or less expressive of the intent of the people than so to construe their Constitution that by its own terms it would substantially hinder rather than help them in defending their national safety.[24]

The Court went on to quote extensively from a celebrated address given during the First World War by Charles Evans Hughes, entitled "War Powers on the Constitution."[25] As the Court noted, Hughes had said:

> [T]he power has been expressly given to Congress to prosecute war, and to pass all laws which shall be necessary and proper for carrying that power into execution. That power explicitly conferred and absolutely essential to the safety of the Nation is not destroyed or impaired by any later provision of the constitution or by any one of the amendments. These may all be construed so as to avoid making the constitution self-destructive, so as to preserve the rights of the citizen from unwarrantable attack, while assuring beyond all hazard the common defence and the perpetuity of our liberties.[26]

Summing up, the *Lichter* court said:

> [T]he primary implication of a war power is that it shall be an effective power to wage war successfully. Thus, while the constitutional structure and controls of our Government are our guides equally in war and in peace, they must be read with the realistic purposes of the entire instrument fully in mind.[27]

These views have been expressed, not only by the Supreme Court, but by wartime leaders such as President Abraham Lincoln. In a public letter of June 12, 1863, Lincoln wrote: "[T]he Constitution is not, in its application, in all respects the same, in cases of rebellion or invasion involving the public safety, as it is in times of profound peace and public security. The Constitution itself makes the distinction...."[28]

These general rules of interpretation apply to the Bill of Rights no less than to other constitutional provisions. Indeed, the text of the Bill of Rights makes plain on its face that its provisions may apply differently in wartime and in peace. Thus, the Third Amendment states (emphasis added): "No soldier shall, *in time of peace* be quartered in any house, without the consent of the Owner, *nor in time of war*, *but in a manner to be prescribed by law*."

Likewise, the Fifth Amendment states (emphasis added): "No person shall be held to answer for a capital, or otherwise infamous crime, unless on a presentment or indictment of a Grand Jury, except in cases arising in the land or naval forces, or in the Militia, *when in actual service in time of War or public danger*..."

In a similar spirit, the "mini-Bill of Rights" in Article I, § 9, cl. 2 provides for the possibility of the suspension of the writ of habeas corpus "when in Cases of Rebellion or Invasion the public Safety may require it."

Although the Third Amendment is something of an orphan in the case law,[29] its language is instructive here. Under standard Fourth Amendment jurisprudence, no private space is more protected than the home.[30] Yet under the Third Amendment, the military may occupy private homes and quarter troops in them during wartime, provided that a "law" so permits.[31] But if the government has the power to *quarter troops* in homes in wartime—a truly massive encroachment on individual privacy—its power to *search* homes must surely also be greatly expanded during war.

Furthermore, it is not necessary for a provision of the Bill of Rights to include a specific wartime exception for it to apply differently (or not at all) during war. Consider the Takings Clause of the Fifth Amendment.[32] A long line of cases going back at least to the Supreme Court's 1887 decision in *U.S. v. Pacific R.R.*[33] establishes that the government is not liable under the Takings Clause for private property that it destroys during military operations in wartime under the compulsion of military necessity. Moreover, this is true whether the destruction occurs on U.S. soil or overseas, and whether the property is owned by U.S. citizens or not. In *Pacific R.R.*, the plaintiff sought compensation for the Union Army's destruction of several of its bridges during operations intended to repulse a Confederate invasion of Missouri. The Court denied the claim, saying:

> The destruction or injury of private property in battle, or in the bombardment of cities and towns, ... had to be borne by the sufferers alone.... Whatever would embarrass or impede the advance of the enemy, as the breaking up of roads, or the burning of bridges, or would cripple and defeat him, as destroying his means of subsistence, were lawfully ordered by the commanding general. Indeed, it was his imperative duty to direct their destruction. The necessities of the war called for and justified this. The safety of the state in such cases overrides all considerations of private loss.[34]

The Supreme Court reaffirmed the doctrine of *Pacific R.R.* in a 1952 case, *United States v. Caltex*,[35] in which a group

of oil companies sought compensation for the U.S. Army's demolition of their facilities and products in Manila (which was then U.S. soil) in order to prevent them from falling into the hands of the advancing Japanese Army. In denying recovery, the *Caltex* Court said that

> [t]he terse language of the Fifth Amendment is no comprehensive promise that the United States will make whole all who suffer from every ravage and burden of war. This Court has long recognized that in wartime many losses must be attributed solely to the fortunes of war, and not to the sovereign.[36]

These and similar cases do not mean, as one court has rightly noted, that the Takings Clause is *suspended* during wartime.[37] But they *do* mean that the Takings Clause is simply *inapplicable* to certain losses that occur as a necessary incident of military operations.

Further, the Supreme Court's case law suggests that the Due Process Clause of the Fourteenth Amendment (and therefore, it would seem, of the Fifth Amendment) would also apply differently in time of war or other grave emergencies (such as insurrection). A case in point is Justice Holmes' decision in *Moyer v. Peabody*,[38] upholding against a due process challenge the authority of a state Governor, in suppressing an insurrection by military means, to detain the leader of a miners' union for two and a half months without criminal charges, until the Governor thought that it was safe to release him. In the course of his opinion, Holmes wrote broadly that "[w]hen it comes to a decision by the head of the state upon a matter involving its life, the ordinary rights of individuals must yield to what he deems the necessities of the moment."[39] While it may be unlikely that *Moyer* would be decided on the same basis today,[40] even the much later case of *Duncan v. Kahanamoku* appeared to acknowledge "the power of the military simply to arrest and detain civilians interfering with a necessary military function at a time of turbulence and danger from insurrection or war."[41]

In sum, then, wartime conditions may profoundly affect the interpretation and application of the provisions of the Bill of Rights. In particular, some of those provisions may apply very limitedly, or even not at all, to military operations during wartime. The 10/23/01 OLC Opinion read the Fourth Amendment in that light.

III.

The Fourth Amendment requires that police searches and seizures not be "unreasonable," and the Supreme Court has held that if the intrusion counts as a full "search" or "seizure"—as distinct, say, from a brief, investigative police "stop"—it must be based upon "probable cause." Further, the Court has also held that, in addition to requiring probable cause, the reasonableness of a "search" normally depends on the government's having obtained a warrant beforehand. The Court has, however, carved out an exception from these (largely, judge-made) requirements in cases of "special needs, beyond the normal need for law enforcement."[42] In a case decided not long before the 10/23/01 OLC Opinion, the Court stated that "the Fourth Amendment would almost certainly permit an appropriately tailored roadblock [at a law enforcement checkpoint] set up to thwart an imminent terrorist attack."[43] Subsequently to the OLC Opinion, courts of appeals, in reliance on the "special needs" doctrine, have upheld governmental searches intended to prevent terrorist attacks.[44]

The requirements of reasonableness, probable cause, and a warrant apply paradigmatically to law enforcement operations. Thus, the Court has said that "[t]he Fourth Amendment was tailored explicitly for the criminal justice system,"[45] and that "[t]he standard of probable cause is peculiarly related to criminal investigations."[46] This is not to say that the Fourth Amendment applies only to police action; it has a wider scope.[47] But law enforcement activity remains the core subject matter to which the Fourth Amendment is addressed. The "driving force" behind the Fourth Amendment "was widespread hostility among the former colonists to the issuance of writs of assistance empowering revenue officers to search suspected places for smuggled goods, and general search warrants permitting the search of private houses, often to uncover papers that might be used to convict persons of libel."[48] There is nothing to suggest that the Fourth Amendment was originally intended to apply to wartime operations carried out for military rather than law enforcement purposes by the armed forces. Application of law enforcement-based tests of reasonableness—let alone the warrant requirement—to military commanders conducting wartime operations on U.S. soil could pose a significant risk of damaging the military's ability to function in combat-type situations and to carry out its missions successfully.

While there appears to be no Supreme Court precedent directly on point, the Court's 1990 decision in *United States v. Verdugo-Urquidez* is highly illuminating. There the Court held that the Fourth Amendment did not apply to the search by Drug Enforcement Agency personnel of the Mexican residence of a Mexican citizen who had no voluntary association with the United States. The plurality found that there was "no indication that the Fourth Amendment was understood by contemporaries of the Framers to apply to activities of the United States directed against aliens in foreign territory or in international waters."[49] Further, acceptance of the claim that the Fourth Amendment had such extraterritorial reach "would have significant and deleterious consequences for the United States in conducting activities beyond its boundaries."[50] Thus, the opinion reasoned, if the Fourth Amendment were held to apply to the extraterritorial activities of the military, it "could significantly disrupt the ability of the political branches to respond to foreign situations involving our national interest.... [and create] a sea of uncertainty as to what might be reasonable in the way of searches and seizures conducted abroad."[51]

If that reasoning holds for military operations overseas, it seems to be no less true of military operations inside the United States. Suppose that the U.S. Coast Guard detects in international waters what it believes may be a vessel heading towards the United States to launch a terrorist attack on American soil—in other words, a situation like the seaborne attack on Mumbai. Assuming that the Coast Guard may intercept the vessel without having to satisfy the requirements

of "reasonableness" and "probable cause," should the analysis be different the moment that the suspect vessel enters U.S. territorial waters? What if the vessel nears the coast and lets off its crew and passengers? What if the U.S. Air Force identifies a civilian aircraft nearing U.S. air space that refuses to respond to signaling or to disclose its flight pattern and is otherwise behaving suspiciously? Is the Air Force *more* restrained in its ability to intercept the aircraft once it enters U.S. airspace than it was before? The reasons that *Verdugo-Urquidez* outlined for thinking that the Fourth Amendment does not apply to *extraterritorial* military operations seem to hold good for military operations such as these *inside* the United States.

Or consider the kind of "urban warfare" scenario that seemed only too plausible in the immediate aftermath of the 9/11 attacks. I cannot improve here on what the 10/23/01 OLC Opinion itself said:

> Consider, for example, a case in which a military commander, authorized to use force domestically, received information that, although credible, did not amount to probable cause, that a terrorist group had concealed a weapon of mass destruction in an apartment building. In order to prevent a disaster in which hundreds or thousands of lives would be lost, the commander should be able to immediately seize and secure the entire building, evacuate and search the premises, and detain, search, and interrogate everyone found inside. If done by the police for ordinary law enforcement purposes, such actions most likely would be held to violate the Fourth Amendment. See *Ybarra v. Illinois*, 44 U.S. 85 (1979).... To subject the military to the warrant and probable cause requirement that the courts impose on the police would make essential military operations such as this utterly impossible. If the military are to protect public interests of the highest order, the officer on the scene must be able to "exercise unquestioned command of the situation." *Michigan v. Summers*, 452 U.S. 692, 703 (1981).[52]

To be sure, the 10/23/01 OLC Opinion's argument relied critically on the distinction between law enforcement and domestic military operations. And it may on occasion be difficult to know whether to characterize a particular use of the military as one or the other. As the 10/23/01 OLC Opinion said, "[i]f the President were to deploy the Armed Forces within the United States in order to engage in counter-terrorism operations, their actions could resemble, overlap with, and assist ordinary law enforcement activity."[53] Nonetheless, the distinction is clearly implicit in the Posse Comitatus Act,[54] which is usually regarded as a bulwark of civil liberties; and courts interpreting the Posse Comitatus Act have been able to understand and apply it without undue difficulty. Subject to certain exceptions, the Posse Comitatus Act prohibits the "willful[] use[]" of "any part of the Army or the Air Force as a posse comitatus or otherwise *to execute the laws*" (emphasis added). But just as the statute generally *forbids* the use of the Armed Forces for domestic law enforcement, so it impliedly *permits* their use for other purposes: it is a sword that cuts both ways. Under the longstanding views of both the Justice Department and the Defense Department, therefore, the Posse Comitatus Act was construed *not* to prohibit the use of the Armed Forces in actions undertaken primarily for a military purpose.[55] The traditional statutory distinction between law enforcement and military operations can readily be understood to have constitutional dimensions.[56] It can be understood to delineate both the reach of the Fourth Amendment and its limits.

The discussion to this point has, no doubt, worn a somewhat abstract air. That is because it has so far failed to mention what would surely emerge as a crucial element in any Fourth Amendment review of a domestic military counter-terrorist operation: the existence of appropriate Rules of Engagement (ROE) governing that operation. Engagement with hostile forces is rarely, if ever, left to the unstructured discretion of military commanders in the field. Rather, it is subject to careful operational controls embedded in ROE. Such ROE (which are typically classified) exist for military operations overseas, and it is overwhelmingly likely that they have been, or would be, issued if counterterrorist operations by the military were expected within the United States.[57] Such ROE could enter into the Fourth Amendment analysis in either of two ways. First, if carefully crafted and conscientiously followed, they might be used to persuade reviewing courts that "reasonableness" review under the Fourth Amendment was unnecessary (since effective safeguards to govern the operation were already in place). Alternatively, defending officers might rely on compliance with such ROE to demonstrate that, taking account of their purposes and the contexts in which they acted, their conduct satisfied the Fourth Amendment's "reasonableness" standard. What then are ROE?

ROE are "most fundamentally... the means by which the National Command Authorities... express their intent as to how force will and will not be used to achieve policy objectives."[58] It is critical to understand that well-crafted ROE use military means to serve an ultimately political purpose:

> [R]ules of engagement must... be carefully written so as to preclude actions that might run counter to national policy. The process requires sensitivity to the distinction between purpose and means.... The proper measure for success... is not the extent to which violations occur, but rather the congruency of the operation's execution with its underlying political purpose."[59]

To take a simple example, ROE may be crafted to minimize the likelihood of civilian deaths in an air campaign targeting an enemy's infrastructure, if only because the nation's political leadership considers it vital to win over the local population's support in pursuing the war. When properly designed, ROE "have three underlying bases that operate in tandem[:] ... policy, law, and operational concerns."[60] Ordinarily, therefore, sound ROE "are best drafted by a team consisting of a judge advocate and [a military] operator [such as a pilot], and must be reviewed at an appropriate policy level."[61] In combat situations overseas, ROE must take account of the restraints on force imposed by international *jus in bello*, specifically including the overarching principles that the use of force must be both necessary and proportionate.[62] Observance of these legal requirements will, of course, tend to restrict and reduce the violence used to achieve an operation's objectives.

If ROE were developed for the use of the military in armed counter-terrorism operations inside the United States, the policy component would obviously be of quite extraordinary importance. Political and military leaders would (rightly) be extremely sensitive to the effect that the use of military force would have on American civilians and American property, and would surely endeavor by all means possible to ensure that mission achievement in purely military terms was not bought at the price of unnecessary loss of the lives of innocent American men, women, or children, avoidable destruction of their property, or offensiveness to domestic public opinion. Military planners and commanders would be acutely aware that the eyes of Congress, of the media, and of the public at large, were constantly upon them. Political leaders facing elections would also go to great lengths to restrict the use of force domestically to the most imperatively demanding situations.[63]

If these assumptions are sound, then the question whether the Fourth Amendment applied to such operations or not could turn out to make little practical difference in many, perhaps most, cases. Even if reviewing courts determined that the Amendment's reasonableness standard *did* apply, they would likely be highly deferential to the operational commander's decisions about the use of force.[64] Moreover, the courts would also be likely to categorize the operation as falling under the case law's "special needs" exception to the warrant requirement—a conclusion that the 10/23/01 OLC Opinion reached.[65]

Although the case law is rather meager, several judicial decisions lend support to the view that military operations against terrorists, even when carried out on U.S. soil, are not subject to the Fourth Amendment's probable cause and warrant requirements. For example, in *United States v. Green*, the Fifth Circuit examined the Fourth Amendment claim of a driver who had been stopped at a roadblock checkpoint on a military base, found to be without a license or proof of insurance, fled, was captured and arrested by military police, and discovered to be in possession of crack cocaine.[66] She sought to have the evidence suppressed on the grounds that it was the fruit of an "unreasonable" search. The checkpoint stop was suspicionless; the military police were stopping every sixth car (of which the driver's was one) to search for terrorists and so deter terrorism. The court agreed that the search would have been unconstitutionally unreasonable if it had been undertaken in a search for criminals or an attempt to prevent general criminal activity. But it found that the "more narrow" purpose of the checkpoint was "to protect a military post, distinct from a general law enforcement mission."[67] Because this purpose was more like a traditional military function than criminal law enforcement, the court held that the driver could not rely on the Fourth Amendment.

Also relevant is the European Court of Human Rights' decision in *Murray v. United Kingdom* under Article 8 of the European Convention on Human Rights—a provision resembling the Fourth Amendment.[68] In *Murray*, the British Army visited Murray's home in order to arrest and question her regarding her suspected terrorist activities on behalf of the Irish Republican Army. After entering, the Army searched the house for other occupants and forced the Murray family to gather in one room. The Court observed that this was not an ordinary law enforcement action but instead a military security one. It noted the "responsibility of an elected government in a democratic society to protect its citizens and its institutions against the threats posed by organized terrorism and... the special problems involved in the arrest and detention of persons suspected of terrorist-linked offences."[69] Accordingly, the Army's actions satisfied Article 8's "necessity" requirement.[70]

Conclusion

I continue to believe that the main conclusions reached in the 10/23/01 OLC Opinion were sound. Despite the shrill criticisms of the opinion, I have yet to see a convincing refutation of it. Let that be a challenge to those of you who think otherwise. I am open to persuasion, as you should be.[71] Let Law and Reason decide.

Endnotes

1 *See* Neil A. Lewis, Memos Released by Justice Dept. Show Wide View of Bush's Power, The New York Times at A1 (Mar. 3, 2009); R. Jeffrey Smith & Dan Eggen, Post-9/11 Memos Show More Bush-Era Legal Errors, The Washington Post at A5 (March 3, 2009).

2 Quoted in Smith & Eggen, *supra* n.1.

3 Rosa Brooks, Bush's big lies: Behind the sordid memos that purported to give legal justification for the war on terror, The Los Angeles Times (March 5, 2009), available at http://www.latimes.com/news/opinion/commentary/la-oe-brooks5,0,2289353.column.

4 Lawrence Rosenthal, Does John Yoo belong in front of a class? No: His work for the Bush administration suggests he's incompetent, The Los Angeles Times (Apr. 9, 2009), available at http://www.latimes.com/news/opinion/commentary/la-oe-rosenthal9-2009apr09,0,592367.story.

5 Newsweek later retracted and apologized for Isikoff's story, but not before it had sparked riots in the Muslim world that resulted in 17 deaths. *See* Katherine Q. Seelye & Neil A. Lewis, Newsweek Says It Is Retracting Koran Report, The New York Times (May 17, 2005), available at http://query.nytimes.com/gst/fullpage.html?res=9E02E5DF1639F934A25756C0A9639C8.

6 Michael Isikoff, Extraordinary Measures: A new memo shows just how far the Bush administration considered going in fighting the war on terror, Newsweek Web Exclusive, available at http://www.newsweek.com/id/187342/output/print. Since the memo in question was addressed to the Fourth Amendment, it might seem remarkable that Isikoff presented it as a threat to the First Amendment. He was apparently alluding to one paragraph in the 37-page single spaced memo that quoted a Supreme Court decision, *Near v. Minnesota ex rel. Olson*, 283 U.S. 697, 716 (1931), for the (surely uncontroversial) proposition that the press has no liberty during wartime to publish the departure times of troop ships.

7 *See* Donna Miles, Obama Considers Deploying National Guard to U.S.-Mexico Border, American Forces Press Service (March 12, 2009), available at http://www.defenselink.mil/utility/printitem.aspx?print_http://www.defensselink.mil/news/.

8 *See* Memorandum for Alberto B. Gonzales, Counsel to the President, and William J. Haynes II, General Counsel, Department of Defense, from John C. Yoo, Deputy Ass't Att'y Gen. and Robert J. Delahunty, Special Counsel, Re: Authority for Use of Military Force To Combat Terrorist Activities Within the United States (Oct. 23, 2001) (hereafter the 10/23/01 OLC Opinion). The opinions released in March were originally available at http://www.usdoj.gov/opa/pr/2009/March/09-ag-181,html.

9 10/23/01 OLC Opinion at 2.

10 I say this deliberately and in full awareness of the Memorandum to Files prepared by Steven Bradbury, Principal Deputy of OLC, on October 6,

2008, which identified the propositions at issue as "either incorrect or highly questionable." *See* Memorandum for the Files from Steven G. Bradbury, Principal Deputy Ass't Att'y Gen., Re: October 23, 2001 OLC Opinion Addressing the Domestic Use of Military Force to Combat Terrorist Activities at 1 (Oct. 6, 2008). Respectfully, I disagree. I would further note that the 10/23/01 OLC Opinion stated that it considered the propositions only to represent "the better view" rather than as incontestable; that it drew attention to the novelty of the situation and the paucity of relevant judicial precedent; and that it argued that even if the Fourth Amendment were held to apply, courts would likely review and uphold properly conducted military operations under the "special needs" doctrine.

11 The Rand Corporation has published a full analysis of the attacks on Mumbai. *See* Angel Rabasa et al., The Lessons of Mumbai (2009), available at http://www.rand.org/pubs/occasional__psapers/2009/RAND0P249.pdf.

12 David A. Klinger & Lt. Col. Dave Grossman (U.S. Army Ret.), Responses to the September 11 Attacks: Who Should Deal with Foreign Terrorists on U.S. Soil? Socio-Legal Consequences of September 11 and the Ongoing Threat of Terrorist Attacks in America, 25 Harv. J. L. & Pub. Pol'y 815, 815 (2002).

13 For a factual account of that decision, *see* 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks Upon the United States 40-44 (2004).

14 For a careful study and chronology of the events, *see* Center for Counterproliferation Research, Working Paper: Anthrax in America: a Chronology and Analysis of the Fall 2001 Attacks (Nov. 2002), available at http://www.ndu.edu/centercounter/ANTHRAX%20CHRONOLOGY.pdf.

15 For example, on October 23, 2001—the day that the OLC Opinion was signed—Representative Richard Gephardt (D-MO) described the anthrax sent to Sen. Daschle's office as "weapons grade... highly sophisticated material" and said that while there was no proof of a link to al Qaeda, "I think we all suspect that." Quoted in *id.* at 44.

16 For an interpretation of the sweeping language of the statute that was close in time to its enactment, *see* Michael Stokes Paulsen, Youngstown Goes to War, 19 Const'l Comm. 215, 222-23, 250-57 (2002).

17 *See* U.N. Security Council Res. 1368 (Sept. 12, 2001); U.N. Security Council Res. 1373 (Sept. 28, 2001) (recognizing and reaffirming "the inherent right of individual or collective self-defence"); *see also* Letter from the U.S. Permanent Representative to the United Nations Addressed to the President of the Security Council (Oct. 7, 2001); Christopher Greenwood, International Law and the Pre-Emptive Use of Force, 4 San Diego Int'l L. J. 7, 17 (2003).

18 *See* Press Release, Statement by the North Atlantic Council on Collective Defense (Sept. 12, 2001).

19 *Compare* Philip Bobbitt, Terror and Consent: The Wars for the Twenty-First Century 128-34 (2008) (arguing that "the phrase 'a war on terror' is not an inapt metaphor, but rather a recognition of the way war is changing"), John Yoo, War by Other Means: An Insider's Account of the War on Terror 1-17 (2006), Daniel Statman, Targeted Killing, 5 Theoretical Inquiries in Law 178, 183-85 (2004), available at http://www.bepress.com/til/default/vol5/iss1/art7/, *and* Christopher Greenwood, International Law and the 'War against Terrorism,' 78 Int'l Aff. 301, 307-09. 314-16 (2002), *with, e.g.,* Monica Hakimi, International Standards for Detaining Terrorism Suspects: Moving beyond the Armed Conflict-Criminal Divide, 33 Yale J. Int'l L. 369, 376-79 (2008), Robert Jervis, American Foreign Policy in a New Era 46-51 (2005), and Noah Feldman, Choices of Law, Choices of War, 25 Harv. J. L. & Pub. Pol'y 457, 458-66 (2002). For a review of recent case law and legal scholarship on the circumstances in which an "armed conflict" exists, *see* Mark Osiel, The End of Reciprocity: Terror, Torture, and the Law of War 114-23, 130 (2009).

20 Of particular interest and value is Gregory E. Maggs, Assessing the Legality of Counterterrorism Measures Without Characterizing Them as Law Enforcement or Military Action, 80 Temp. L. Rev. 661 (2007).

21 *See* the 9/11 Commission Report, *supra* n.13, at 72-3; Robert Chesney & Jack Goldsmith, Terrorism and the Convergence of Criminal and Military Detention Models, 60 Stanford L. Rev. 1079, 1094-95 (2008) (outlining evolution of "crime versus war" debate before 9/11). It should be pointed out, however, that the United States' attacks in 1998 on a pharmaceutical plant in Sudan that was believed to be producing chemical weapons for al Qaeda assumed a state of war between the United States and that organization. The Clinton Administration had argued that the attacks were justifiable as lawful self-defense under Article 51 of the U.N. Charter—thus implicitly assuming the applicability of the war paradigm to the conflict with al Qaeda. Indeed, National Security Adviser Sandy Berger described the attack in Afghanistan as directed to "a military target," and anonymous Administration sources stated that that attack was designed to kill Osama bin Laden and as many of his associates as possible. *See* Sean D. Murphy, Legal Regulation of Use of Force, Missile Attacks on Afghanistan and Sudan, Contemporary Practice of the United States Relating to International Law, 93 Am. J. Int'l L. 161, 163-5 (1999). For legal analyses close in time to the 1999 missile strikes, *see* Ruth Wedgwood, Responding to Terrorism: The Strikes Against bin Laden, 24 Yale J. Int'l L. 559 (1999); W. Michael Reisman, International Legal Responses to Terrorism, 22 Hous. J. Int'l L. 3, 47-49 (1999).

22 *See* Chesney & Goldsmith, *supra* n.21, at 1096-99 (summarizing reasons why criminal model was thought inadequate to new threat posed by al Qaeda); *see also* the exchange Kenneth Roth and Ruth Wedgwood, Combatants or Criminals? How Washington Should Handle Terrorists, Foreign Affairs (May/June 2004).

23 334 U.S. 742 (1948).

24 *Id.* at 779-80.

25 42 A.B.A. Rep. 232 (1917). Hughes had been the Republican candidate for President in the year preceding this address. His distinguished career, it will be recalled, included service as an Associate Justice, as Chief Justice, as Secretary of State, and as Governor of New York.

26 Quoted in 334 U.S. at 781.

27 *Id.* at 782.

28 Letter from Abraham Lincoln, President of the United States, to Erastus Corning and Others (June 12, 1863), quoted and discussed in Michael Stokes Paulsen, The Constitution of Necessity, 79 Notre Dame L. Rev. 1257, 1277-79 (2004).

29 On the background of the Third Amendment, its ratification history, and the practice and (sparse) case law under it, *see* Tom W. Bell, The Third Amendment: Forgotten but Not Gone, 2 Wm. & Mary Bill of Rights J. 117 (1993).

30 *See, e.g.,* Hudson v. Michigan, 547 U.S. 586, 603 (2006) (Kennedy, J., concurring in part and concurring in the judgment) ("Privacy and security in the home are central to the Fourth Amendment's guarantees as explained in our decisions and as understood since the beginnings of the Republic.").

31 This possibility, moreover, is not in terms limited to occasions of *declared* war, but arguably may arise when war exists as a matter of fact, and not only when it is has been formally declared. *Cf.* The Prize Cases, 67 U.S. 635, 668-69 (1862).

32 "[N]or shall private property be taken for public use, without just compensation."

33 120 U.S. 227 (1887).

34 *Id.* at 234.

35 344 U.S. 149 (1953).

36 *Id.* at 155-56.

37 *See* El-Shifa Pharmaceutical Industries Co. v. United States, 378 F.3d 1346, 1356 (Fed. Cir. 2004), cert. denied, 545 U.S. 1139 (2005) (President Clinton had authority to determine that pharmaceutical plant in Sudan was part of al Qaeda weapons supply chain and thus "enemy" property constituting legitimate target for missile strike).

38 212 U.S. 78 (1909).

39 *Id.* at 85.

40 For an explanation of the legal culture's changing views of civil liberties in wartime, *see* Jack Goldsmith & Cass Sunstein, Military Tribunals and the Legal Culture: What A Difference Sixty Years Makes, 19 Const'l Comm. 261, 284-89 (2002).

41 327 U.S. 304, 314 (1946). Also of note, in its 1987 decision in *United States v. Salerno*, the Court said that "in times of war or insurrection, when

society's interest is at its peak, the Government may detain individuals whom the Government believes to be dangerous." 481 U.S. 739, 748 (1987). And even in dissent, Justice Stevens acknowledged that "it is indeed difficult to accept the proposition that the Government is without power to detain a person when it is a virtual certainty that he or she would otherwise kill a group of innocent people in the immediate future." *Id.* at 768 (Stevens, J., dissenting).

42 Vernonia School Dist 47J v. Acton, 515 U.S. 646, 653 (1995).

43 Indianapolis v. Edmond, 531 U.S. 32, 44 (2000).

44 *See, e.g.,* Cassidy v. Chertoff, 471 F.3d 67, 82 (2d Cir. 2006) (Sotomayor, J.) ("[T]he prevention of terrorist attacks on large vessels engaged in mass transportation and determined by the Coast Guard to be at heightened risk of attack constitutes a 'special need.' Preventing or deterring large-scale terrorist attacks present problems that are distinct from standard law enforcement needs and indeed go well beyond them."); *see also* MacWade v. Kelly, 460 F.3d 260. 272 (2d Cir. 2006).

45 Gerstein v. Pugh, 420 U.S. 103, 125 n.27 (1975).

46 South Dakota v. Opperman, 428 U.S. 364, 370 n.5 (1976).

47 *See, e.g.,* Griffin v. Wisconsin, 483 U.S. 868 (1987) (search by probation officers); Marshall v. Barlow's, Inc., 436 U.S. 307 (1978) (administrative search).

48 United States v. Verdugo-Urquidez, 494 U.S. 259, 267 (1990)(plurality op.); *see also* Boyd v. United States, 116 U.S. 616, 625-26 (1886); City of West Covina v. Perkins, 525 U.S. 234, 247 (1999) (Thomas, J., concurring in the judgment).

49 494 U.S. at 267.

50 *Id.* at 273.

51 *Id.* at 274.

52 10/23/01 OLC Opinion at 28-9.

53 *Id.* at 18.

54 18 U.S.C. § 1385.

55 *See, e.g.,* Application of the Posse Comitatus Act to Assistance to the United States National Central Bureau, 13 Op. O.L.C. 195, 197 (1989) (cited and discussed in 10/23/01 OLC Opinion at 17).

56 *See* Bissonette v. Haig, 776 F.2d 1384, 1388 (8th Cir. 1985).

57 *See* Chairman of the Joint Chiefs of Staff Instruction: Standing Rules of Engagement for US Forces, Jan. 15, 2000, available at http://www.fas.org/man/dod-101/dod/docs/cjcs__sroe.pdf. In addition to "standing" ROE, operation-specific ROE are also issued. ROE may be developed for domestic as well as overseas operations. *See also* Lt. Col. Guy R. Phillips, Rules of Engagement: A Primer, 1993 Army Law. 4, 8. *See generally* Richard J. Grunawalt, The JCS Standing Rules of Engagement: A Judge Advocate's Primer, 42 A.F.L. Rev. 245 (1997).

58 Michael N. Schmitt, Clipped Wings: Effective and Legal No-Fly Zone Rules of Engagement, 20 Loy. L.A. Int'l & Comp. L. J. 727, 741-2 (1998).

59 *Id.* at 742.

60 *Id.* at 741.

61 *Id.* at 751.

62 *Id.* at 744-45.

63 Critics might question this, citing, for example, Attorney General Janet Reno's use of deadly force against American civilians, including young children, in Waco, Texas—an incident that seems not to have damaged her standing. For example, the New Yorker magazine published a sympathetic profile of Reno two years after the event. *See* Peter Boyer, The Children of Waco, The New Yorker (May 15, 1995), available at http://www.pbs.org/wgbh/pages/frontline/waco/childrenofwaco1.html.

During the 51-day "siege" of the Branch Davidians in Waco, the government deployed nine Bradley fighting vehicles, five combat engineering vehicles, one tank retrieval vehicle, and two Abrams tanks. The FBI reported that a minimum of 719 law enforcement personnel were committed to being on-site during any given day of the "siege." In the fire that incinerated the Branch Davidians' compound, 76 people died—more than 20 of them children. The government is thought to have used CS gas (orthochlorobenzyladine malononitrile) against the Branch Davidians in a manner that would have violated the instructions in an Army Field Manual, had they applied. For an account of the events and the reactions to them, *see* Dean M. Kelley, Waco: A Massacre and Its Aftermath, First Things (May 1995), available at http://www.firstthings.com/article/2008/09/001-waco-a-massacre-and-its-aftermath--12. But perhaps Waco was an unusual case.

64 *See* Paulsen, The Constitution of Necessity, *supra* n.28, at 1279 (arguing that "[a] search that might otherwise be unreasonable might become reasonable in an emergency" such as would be posed by the "danger of a nuclear bomb being detonated" by terrorists).

65 *See* 10/23/01 OLC Opinion at 34-5.

66 293 F.3d 855 (5th Cir.), cert. denied. 123 S. Ct. 403 (2002).

67 *Id.* at 859.

68 19 Eur. Ct. H.R. 193 (1994). This is not, of course, to say that Murray can provide a rule of decision in constitutional cases in the United States. *See* Robert J. Delahunty & John C. Yoo, 29 Harv. J. L. & Pub. Pol'y 291 (2005).

69 19 Eur. Ct. H.R. at at 234.

70 Section 1 of Article 8 guarantees the "right to respect for [one's] private and family life, his home and his correspondence." Section 2 states that there shall be no interference by a public authority with this right "except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or [in certain other cases]." European Convention for the Protection of Human Rights and Fundamental Freedoms, Nov. 4, 1950, 312 U.N.T.S. 221, E.T.S. 5, as amended.

71 For example, it would be instructive to discover whether President George Washington considered it necessary or advisable to seek search or arrest warrants when leading federal forces in suppressing the Whisky Rebellion in Pennsylvania in 1794. This episode occurred only three years after the ratification of the Fourth Amendment, and became "a precedent for all future use of federal military force in domestic disorders." Stephen I. Vladek, Emergency Power and the Militia Acts, 114 Yale L.J. 149, 162 (2004). Washington's most recent scholarly biographer states that the federal forces (under Colonel Henry Lee) "rounded up some 150 alleged rebels—many were dragged from their homes in the middle of the night." John Ferling, The Ascent of George Washington: The Hidden Political Genius of An American Icon 337 (2009).

The leading roles played by two pre-eminent Founders—George Washington and Alexander Hamilton, who accompanied him on the campaign—in putting down the rebellion give great weight to the precedent in determining whether the Fourth Amendment is applicable in this context. The statute under which Washington acted, the Calling Forth Act of May 2, 1792, ch. 18, 1 Stat. 264, was enacted by the Second Congress. Sections 1-2 vested the authority in the President to call out the militia in certain emergencies, and sections 3-8 prescribed detailed rules to be followed during their deployment. (For example, section 3 required the President to order insurgents to disperse before using force against them.) But nothing in the statute indicated any need to seek or obtain search or arrest warrants; indeed, given that section 2 became operative when the execution of federal law was obstructed "by combinations too powerful to be suppressed by the ordinary course of judicial proceedings," it might have been impossible to obtain warrants from local federal (though perhaps not state) magistrates. Furthermore, the Supreme Court construed section 9 of the same statute, relating to federal marshals, to demonstrate Congress' early approval of *warrantless* arrest authority. *See* United States v. Watson, 423 U.S. 411, 420 (1976) ("'The Second Congress thus saw no inconsistency between the Fourth Amendment and legislation giving United States marshals the same power as local police officers to arrest for a felony without a warrant"). *But see* Thomas Y. Davies, Recovering the Original Fourth Amendment, 98 Mich. L. Rev. 547, -- n.171 (1999).

Litigation

Constitutional Limits on Punitive Damages: An Evaluation of the Role of Economic Theory in Prescribing Constitutional Constraints on Punitive Damage Awards

*by Dorothy Henderson Shapiro**

Over the last fifteen years, the Supreme Court has formulated new constitutional principles to constrain punitive damages awards imposed by state courts, invoking its authority under the Due Process Clause of the Fourteenth Amendment. This intervention has been controversial from the start, generating dissents from several Justices asserting that the actions of the Court are unwarranted and amount to unjustified judicial activism. Over the ensuing years lower courts and commentators have criticized the Court's prescription of procedural and substantive limitations, finding them to be vague and unnecessarily restrictive of state common law prerogatives. Some observers with an economic orientation have entered the debate, motivated by runaway punitive damage awards in some states. Their core premise is that actual damages alone are sufficient in most tort cases, and that adding punishment on top of compensation creates inefficient incentives to adopt unnecessary precautions against negligent or reckless harm.

Only in instances in which the tort would escape detection is it economically efficient for courts to impose punitive damages. This is an application of deterrence theory, which reflects the probability of a covert injury that is never remedied. Deterrence theory aims at eliminating tortious behavior by eliminating incentives to commit a tort; to achieve appropriate deterrence, injurers should be made to pay only for the amount of harm their conduct generates. As we will see, actual damages result in optimal deterrence except for cases in which the injurer has escaped detection for similar torts, and in these instances, a punitive award would be appropriate.

The Supreme Court does not, however, possess general authority to impose economically efficient standards for the award of punitive damages. As it stated in cases like *BMW v. Gore*,[1] its authority under the Due Process Clause is different and more limited: it sits to review punishments that are "arbitrary" in procedure or amount, and to reject "outlier awards" which cannot reasonably be anticipated by persons accused of torts. Within this limited framework, the Court has intervened in state court cases and reversed punitive damage judgments that have shocked the judicial conscience to the degree that they are deemed to be in violation of the U.S. Constitution. The Court has so far articulated a few constitutional standards on an incremental and tentative basis. Those standards focus chiefly on the concept of reprehensibility and the ratio between actual and punitive damages.

* *The author wishes to thank Professor Stephen V. Marks, PhD, Professor of Law and Economics at Pomona College, who supervised her work.*

While economic evaluations of the issues have been very pertinent to punitive damages reform at the state level by common law evolution or statutory amendment, they have had little noticeable impact on the Supreme Court's articulation of constitutional standards and safeguards. It is not obvious how economic analysis fits within the limited legal principles the Supreme Court applies under the Due Process Clause. In addition, the proposals of economists seem very difficult to implement in a practical way, because lay juries must apply the law to complex and disputed facts. Deterrence theory asks jurors to estimate probabilities that cannot be pinpointed and are largely a matter of speculation. Also, there are differences within the published economic literature addressing punitive damage reforms. We will see that in significant ways expert economists disagree among themselves. Given this background, it is perhaps not surprising that the Supreme Court pays little attention to economic teaching in formulating constitutional limitations.

I. MAJOR SUPREME COURT CASES AND THEIR OUTCOMES: THE EVOLUTION OF PUNITIVE DAMAGE RESTRICTIONS

In a 1996 landmark decision, *BMW v. Gore*, the Supreme Court struck down a punitive damage award that it found unacceptable under the Fourteenth Amendment of the Federal Constitution. In his majority opinion, Justice Stevens declined to draw a bright-line limitation or set a cap on punitive damages. He did, however, prescribe three guideposts that he believed should guide the lower courts in their deliberations—"the degree of reprehensibility [of the action], the disparity between the harm or potential harm suffered by the plaintiff and the punitive damages award, and the difference between the punitive remedy and the civil penalties authorized."[2] The Court found that BMW's action[3] was not particularly reprehensible (no reckless disregard for health and safety or evidence of bad faith). The ratio of punitive to actual damages was 500/1, which was suspiciously steep. And the civil penalty sanctioned by the legislature for similar conduct could not exceed $2,000. Thus, the Court concluded that the punishment meted out was "grossly excessive" and violated substantive due process.[4] The Court added, with an element of deliberate vagueness, that these guideposts could be overridden as necessary to deter intentional torts in the future.[5] The dissent, written by Justice Scalia, argued that the identification of a "substantive due process right" against a grossly excessive award is not specified in the U.S. Constitution, and is thus "an unjustified incursion into the province of state governments."[6] Justice Scalia also found fault with the vagueness of the guideposts, calling them a "road to nowhere."[7]

In this case, the Court struggled with the idea of global punishment. Justice Stevens clearly explained that a State may not "impose economic sanctions on violators of its laws with the intent of changing the tortfeasors' lawful conduct in other states,"[8] but later acknowledged that repeated misconduct is more reprehensible, and thus must be taken into consideration.[9] Of special relevance to economic analysis, we see in *BMW v. Gore* an early attempt to delimit the kind of punishment that is necessary to deter conduct that may recur and affect more than one litigant.[10]

In *State Farm Mutual Automobile Insurance Company v. Campbell*,[11] the Court deemed a punitive damages award of $145 million on top of $1 million in actual damages to be unconstitutional, applying the guideposts articulated in *Gore*.[12] The Court once again stated that the purpose of punitive damages is "deterrence and retribution," while actual damages serve to redress a plaintiff's loss.[13] It also refused to articulate a "bright line" ratio that a punitive damage award cannot exceed. Significantly, however, the Court observed that "few awards exceeding a single-digit ratio between punitive and compensatory damages, to a significant degree, will satisfy due process."[14] Justice Kennedy, writing for the majority, explained that a ratio of 4-1 might be "close to the line of constitutional impropriety," but later noted that greater ratios can comport with due process in cases in which a particularly egregious act results in a small amount of actual damages.

Finally, the Court in *State Farm* suggested that there is a danger of abuse in punitive awards because they may be meted out in an extravagant manner to strike at a defendant simply because of its size or wealth, without regard to the legal rationale for a punitive award. The majority opinion explained: "The wealth of a defendant cannot justify an otherwise unconstitutional punitive damages award."[15] This issue has received some attention from law and economics commentators, most of whom conclude that corporate wealth should not influence the size of the punitive damage award.[16]

In both *Gore* and *State Farm*, the Court emphasized that punitive sanctions should not be imposed on behavior deemed lawful in other states, and that evidence of alleged misdeeds in other states should not be used to influence the jury's decision. In *Philip Morris v. Williams*,[17] the Court went a step further and held that due process precludes punitive damages awarded for harm to non-parties (parties not involved in the litigation). The Court explained that juries should not be allowed to speculate about harm to non-parties because the defendant could not mount a defense to claims of that sort based on facts and circumstances not truly before the court, and this would violate the defendant's due process rights. The Supreme Court also emphasized that "permitting such punishment would add a near standardless dimension to the punitive damages equation and magnify the fundamental due process concerns of this Court's pertinent cases—arbitrariness, uncertainty, and lack of notice."[18]

In this opinion, the majority addressed an important unresolved question when it ruled evidence of harm to nonparties out of bounds. It distinguished sharply between considering harm to nonparties in general terms as it may bear on reprehensibility, which it would allow, and considering harm to nonparties as direct evidence of the amount of damages to impose, which it would not allow. A jury instruction is now required on this elusive issue, in compliance with *Williams*. The Court declared that the state courts may no longer authorize a procedure "that creates an unreasonable and unnecessary risk of any such confusion occurring," such as this one, calling the risk of misunderstanding "a significant one."[19]

In the latest case to be decided involving punitive damages restrictions, *Exxon Shipping Company v. Barker*,[20] the Court hinted that in the future it will consider adopting a 1:1 test (limiting the punitive award to the amount of actual damages) as a constitutional matter, assuming actual damages are substantial.[21] Prior decisions, in which the Court explained that certain restrictions should apply in certain cases, have not had the limiting effect that the Court hoped. The Supreme Court expressed concern that the lower courts have treated the high Court's flexible, multi-factor standard as no standard at all. The Court in *Exxon* explained that "the real problem is the stark unpredictability of punitive awards."[22] In this opinion, Justice Souter evaluated several ideas that might restore substantive and procedural fairness to the assignment of punitive damages. He rejected the idea of a specific dollar cap because there is no "standard" tort or contract injury. But Justice Souter, writing for the majority, believed that a strict ratio or maximum multiple is a promising alternative. Along these lines, he noted that the median ratio of punitive to compensatory awards found in a series of studies was less than 1:1,[23] meaning that the compensatory award exceeded the punitive award in most cases. Justice Souter explained that, "In a well functioning system, awards at or below the median would roughly express jurors' sense of reasonable penalties in cases like this one that have no earmarks of exceptional blameworthiness. Accordingly, the Court finds that a 1:1 ratio is a fair upper limit in such maritime cases."[24] Although this case dealt with federal maritime law,[25] there are many hints in the opinion that the Court is now considering extending this reasoning to all state court cases evaluated under the Due Process Clause.

In each of these extensive Supreme Court opinions, economic theory is conspicuous by its absence. The Justices occasionally mention the ideas proposed by economists, but in most cases ignore their theories altogether. I will give examples of this trend in later sections of this paper. In order to see what is missing, it is important to understand the economic theory of punitive damages as it stands today.

II. ECONOMIC THEORY OF PUNITIVE DAMAGES

Robert Cooter, professor of law and economics at Boalt Law School at the University of California Berkeley, once wrote, "Litigating a tort dispute involving punitive damages, much like navigating the Straits of Magellan, runs the risk of incurring grave losses from colliding with unseen objects."[26] Some economists have addressed the issue of punitive damages, trying to suggest a system that would achieve legal certainty in uncharted territory. The consequences of miscasting the governing legal rules are grave, say these scholars. By erring on the side of too great a punishment, the Court runs the risk of over-deterrence, which would have harmful effects for innovators and would raise product prices, thus harming

American consumers and industries. On the other hand, too little punishment would result in under-deterrence, which would invite repeat offenses by parties who found it cost-effective to continue unlawful behavior. (It has been widely asserted by economics scholars that punitive damages should only be awarded in cases that involve intentional torts, although in practice in state courts today punitive damages are awarded based on asserted "reckless" behavior, which often shades into "gross negligence").

Legal treatises and the previously cited Supreme Court opinions consistently state that the purpose of punitive damages is "punishment and deterrence." In *Williams*, the Court acknowledged this basic proposition only to abandon the deterrence aspect and focus on the idea of punishment alone throughout the remainder of the opinion. It seems as though the Court has neglected to take into account the proper role of optimal deterrence in punitive damages law. Economists, aware of this somewhat confusing legal definition, articulate the purpose of punitive damages in different but related ways, and all focus on deterrence with little mention of punishment. Economists are able to justify punitive damages from a deterrence point of view because the incentives to conform to the law are insufficient without them.

The only time that the violation of a legal standard is profitable, according to Cooter, is when enforcement error reduces the injurer's expected liability. This is the central point in Cooter's "rule of the reciprocal," a formula that accounts for situations in which violators will avoid paying full compensatory damages because not all victims will bring suit in a court of law. He writes, "in general, the punitive multiple should equal the reciprocal of the *enforcement error* for the sake of deterrence,"[27] The major forms of enforcement error include a victim's failure to assert claims that they are legally entitled to recover and "under compensation" of successful plaintiffs. Deterrence theory does not aim to make the victim whole, but rather to punish the injurer for the amount of harm for which he or she is responsible and to encourage economically sensible precautions.

Economists Mitchell Polinsky and Steven Shavell have continued the discussion of the appropriate role of punitive damages.[28] Like Cooter, they believe that the goal of punitive damages is to achieve optimal deterrence so that parties take efficient precaution against harm. Polinsky and Shavell explain: "to achieve appropriate deterrence, injurers should be made to pay for the harm their conduct generates, not less, not more."[29] Thus, the only time punitive damages can be justified at all, according to Polinsky and Shavell, is when injurers can escape liability for harms for which they are responsible. "If they do," write Polinsky and Shavell, "the level of liability imposed on them when they are found liable needs to exceed compensatory damages so that, on average, they will pay for the harm that they cause."[30] Thus, the authors offer an economic justification for the efficient allocation of punitive damages.

Polinsky and Shavell expand on Cooter's initial point that not every illegal action will be detected and reported. Polinsky and Shavell set forth an equation that can be used to calculate the appropriate amount of total damages: if an injurer has been found liable, total damages should equal the harm caused multiplied by the reciprocal of the probability of being found liable. They explain, "we believe that courts and juries often will be able to obtain enough information about the likelihood of escaping liability to apply the theory reasonably well."[31] Despite this assertion, evidence to the contrary has been amassed by critics of the Polinsky and Shavell economic model, as I explain at the end of this essay.

The proposition that punitive damages should be based solely on deterrence of behavior that is not caught by the compensatory damages web is supported by eminent legal scholars such as Yale Law School Professor George Priest. But Priest broadens the set of possible rationales for punitive damage awards, explaining that punitive awards may be needed to remedy defects in compensatory damage awards, "such as juries awarding damages that are too low in some dimension or some set of injuries going undetected or perhaps being too insignificant individually to justify litigation."[32]

III. WHY IS ECONOMIC ANALYSIS MISSING FROM THE SUPREME COURT'S OPINIONS?

It can be seen from the above that many economists concur in the idea of optimal deterrence. They propose detailed theories and even mathematical formulas that they hope will aid the courts in formulating a systematic and exact method for calculating the appropriate amount of punitive damages. It surely must be disconcerting to these scholars that the Supreme Court has ignored the economic analysis proposed, and contradicted large portions of economic thinking.

In *Browning Ferris v. Kelco*,[33] the defendant attempted to argue that excessive damages awards were invalid under the Excessive Fines Clause of the Eighth Amendment. This was the first time that punitive awards received constitutional scrutiny, and in the end the Court (per Justice Blackmun) rejected that argument. In her dissenting opinion, Justice O'Conner dismissed the economic theory that both sides had employed in their arguments, noting, "the Constitution does not incorporate the views of the Law and Economics School," nor does it "require the States to subscribe to any particular economic theory."[34] Neither the majority opinion nor Justice O'Conner in dissent accepted a link between constitutional law and economic theory.

In later opinions, the Justices mention prominent economists by name, but still refuse to put their precepts into practice. In his concurring opinion in *BMW v. Gore*, Justice Breyer mentions Professors Shavell and Cooter and summarizes his understanding of their standard for punitive damages. Justice Breyer correctly interprets the economic theories introduced by Shavell. He notes:

> Some economists… have argued for a standard that would deter illegal activity causing solely economic harm through the use of punitive damages awards that, as a whole, would take from a wrongdoer the total cost of the harm caused…. My understanding of the intuitive essence of some of these theories, which I put in crude form (leaving out various qualifications), is that they could permit juries to calculate punitive damages by making a rough estimate of global harm, dividing that estimate by a similarly rough estimate of the number of successful lawsuits that would likely be

> brought, and adding generous attorney's fees and other costs. Smaller damages would not sufficiently discourage firms from engaging in the harmful conduct, while larger damages would 'over-deter' by leading potential defendants to spend more to prevent the activity that causes the economic harm, say, through employee training, than the cost of the harm itself.[35]

He concludes by explaining that the record contained nothing that might suggest that the Alabama Supreme Court applied any economic theory that might explain the high punitive award. He then rejects "reference to a constraining 'economic' theory, which might have counseled a more deferential review" by the Court.

After Justice Breyer's summation of economic theory, Shavell and Cooter disappear from the remainder of the *BMW v. Gore* opinion. In *State Farm*, Justice Kennedy summarizes some of the arguments for the plaintiffs, including "the fact that State Farm will only be punished in one out of every 50,000 cases as a matter of statistical probability."[36] In dismissing this argument, Kennedy joins this economically legitimate argument with another; he writes, "Here, the argument that State Farm will be punished in only the rare case, coupled with the reference to its assets, had little to do with the actual harm sustained by the Campbells. The wealth of a defendant cannot justify an otherwise unconstitutional punitive damages award."[37] The Court deals with the corporate wealth argument, while pushing the detectability issue into the background, claiming that it lacks proximity to the victim. This misapprehends the economic substance of deterrence theory. The Court requires the jury to focus on total harm caused, and not the amount needed to effectively deter future misdeeds. The Court's analysis pulls the issue in a different direction than the economists would like to go.

Finally, in *Exxon*, Justice Souter mentions Shavell and Polinsky's detection theory without coming to grips with it. Justice Souter spends several pages debating the pros and cons of limiting punitive damages in a specific way (something the Court had explicitly declined to do just fifteen years previously). In doing so he writes, "Heavier punitive awards have been thought to be justifiable when wrongdoing is hard to detect (increasing chances of getting away with it),"[38] and then cites snippets from two other Supreme Court opinions. This is the last we hear of Shavell, as Justice Souter moves forward and leaves this trace of economic theory in the dust.

Despite the variety of articles written by economists on the subject of punitive damages in the last twenty years, economic theory has been largely brushed aside by the Supreme Court, notwithstanding the Court's obvious desire to fashion a more exacting rationale intended to limit punitive damages to promote certainty and predictability. It remains to be seen whether economic analysis can be fitted within the requirements of constitutional law, and whether it can operate in a practical manner to serve as the purposes the Supreme Court seeks to accomplish.

IV. MAJOR CONTRADICTIONS BETWEEN SUPREME COURT OPINIONS AND ECONOMIC THEORY

The *BMW v. Gore* guideposts ask the courts first and foremost to consider "reprehensibility" and "the ratio between punitive and actual damages," considerations not grounded in economic thinking. In fact, Shavell and Polinsky explicitly caution against using "reprehensibility" as a litmus test. They write, "That a defendant's conduct can be described as reprehensible is in itself irrelevant. Rather, the focus in determining punitive damages should be on the injurer's chance of escaping liability."[39] Cooter explains that punitive damages are appropriate to curb "gross faults" only because the perpetrator will most likely need a serious sanction in order to influence his incentives, not because the faults are especially egregious. But the Court has not made this distinction. The *Gore* opinion explains: "The most important indicium of the reasonableness of a punitive damages award is the degree of reprehensibility of the defendant's conduct."[40] It is true that using reprehensibility as a gauge does not equate to throwing economic analysis out the window. But by adopting a system that allows jury outrage to trump careful evaluation of deterrence, the Court is opening a Pandora's box of runaway jury decisions.

One problem with allowing "reprehensibility" to be a main criterion for evaluating punitive damages is that no universal moral code exists that would guide jury discretion. In evaluating the reprehensibility of a tortious action, the Supreme Court has provided a vague outline:

> [T]he Court must consider whether the harm was physical rather than economic, the tortious conduct evinced an indifference to or a reckless disregard of the health or safety of others, the conduct involved repeated actions or was an isolated incident, and the harm resulted from intentional malice, trickery, or deceit, or mere accident.[41]

All of these factors serve as proxies for especially egregious behavior. But the unpredictability problem stems from more than the characterization of a reprehensible act. Perhaps a particular jury may feel that an action is really bad, and perhaps they will unanimously feel a strong sense of outrage toward the perpetrator, and believe that he should be punished. But it is impossible to translate this feeling into a dollar amount that is predictable and consistent. Cass Sunstein and other prominent legal scholars conducted hundreds of controlled experiments in order to study this phenomenon. Their conclusion: "people have a hard time in arriving at consistent, predictable judgments when using the scale of dollars—even when their moral judgments are both consistent and predictable."[42] In addition, they found that juries are extremely vulnerable to lawyer suggestion when dealing with punitive damage instructions; evidently, it is not too difficult to convince a jury that someone's behavior is especially egregious when the right rhetorical skills are employed.

It is obvious why a focus on reprehensibility is unsettling to economists. But the second guideline enunciated in the *Gore* opinion seeks to limit the ratio between punitive and compensatory damages and may have a closer link to economic concerns. As mentioned above, the latest *Exxon* opinion hints

that a decisive "bright-line" one-to-one ratio may be the next limit placed on punitive damages. Seen through the Shavell and Polinsky lens, this means that the tortfeasor has at least a 50% chance of detection. Although the ratio cap serves the basic purpose of creating a concrete limit on excessive punitive damage awards, it does not allow for flexibility to apply economic analysis. Instead, Justice Souter explains in *Exxon* that, "an acceptable standard can be found in the studies showing the median ratio of punitive to compensatory awards… in a well functioning system, awards at or below the median would roughly express juror's sense of reasonable penalties in cases like this one that have no earmarks of exceptional blameworthiness."[43] Justice Souter offers a statistical theory, but economic theory is absent. The only exception that is allowed bows to the vague reprehensibility standard once again.

In "An Economic Evaluation of Punitive Damages," David Friedman remarks,

> If the common law does not follow the rule we think is economically efficient, that may be evidence that our economic analysis is wrong. It also may be evidence that something has gone wrong with the common law, or that whatever forces push it toward economic efficiency apply in only some areas and not others.[44]

Perhaps a third explanation for this discrepancy is that there is no practical way to implement the principles of deterrence theory that the economists advocate. The Supreme Court is struggling with strong economic and constitutional pressures, which cut in different directions. As a result, the current state of punitive damages law is understandably a hodgepodge of ambiguous formulations which serve no clear purpose other than "cutting back on awards" that strike the judicial conscience as excessive in particular circumstances. Lower courts, as a result, have little practical guidance.

V. RECOMMENDATION FOR POLICY

How can economic theory be reconciled with constitutional law in the field of punitive damages? As several of the Justices have noted, the Due Process Clause does not enshrine any particular economic theory as the law of the land, whether it is the theory of Mr. Herbert Spencer or that of Professor Shavell. States are entitled to have their own policies on punitive damages, both wise and unwise. In our federal system, individual states serve as laboratories to experiment with different economic policies. They learn from each other. And state legislatures supervise the handiwork of common law courts, including tort reform statutes. Congress too has the power to place limits on punitive damages. It is arguably undemocratic for the Supreme Court to intervene in this process and lay down its own, national standards, particularly standards that have proven to be so vague and controversial.

But there is a powerful constitutional concern at work in the field of punitive damages. If awards are freakish and unpredictable, the defendant has no fair notice of the magnitude of potentially crushing punishments. And if astronomical awards are imposed in response to jury outrage over "reprehensibility," inflamed by the rhetoric of lawyers, the award in the end may serve no legitimate purpose—punishment that exceeds the outer limits of substantive due process.

As previously noted, the Supreme Court is groping toward greater limitations, inching its way toward a 1-to-1 ratio limit. It sees greatest hope in adopting a mechanical litmus test of this kind. Such a standard can be understood by defendants and applied readily by state courts and juries. It is preferred to an open-ended "reprehensibility" analysis, which invites unlimited inflation in an award based on overblown indignation or even prejudice against a large, out-of-state defendant brought before the bar of a local court. One trouble with the 1-to-1 standard, however, is that it may not allow optimal deterrence to be achieved. The award might underdeter by failing to punish misconduct occurring in secrecy that could reoccur absent a large punitive award. And in cases involving a large compensatory damage awards, it might not effectively limit an excessive punitive award.[45]

The key to future reform lies in separating deterrence theory from the idea of punishment and making it the focus of punitive awards. Vague limitations by the Court have not been successful in combating excessive awards of punitive damages. Thus, the Court is right in moving toward a "bright-line" limit. But there has been too much focus in the Court's opinions on the retributive aspect of punitive damages, and too little focus on changing incentives by forcing injurers to internalize costs. Juries should be instructed in a way that removes their focus from the total harm generated by the injurer and sustained by the victims and instead focuses on ways to deter tortious behavior.

Is a 1:1 ratio cap appropriate? The predictability secured by a concrete limit has its pitfalls. Although a ratio cap would be a helpful guide in most situations, there are a few in which such an inflexible limit could have adverse results. As previously mentioned, if actual damages are quite small, a higher multiplier may be necessary to result in effective deterrence. There would be little incentive for a party to bring a suit in which actual damages are minimal, even if the behavior was especially egregious and warranted some kind of punishment. In these cases, the judicial system would not be able to deter wrongdoing as effectively, and worse, would deprive private parties of incentive to pursue wrongdoers, which is a basic function of tort law.

Justice Scalia mentions in his *Gore* dissent that the Fourteenth Amendment's Due Process Clause should not be seen as a guarantee against "unfairness." He explains, "What the Fourteenth Amendment's procedural guarantee assures is an opportunity to contest the reasonableness of a damages judgment in state court; but there is no federal guarantee a damages award actually *be* reasonable."[46] This textualist argument does not allow for the evolution of due process law, and certainly does not suggest any implied constitutional limit on punitive damage awards. But this legal view does sometimes mesh with economic theory in certain cases: for example, if an injurer had only been made to pay for harm generated in one of a hundred cases, Justice Scalia would not oppose a punitive award one hundred times greater than the actual.[47] But Justice Scalia would also tolerate economically inefficient awards more substantial than necessary to achieve deterrence.

It is true that for most cases, a 1-to-1 cap would have the beneficial effect of limiting runaway punitive awards and preventing gross over-deterrence that economists caution against. But a rule that makes economic sense would allow for

exceptions to the 1:1 rule in specific situations.[48] A case that would allow for a larger punitive award would not be judged based on reprehensibility. Instead, the jury would be given simple but specific instructions: is the tortious act easily detectable,[49] or is it likely to have occurred over time without disclosure?[50] Is the compensatory award quite small? A brief summary of basic deterrence theory may be appropriate, highlighting the risks of under- and over-deterrence for the jury's consideration. However, these exceptions would be rare: serious injuries are usually detectable by victims, and individual victims of egregious torts can be counted on to bring their own suits.

Why have a rule at all, only to allow exceptions? The advantages of a concrete limit are significant and result in a punitive damages framework that is predictable, which is necessitated by the Constitution. Also, civil liability critics have a point when they argue that massive awards defeat consideration of the merits. If massive awards are threatened, no one can run the risk of defending themselves on the merits. Settlement outside of court becomes the only feasible option, which deprives the defendant of opportunity to defend him or herself.

A mathematical formula like that advocated by the economic theorists has many advantages. However, it remains to be seen whether the economic calculus suggested is implementable in the real world. Although deterrence theory seems simple enough, Cass Sunstein's studies found that most jurors are not attentive to the judge's instructions: individuals averaged 5% correct on a test of memory for comprehension of the instructions (although the studies found that the more discussion a jury devoted to the judge's instructions, the less likely they were to award punitive damages).[51] The outrage and punishment factor is much more appealing to the average juror than predictions about concealment and discovery, especially when enhanced by the rhetoric of an accomplished lawyer. And the formula advocated by Shavell and Polinsky asks the jury to estimate probabilities of detection that can never be known as certain. In addition, the idea of total harm, which makes theoretical sense, is almost impossible to calculate (especially when the jury is barred from using evidence not brought by parties before the court, as was decided in *Philip Morris*). It is also impossible to place accurate monetary values on subjective losses, such as human life, which is almost always undervalued.[52] The only way to determine appropriate punitive damages in these situations is to make a speculative guess about "subjective probability." Due to these difficulties, I find that there is currently not a principled basis for implementing the precepts of deterrence theory. We are left where we started: the judicial system needs clear rules and predictability, but the facts hypothesized by economists are difficult to measure, especially when a jury is making the calculations.[53] Until economists resolve this practical issue, abstract theories involving deterrence will be ignored by the courts in favor of more concrete solutions.

At the present, the courts' best solution is to graft the principles of deterrence theory upon the current trend in Supreme Court jurisprudence. Thus, awards of punitive damages that exceed a 1:1 ratio should be allowed to stand only in cases of a high probability of a lack of detection, or in cases in which actual damages are next to nothing. The jury should not be allowed to speculate on such matters, and it should be the plaintiff's burden of proof on the need for an extra enhancement in the ratio to reflect concealment. But because of these exceptions, the ratio cap would not stand as an absolute rule, but rather as a guidepost in order to limit unpredictable outlier jury awards.

Of course, as the Supreme Court's opinions demonstrate, it is ultimately the job of the judiciary, including the trial judge and all reviewing appellate courts, to make sure that a jury award of punitive damages is justified. The 1:1 ratio for cases involving substantial amounts of actual damages, enlarged only in instances that involve a clear risk of concealment, should be easy for reviewing courts to apply and thereby achieve the Supreme Court's due process goal of promoting certainty and predictability in the punitive damages field. The "outlier" cases (cases involving especially high ratios of punitive to compensatory damages) that have caused the Supreme Court such concern in the past two decades can be dealt with efficiently in this manner.

By capping the permissible ratio, the Court would promote certainty in the law and avoid excessive litigation due to the misunderstanding of the constitutional requirements. For example, there would be fewer cases like *Phillip Morris v. Williams*,[54] that have gone to the Supreme Court several times because lower courts misperceived the Supreme Court's message on the permissibility of punitive damages. The additional predictability and legal efficiency resulting from this ratio cap would benefit society generally, even when discounted by occasional instances of under-deterrence due to reduced legal flexibility. Application of due process law may not achieve optimal deterrence in every case, but parties can predict a consistent level of punishment when their conduct warrants deterrence. Consistency and predictability are benefits not only to the efficient administration of justice in the courts and fairness to those subjected to punishment, but also to settlement negotiations and the formulation of insurance rates. Large economic penalties, imposed unpredictably, prevent parties from arriving at reasonable settlements in civil cases and make liability insurance more costly and sometimes wholly unobtainable.[55]

Endnotes

1 517 U.S. 559 (1996).

2 *Id.* at 575.

3 The case against BMW involved BMW's failure to disclose that the automobile sold to Gore had been repainted after being damaged prior to delivery.

4 *Gore*, 517 U.S. at 586.

5 For example, after outlining the ratio guidepost, Justice Stevens explained,

> Low awards of compensatory damages may properly support a higher ratio than high compensatory awards, if, for example, a particularly egregious act has resulted in only a small amount of economic damages. A higher ratio may also be justified in cases in which the injury is hard to detect or the monetary value of noneconomic harm might have been difficult to determine. It is appropriate, therefore, to reiterate our rejection of a categorical approach.

Id. at 582.

6 *Id.* at 599 (Scalia, J., dissenting).

7 *Id.* at 605.

8 *Id.* at 573 (majority opinion).

9 This difficulty was directly addressed less than ten years later in *Phillip Morris v. Williams,* 549 U.S. 346 (2007), as we will see.

10 But the Court pointed out that a state court class action could not attempt to punish conduct that may be lawful "in other jurisdictions." *Gore*, 517 U.S. at 573.

11 538 U.S. 408 (2003).

12 Once again, Justice Scalia filed a dissenting opinion emphasizing his view that "the Due Process Clause provides no substantive protections against 'excessive' or 'unreasonable' awards of punitive damages," and that "the punitive damages jurisprudence which has sprung forth from *BMW v. Gore* is insusceptible of principled application." *Id.* at 429 (Scalia, J., dissenting). Justice Thomas and Justice Ginsberg wrote concurring opinions.

13 *Id.* at 409 (majority opinion).

14 *Id.* at 425.

15 *Id.* at 427.

16 *See* E. Donald Elliot, "Why Punitive Damages Don't Deter Corporate Conduct Misconduct Effectively" 40 Alabama L. Rev. 1053 (1989); Andrew Frey, "Corporate Wealth: The 800-Pound Gorilla that Sabotages Fair Adjunction of Punitive Damages" *Litigation* (2004); Peter Huber, *Liability: The Legal Revolution and Its Consequences* (1988).

17 549 U.S. 346 (2007).

18 *Id.* at 349.

19 *Id.* at 355.

20 128 S.Ct. 2605 (2008).

21 The Court left open the possibility of a large punitive award when actual damages are very small.

22 *Exxon*, 128 S.Ct. at 2610.

23 *See* Eisenberg & Heise "Juries, Judges, and Punitive Damages: An Empirical Study," 269 (reporting median ratios of 0.62:1 in jury trials and 0.66:1 in bench trials using Bureau of Justice Statistics data from 1992, 1996, and 2001); Vidmar & Rose, "Punitive Damages by Juries in Florida," 38 Harv. J. Legis. 487, 492 (2001) (studying civil cases in Florida state courts between 1989 and 1998 and finding a median of 0.67:1).

24 *Exxon*, 128 S.Ct. at 2611.

25 It is true that maritime law could evolve in a different direction from federal due process law. But the significant aspect of this maritime decision is that it cites and relies on due process case law repeatedly. In a footnote, Justice Souter pointedly comments, "Indeed, any argument for more generous punitive damages would call into question the maritime applicability of the constitutional limit on punitive damages as now understood." *Id.* at 2633. This footnote suggests that the Court intends to rely on these same principles in the due process context.

26 Robert Cooter, "Economic Analysis of Punitive Damages," 56 S. Cal. L. Rev. 79 (1982-83).

27 Robert Cooter, "Punitive Damages for Deterrence: When and How Much?," 40 Alabama L. Rev. 1148 (1989).

28 Mitchell Polinsky & Steven Shavell, "Punitive Damages: An Economic Analysis," 111 Harv. L. Rev. 868 (1997-1998).

29 *Id.* at 873.

30 *Id.* at 875.

31 *Id.*

32 George Priest, "Punitive Damage Reform: The Case of Alabama," 56 La. L. Rev. 825, 831 (1995-1996).

33 492 U.S. 257 (1989).

34 *Id.* at 300-301 (O'Connor, J., dissenting).

35 BMW v. Gore, 517 U.S. 559, 593 (1996) (Breyer, J., concurring).

36 State Farm Mutual Automobile Insurance Company v. Campbell, 538 U.S. 408, 426 (2003).

37 *Id.* at 427.

38 Exxon Shipping Company v. Barker, 128 S.Ct. 2605, 2622 (2008).

39 Polinsky & Shavell, *supra* note 29, at 906.

40 *Gore*, 517 U.S. at 559.

41 *Id.* at 568.

42 Cass R. Sunstein et al., *Punitive Damages: How Juries Decide*, 29 (2002).

43 *Exxon*, 128 S.Ct. at 2611.

44 David Friedman, "An Economic Evaluation of Punitive Damages," 40 Alabama L. Rev. 1125, [page number?] (1990).

45 *Exxon* is a good example of this. In most cases, like *Exxon*, economists would call for no punitive damages at all (there was virtually no chance that Exxon could escape liability), finding that accurately calculated compensatory damages served deterrence purposes appropriately. The cap on punitive damages had the useful function of reducing the award, but in this case where the actual award was substantial, there remained a lot of latitude for unnecessary punishment: the final punitive award was two billion dollars. This, by our economic analysis, could result in unnecessary deterrence (perhaps forcing the company to hire several seamen to steer the ship at the same time, a proverbial unwise strategy). Thus, the risk of over-deterrence still remains even with a multiplier cap on punitive damage awards.

46 *Gore*, 517 U.S. at 598 (Scalia, J., dissenting).

47 This is not to say Justice Scalia explicitly advocates economic theory in his dissent, nor does he support large and unconstrained punitive awards. Later, he paradoxically faults the Court for constructing a framework "that does not genuinely constrain." *Id.* at 606.

48 These exceptions to the general 1:1 ratio rule, of course, are easy to defend on constitutional grounds. They leave more room for the states to impose enhanced punishment, if state policy so provides. Arguably, goals of optimal deterrence and state sovereignty are served by recognizing these narrow exceptions to the bright line limitation.

49 To obtain an enhancement in the damages award based on concealment, the plaintiff need not adduce evidence of other persons who were deceived or secretly injured by the tortious conduct. The Supreme Court's decision in *Phillip Morris v. Williams* makes clear that juries are not entitled to speculate about the merits of cases not before the court. Some civil wrongs, however, are covert by their nature. For example, in *BMW v. Gore*, the hidden refinishing activities could have escaped detection entirely, as shown by the facts of the particular case before the Court. This would warrant an enhanced award based on concealment. Some antitrust violations, such as price fixing conspiracies, are covert and may escape detection,

50 Judge Posner and William Landes have suggested that punitive damages may be expanded in cases that involve actual damages that are hard to calculate. William M. Landes & Richard A. Posner, *The Economic Structure of Tort Law* 161 (1987). That is a rationale sometimes applied in antitrust cases, where treble damages are allowed. But in that context Congress has made the necessary judgment. Judge Posner's proposal does not appear to be a rule that juries could administer systematically. The vagueness of such a standard would also create serious problems for appellate judges, moving the law in the exact opposite direction from that intended by the Supreme Court. Actual damages can be proven based on reasonable estimates, and do not require scientific precision. This gives claimants sufficient latitude to make their "best case" to the jury, aided with expert testimony as appropriate. I would not confuse this field of law further with the exception suggested by Judge Posner.

51 Sunstein et al., *supra* note 42, at 23.

52 William Landes and Judge Posner discuss the undervaluation of human life in wrongful death cases:

> The limitation of damages to survivors' pecuniary loss is very peculiar. It implicitly assumes—if, as we generally believe to be the case, tort law seeks to internalize the costs of accidents—that the average person derives no utility from living. He does not work for himself, he works

> solely for his family. This cannot be right, and it results in a systematic underestimation of damages in wrongful death cases.

Landes & Posner, *supra* note 50, at 161.

53 Although perhaps unfortunate, it appears that a clear ratio cap rule, with a few narrow exceptions, is the only way to prevent runaway awards that violate Due Process. David Friedman addresses the clumsiness of the punitive damages system, noting that "one of the most important factors determining the form of efficient law" is the

> fact that courts are a very poor way of finding the correct answer to a difficult question. If you wish to diagnose an illness, design a computer, or discover a new scientific law, you do not do it by picking a dozen people at random, forming them into a committee, and demanding that they give you an answer. Given the limitations of courts, it is sensible to try to avoid, so far as possible, asking them to do difficult things.

Friedman, *supra* note 44, at 1138.

In his *Gore* dissent, Justice Scalia evaluates the consequences of elevating "fairness in punishment" as an aspect of substantive due process, writing "by today's logic, every dispute as to evidentiary sufficiency in a state civil suit poses a question of constitutional moment, subject to review in this Court." BMW v. Gore, 517 U.S. 559, 606 (1996). It would follow that asking the courts to make calculations outside of their ability would violate due process; the jury almost certainly lacks the means of determining exact probabilities, running the risk of imposing an unfair penalty. The 1-to-1 rule, however imperfect, offers an easily implementable and consistent solution to the problem of excessive "outlier" awards.

54 549 U.S. 346 (2007).

55 *See* Priest, *supra* note 32; George Priest, "The Current Insurance Crisis and Modern Tort Law," 96 Yale L.J. 1521 (1986-1987).

A Wave of ADA Public Accommodation Lawsuits Moves from Florida and California to the Rest of the United States

*by Gregory P. McGuire**

There recently has been a surge of private plaintiff lawsuits filed under the public accommodation provisions of the Americans with Disabilities Act of 1990 ("ADA").[1] Public accommodation lawsuits have been common for several years in states such as Florida and California, but in the past two years have been filed with increasing frequency in several states including North Carolina, Virginia, and Alabama.

Several lawyers and law firms have made a "cottage industry" out of suing places of public accommodation, in part because the law may permit the recovery of attorneys' fees.[2] These lawyers will use the same disabled person, often someone who is wheelchair-bound, as the plaintiff in multiple lawsuits. The individual plaintiff may be joined by, or at least supported by, a disabled "advocacy" group. The lawsuits are being filed primarily against restaurants, hotels, retail stores, and shopping centers. The complaint filed with the court will allege that the plaintiff visited the facility but encountered physical "barriers" to his or her full enjoyment of the available goods and services. For example, the plaintiff may allege that there was inadequate handicapped parking; that access into and out of the building was difficult or impossible because of steep inclines, lack of curb cuts, or narrow doorways; that goods or services inside the facility were difficult to reach because of counter or shelf heights; or that it was difficult or impossible to use the public restroom because of inadequate maneuvering space or inaccessible toilet stalls. The lawsuit typically will seek a court order (injunction) requiring the public accommodation to remove or alter the alleged physical barriers, and ask the court to award other damages and attorneys' fees for the plaintiff's lawyer who brought the lawsuit. One Florida-based plaintiff has filed 42 separate ADA public accommodation lawsuits in North Carolina's three federal district courts since July 2008, and he has been represented by the same attorney in 39 of those lawsuits. Another Pennsylvania-based lawyer, using the same plaintiff, filed 8 ADA public accommodation lawsuits in North Carolina's Federal District Courts for the Middle and Western Districts between September 22 and October 8, 2009.[3]

This article is intended to provide a general overview of the requirements placed upon places of public accommodation (and "commercial facilities") by Title III of the ADA to make facilities physically accessible by the disabled, the most commonly raised violations of the accessibility standards, and some practical pointers for avoiding and defending public accommodation lawsuits.

The ADA was signed into law on July 26, 1990[4] and is divided into three sections, or "Titles": Title I covers discrimination in employment; Title II prohibits discrimination by state and local "public" entities in providing access to programs, services and activities; and Title III prohibits discrimination by privately owned places of "public accommodation" and prohibits certain types of discrimination by "commercial facilities."

Title III of the ADA provides that "no individual shall be discriminated against on the basis of disability[5] in the full and equal enjoyment of the goods, services, facilities, privileges, advantages, or accommodations of any place of public accommodation by any person who owns, leases (or leases to), or operates a place of public accommodation."[6] The discrimination prohibited by Title III clearly includes overt discrimination against a disabled individual by a public accommodation, such as refusing entry to the facility or service to that individual. However, Title III also specifically defines "discrimination" to include (1) the failure to "design and construct" new facilities so that they are "readily accessible to and usable by individuals with disabilities,"[7] (2) where alterations are made to a facility, the failure "to make alterations in such a manner that, to the maximum extent feasible, the altered portions of the facility are readily accessible to and usable by individuals with disabilities, including individuals who use wheelchairs,"[8] and (3) in "existing facilities" the failure to remove architectural (physical) barriers to access for the disabled, and the failure to provide appropriate alternative services to the disabled when physical barriers cannot be removed.[9] As explained below, which of the accessibility standards a public accommodation is required to meet will depend upon the date that the facility was constructed and first occupied, or when and what types of alterations are made to a facility.

Under the ADA, a place of "public accommodation" is a facility or establishment owned or operated by a private entity that is open to the public.[10] Public accommodations include, but are by no means limited to:

- hotels, restaurants, retail stores and shopping centers;
- service establishments like banks, doctor's offices, and the portions of other business offices open to the public;
- recreational facilities, including health clubs and spas;
- theaters, auditoriums, and stadiums;
- private places of education and day care centers; and
- many other establishments and facilities open to the public.

A "commercial facility" is different from a public accommodation. Title III of the ADA defines a "commercial facility" as a facility "intended for nonresidential use" and "whose operations will affect commerce," but which is not open to the general public.[11] Commercial facilities include certain office buildings, factories, warehouses and distribution centers, and other buildings in which employment might occur, as well as wholesale establishments that sell exclusively to other businesses and private airports.

* *Shareholder, Ogletree Deakins Nash Smoak & Stewart, P.C., Raleigh, North Carolina*

It also should be noted that the ADA's public accommodation provisions cover more than just the physical attributes of the facility such as parking lots, doorways, access routes, and restrooms. It also requires the removal of "barriers" to the use and enjoyment of the facility and the services offered by the facility. For example, a restaurant may be required to provide Braille menus for blind customers, and a bank must provide reasonable means for a deaf individual to transact business.[12]

Whether a facility is considered accessible is determined by reference to the "ADA Accessibility Guidelines for Buildings and Facilities" adopted by the United States Department of Justice ("the ADAAG").[13] The ADAAG provides detailed requirements and architectural standards for virtually all aspects of a public accommodation, including parking lots, access and entrances to the facility, paths of travel within the facility, restrooms, drinking fountains, access to goods and services in the facility, and safety features. A reputable architect or engineer should be familiar with the requirements for compliance with the accessibility ADA. In addition, there are architectural and engineering firms that have special expertise in ADA accessibility issues, and provide design, consulting, and inspection services.

As noted above, the non-discrimination requirements of Title III apply to all places of "public accommodation." However, the requirement that a public accommodation "design and construct" accessible facilities or remove "architectural barriers to accessibility," will depend on when a facility was built, and whether the facility has been renovated since the ADA became effective.

I. New Construction

As used in the ADA, "new construction" refers to buildings which were constructed for first occupancy on or after January 26, 1993.[14] Such "new" buildings must be constructed so that they are "readily accessible to and usable by individuals with disabilities."[15] The only exception to this requirement is where the public accommodation can demonstrate that it is "structurally impractical" to meet the ADAAG requirements because "the unique characteristics of the terrain prevent the incorporation of the accessibility features."[16] In other words, a public accommodation first occupied after January 23, 1993 usually is required to come into compliance with the ADAAGs. The new construction requirements of the ADA apply both to public accommodations and to commercial facilities.

An interesting issue that remains unclear at this time is whether an entity that buys a facility that was first occupied after January 23, 1993, but which was not built in compliance with the ADA "new construction" standards, is required to bring the facility into compliance. There are at least some court decisions that suggest that a purchaser who had no involvement in the "design and construct[ion]" of the facility would not be required to come into compliance, but would be subject to the "existing facility" standards described below.[17]

II. Alterations (Renovations) to Facilities

If an existing building is "altered" or renovated after January 26, 1992, the owner or operator must "to the maximum extent feasible" remove architectural barriers and make the altered portion of the facility accessible to disabled individuals in compliance with the ADAAGs.[18] The alterations requirements of the ADA apply both to public accommodations and to commercial facilities.

An "alteration" to a facility "is a change to public accommodation or commercial facility that affects or could affect the usability of the building or facility or any part thereof."[19] The regulations further provide that when such alterations affect the usability or access to an area of the public accommodation containing its "primary function," the alterations

> shall be made so as to ensure that, to the maximum extent feasible, the path of travel to the altered area and the restrooms, telephones, and drinking fountains serving the altered area, are readily accessible to and usable by individuals with disabilities, including individuals who use wheelchairs, unless the cost and scope of such alterations is disproportionate to the cost of the overall alteration.[20]

The "primary function" is the major activity for which the facility is used, such as the lobby and customer service areas of a bank, the dining and bar areas in a restaurant, and the common areas and guest rooms in a hotel.[21] The ADA regulations provide the following examples of alterations that would affect the "primary function" of a public accommodation or commercial facility: "(i) Remodeling merchandise display areas or employee work areas in a department store; (ii) Replacing an inaccessible floor surface in the customer service or employee work areas of a bank; (iii) Redesigning the assembly line area of a factory; or (iv) Installing a computer center in an accounting firm."[22]

There are detailed regulations regarding the extent of the duty to remove "path of travel" barriers and make the facility accessible depending on the nature of the alterations, the areas of the facility being altered, and the cost of making the changes.

III. "Existing" Facilities

For public accommodations that existed as of January 26, 1990, and were first occupied before January 26, 1993, the ADA provides that the failure to remove "architectural barriers" where the removal is "readily achievable" is a violation of the ADA. The failure to remove existing barriers from existing facilities applies only to "public accommodations" and not to "commercial facilities." In other words, the owners of an office building used primarily for private employment, or of a manufacturing facility, have no obligation to remove existing barriers from the non-public areas of the building.

The ADA defines "readily achievable" as "easily accomplishable and able to be carried out without much difficulty or expense."[23]

The ADA provides that

> [i]n determining whether an action is readily achievable, factors to be considered include:
>
> (A) the nature and cost of the action needed under this chapter;
>
> (B) the overall financial resources of the facility or facilities involved in the action; the number of persons employed at such facility; the effect on expenses and resources, or

the impact otherwise of such action upon the operation of the facility;

(C) the overall financial resources of the covered entity; the overall size of the business of a covered entity with respect to the number of its employees; the number, type, and location of its facilities; and

(D) the type of operation or operations of the covered entity, including the composition, structure, and functions of the workforce of such entity; the geographic separateness, administrative or fiscal relationship of the facility or facilities in question to the covered entity.[24]

Neither the courts nor the Department of Justice have attempted to state any mathematical formula for determining whether the cost of removal of a barrier is "readily achievable." Instead, it is well established that such decisions must be made on a case-by-case basis. Obviously, what might be cost prohibitive for a single location retail store might not be for a location of a national retail chain.

In addition, the United States Supreme Court has held that whether barrier removal is readily achievable "extends to considerations beyond cost."[25] In fact, courts have held that the fact that the public accommodation can afford to remove the barrier does not necessarily mean that the removal is readily achievable.[26] One commentator has stated that:

> In determining whether the cost of barrier removal is "readily achievable," a public accommodation can consider the effect of the barrier removal on the operation of its business in addition to the initial cost of simply removing the barrier. In certain circumstances, the actual cost of removing a physical barrier may not be very large; however, the process of removing the barrier may cause additional costs by disrupting the operation of the public accommodation.[27]

IV. The Recent Public Accommodation Lawsuits

Although Title III of the ADA covers, and the ADAAGs provide access standards for, a multitude of physical facilities and services provided by public accommodations, the recent slew of lawsuits tend to focus on the same physical, or "architectural," barriers: parking lots and spaces, entrances to the facility, and restrooms.

A. Parking Spaces and Access From Parking Lot

Almost all of the recent public accommodation lawsuits filed against retailers, restaurants, and shopping centers allege that the disabled parking spaces are inadequate. The ADAAG requires that a place of public accommodation provide a certain minimum number of handicapped parking spaces and dictates the size, position of, and access from such parking spaces. For example, a public accommodation that has 51-75 parking spaces must have at least 3 handicapped accessible parking spaces, and a parking lot with 76-100 spaces must have 4 accessible parking spaces.[28] The handicapped accessible parking spaces must be at least 96 inches wide and must have a "parking access aisle" next to the space of at least 60 inches width (although 2 accessible spaces can share the same access aisle).[29] The parking spaces and access aisles must be "level, with surface slopes not exceeding 1:50 (2%)."[30] The accessible spaces must be located "on the shortest possible route of travel from adjacent parking to an accessible entrance."[31]

The ADAAG also provides that "one of every eight accessible spaces, but not less than one" shall be "van accessible." A "van accessible" parking spot must have next to it an "access aisle" at least 96 inches wide.[32]

The accessible parking spaces must have an accessible route to the entrance of the building at least 36 inches wide.[33]

Parking space problems tend to be the easiest and least expensive to fix, particularly when the problems relate solely to the number and size of the spaces, which often can be remedied merely by restriping the parking lot. However, when the problems involve the slope of the parking spaces or access lane, or a lack of curb cuts, the remedies can become much more expensive.

B. Entrances and Doors

The ADAAG requires that a public accommodation provide at least one accessible entrance doorway that is a minimum of 32 inches in width.[34] The doorway also must provide certain minimum clear space around the door to permit maneuvering of a wheelchair, and the "floor or ground area within the required clearances shall be level and clear."[35] The exact size and space requirements vary depending on the type of doors involved. The door on this entranceway must have a handle that is "easy to grasp with one hand and does not require tight grasping, tight pinching, or twisting of the wrist to operate."[36]

C. Restrooms

Most of the current wave of public accommodation lawsuits also allege significant deficiencies with accessibility to restrooms and toilet facilities. Renovation of restroom facilities can be expensive, disruptive, and structurally challenging.

The ADAAG provides that a public accommodation must have at least one handicapped accessible stall. A "standard" accessible stall must be at least 56 inches in length, and 60 inches wide.[37] However, when it is "infeasible" to provide a stall width of 60 inches, the ADAAG permits alternative configurations with a 36-inch width or a 48-inch width.[38] There are specific requirements for the placement of handrails depending on the configuration of the stall.[39]

Door handles on stall doors should have a "shape that is easy to grasp with one hand and does not require tight grasping, tight pinching, or twisting of the wrist to operate." The Guidelines recommend the use of "lever-operated mechanisms."[40]

Urinals must be positioned so that the front rim is no higher than 17 inches off the ground, and must have a minimum clear floor space in front of the urinal of at least 30 inches by 48 inches.[41]There must also be minimum amounts of floor "maneuvering" space provided inside the rest room, and the ADAAG also provides specific requirements for the placement of the sinks and mirrors in the restrooms.[42] A handicapped accessible restroom must have a handicapped accessible sign on the door.

V. Enforcement of Title III

The authority to enforce the nondiscrimination and accessibility requirements of the ADA are placed in the hands of both the United States Department of Justice and private individuals.[43] The United States Attorney General is authorized to "investigate alleged violations of [Title III], and shall undertake periodic reviews of compliance of covered entities." The Attorney General also may bring a lawsuit where there is a "pattern or practice of discrimination," or when discrimination "raises issues of general public importance."[44] The Attorney General may seek injunctive relief, damages for impacted persons who are discriminated against, and civil monetary penalties for non-compliance of $50,000 for a first violation, and $100,000 for subsequent violations.[45] The Attorney General cannot seek an award of punitive damages.[46]

A private individual who has been the victim of discrimination may also bring a civil lawsuit under Title III of the ADA.[47] Generally, the individual must show that he or she was discriminated against by being denied access to or enjoyment of the public accommodation, and that he or she has a desire to return and use the facility in the future. A private plaintiff may seek an injunction requiring that barriers to access be removed and an award of attorneys' fees if they win the lawsuit.[48] A private plaintiff cannot be awarded compensatory or punitive damages under Title III of the ADA.

VI. Avoiding and Defending Public Accommodation Lawsuits

The best means of avoiding a public accommodation lawsuit is to make certain that a facility is in compliance with the ADAAG standards to the maximum extent feasible. Obviously, the best time to do this is during the construction, or renovation, of a facility, when the plans are being drawn and the construction is taking place. Although most good architects are very familiar with the ADA's accessibility standards, there are a surprising number of lawsuits that arise out of the failure to construct or renovate facilities in compliance with the ADAAG, including lawsuits involving buildings constructed and opened in the last 5 to 10 years. A company or individual engaged in designing and building a facility to be used as a public accommodation or a commercial facility may want to hire an ADA accessibility expert to independently review architectural plans and confirm compliance with the ADA requirements. There are a number of engineering firms and other qualified experts who are in the business of conducting ADA compliance reviews and are expert in the accessibility requirements. Another potential means of protection for new construction or renovations is to negotiate an indemnity provision that specifically makes the architects and/or construction contractor liable to indemnify the facility's owner if the building is not constructed in compliance with the ADA requirements.

For owners and operators of public accommodations in "existing facilities," an ADA accessibility study or audit by the aforementioned accessibility expert is the best means of discovering and remedying accessibility problems before being hit with a lawsuit. Renovations of parking lots, restrooms, entranceways, and other parts of a facility are much less costly and easier to manage when they are done on the owner or operator's own schedule than when they are done on the forced schedule that comes with a court-ordered injunction. It may also be wise to have the audit or inspection conducted with the assistance, and under the direction of, an attorney so that the report may be protected from discovery in a lawsuit. This is particularly important where the owner or operator decides that it will fix some, but not all, of the accessibility problems discovered by an inspection. If the inspection report is not prepared at the direction of counsel, it may have to be provided to a plaintiff who brings a later public accommodation lawsuit.

If an owner or operator of a public accommodation is sued under Title III of the ADA, it should immediately report the lawsuit to its insurance carriers for potential coverage. The owner or operator, or its insurance carrier, will also want to quickly retain competent legal counsel to represent it in the lawsuit. Among the first things that must be determined are: was the facility constructed for first occupancy after January 26, 1993?; have there been substantial renovations made to the facility since January 26, 1992?; are there any indemnity provisions regarding ADA compliance contained in contracts surrounding construction or renovation of the facility?; and are the allegations in the lawsuit accurate—is the facility out of compliance with the ADA accessibility requirements? In addition, for an "existing facility," it will be important to determine how much it will cost to make any needed renovations to the facility so that it can be determined if such renovations are "readily achievable."

The wave of ADA public accommodations lawsuits does not show signs of breaking anytime soon. Preparation before they hit is the best protection.

Endnotes

1 42 U.S.C. §12182 et seq. Statistics compiled by United States federal court system showed a 56% increase in the number of non-employment lawsuits filed under the ADA from 2005 to 2008, and a 24% increase from 2007 to 2008.

2 The United States Department of Justice also is authorized to enforce the public accommodation requirements of the ADA and to bring a civil lawsuit where it believes there is a "pattern or practice" of violations, and it may seek damages and civil monetary penalties for non-compliance.

3 Information compiled from search of federal district court's electronic PACER dockets.

4 42 U.S.C. §12101 et seq.

5 Title III uses the same definition of "disability" as is used in Title I prohibiting employment discrimination. 42 U.S.C. §12102 (2).

6 42 U.S.C. §12182(a).

7 42 U.S.C. §12183(a)(1).

8 42 U.S.C. §12183(a)(2).

9 42 U.S.C. §12182(b)(2)(A)(iv) and (v).

10 42 U.S.C. §12181(7). Title II of the ADA places the same or similar requirements on state and local publicly owned and operated facilities open to and utilized by the public. 42 U.S.C. §12181 et seq. Federal government facilities are not covered by the ADA, but are covered by the Rehabilitation Act of 1973.

11 42 U.S.C. §12181(2).

12 42 U.S.C. §12182(b)(2)(ii); 28 C.F.R. Part 36, §36.303.

13 Appendix A to the ADA regulations at 28 C.F.R. Part 36. A revised version of the ADAAG, first introduced in 2004, was submitted to the Office of Management Budget in late 2008 by the Bush Administration as part of a proposed revision of the Title III regulations, but was not acted upon before the change of administration.

14 42 U.S.C. §12183(1).

15 *Id.*

16 *Id.*

17 *See, e.g.,* Rodriguez v. Inestco, L.L.C., 305 F. Supp.2d 1278 (M.D. Fla. 2004).

18 42 U.S.C. §12183(a)(2); 28 C.F.R. Part 36, §36.402.

19 28 C.F.R. Part 36, §36.402(b).

20 28 C.F.R. Part 36, §36.403.

21 *Id.*

22 *Id.* §36.403(c).

23 42 U.S.C. §12181(9).

24 *Id.*

25 Spector v. Norwegian Cruise Lines, Ltd., 545 U.S. 119 (2005).

26 Brother v. CPL Investments, Inc., 317 F.Supp.2d 1358, 1371 – 72 (S.D. Fla. 2004).

27 Mook, Americans with Disabilities Act: Public Accommodation and Commercial Facilities, Ch. 3, § 3.02(4)(a) (Matthew Bender).

28 ADAAG 4.1.2(5).

29 *Id.* 4.1.2.

30 ADAAG 4.6.3.

31 *Id.* 4.6.2.

32 ADAAG 4.1.2(5)(6).

33 ADAAG 4.3.3.

34 ADAAG 4.13.5.

35 *Id.* 4.13.6.

36 ADAAG 4.13.9.

37 ADAAG 4.17.3.

38 *Id.*

39 ADAAG 4.26.

40 ADAAG 4.13.9.

41 *Id.* 4.18.2 and 3.

42 ADAAG 4.19.

43 42 U.S.C. §12188.

44 42 U.S.C. §12188(b).

45 42 U.S.C. §12188(b)(2).

46 *Id.*

47 42 U.S.C. §12188(a)(1).

48 *Id.*; 42 U.S.C. §2000a-3(a); 42 U.S.C. §12205.

Professional Responsibility & Legal Education

Government Ethics in President Obama's First Year: A Preliminary Assessment

*by Richard W. Painter**

Barack Obama campaigned on a promise to bring ethics to Washington. On January 21, 2009, one day after becoming President, he delivered on this promise by signing an Executive Order that purported to slam shut the revolving door between the private sector and his Administration. He made clear his expectation that members of his Administration are to serve the public free of whatever conflicts of interest might get in the way.

Will government ethics really change for the better under President Obama, or will things get worse? Will it be more of the same? It is too early to tell, but the signs we have seen thus far demonstrate how difficult it will be for the President to make headway against imbedded conflicts of interest in Washington. The President's chances of success in this area will depend on the priority he gives to government ethics over his other policy and political objectives. It is not at all clear where government ethics stands on his priority list relative to the many other promises he made to the people who elected him.

This essay discusses a sampling of important issues in government ethics and an assessment of what the Administration has accomplished in each. A more complete assessment will require more space than this essay will allow and also will require more time to observe what the President is able to accomplish.

A first priority at the beginning of an Administration is appointing senior officials who have demonstrated integrity and sound judgment. Some Presidents get off to a bad start by indiscriminately relying on friends from their home state for senior appointments (recall Bert Lance in the Carter Administration and many others since then). For a President coming from Chicago, indiscriminate home-state appointments could have been an ethical disaster (in the past year the President's Senate seat was added to the long list of things politicians put up for sale in Illinois). Although the President brought some people from Illinois to Washington, fortunately he avoided the more sordid elements in state politics and turned instead to persons who like himself keep a respectable distance from corruption (former Illinois Congressman Rahm Emanuel as White House Chief of Staff, Chicago schools chief Arne Duncan as Education Secretary, and Valerie Jarrett as White House director of Intergovernmental Affairs are a few of his better home-state picks). Surprisingly, most of the vetting problems in this Administration have not been with people the President brought from Illinois.

** S. Walter Richey Professor of Corporate Law, University of Minnesota, and former Associate Counsel to the President and chief White House ethics lawyer 2005-07. In January 2009, Professor Painter published his book Getting the Government America Deserves: How Ethics Reform Can Make a Difference (Oxford U. Press).*

There have, however, been problems with the vetting process for senior appointments, particularly in the first few months of the Administration. This process—sometimes called a "sex, drugs and rock and roll" review—involves an FBI background check before an appointment is announced and is supposed to ferret out candidates who don't pay their income taxes, don't pay their nanny taxes or hire illegal immigrants, have prior criminal convictions or adverse civil judgments, have problematic corporate or charitable board memberships, have difficulty living in monogamous relationships, or have anything else in their background that could reflect badly on the President. This process—which infamously failed the Bush White House in the Bernie Kerik nomination but rarely since then—appeared to be broken for the Obama team in the months after the 2008 election. Vetting failed to detect problems with at least three cabinet level nominations: the successful nomination of Treasury Secretary Timothy Geithner, who failed to pay some of his taxes; the failed nomination for Heath and Human Services Secretary of former Senator Thomas Daschle, who did not pay some of his taxes and who had conflicts of interest within the private sector; and the failed nomination for Commerce Secretary of New Mexico Governor Bill Richardson, who was in the midst of an uncomfortable criminal investigation into awards of state contracts to campaign contributors. There were problems with vetting some lower level appointments as well.

One of the President's key appointments has not received much attention as a vetting problem, although the facts suggest it was a problem. The appointment lies at the heart of a dilemma that the President inherited but that, if mismanaged, could damage his Presidency: Afghanistan.

Richard Holbrooke is a talented if controversial diplomat with a track record in Kosovo, and he brings this experience to his present position as liaison between the United States and parties interested in the War in Afghanistan. Holbrooke was also, however, a director of AIG between 2001 and 2008, was on AIG's compensation committee, which handed out millions in bonuses to executives who brought AIG to disaster, and resigned from AIG in the summer of 2008, just as things were falling apart. The badly managed AIG has since become an enormous money pit for federal tax dollars as a cornerstone of the bailout of the financial services industry. President Obama has observed that "[n]obody here was responsible for supervising AIG and allowing themselves to put the economy at risk by some of the outrageous behavior that they were engaged in." The President presumably meant to say that the persons who were responsible were at AIG, and those persons did not do their jobs.

Earlier, Holbrooke also had trouble with a core statute regulating the revolving door between government and the private sector. He left the Clinton Administration to pursue investment banking. The Department of Justice subsequently

alleged that he violated post-employment rules in a criminal statute by representing back to the State Department on behalf of an investment bank when he was prohibited from doing so, charges which were later settled with payment of a $5000 fine.[1]

There is reason to wonder whether a man who could not identify conflicts of interest of his own as a former government employee and then failed again to identify conflicts of interest and business risks at AIG can effectively deal with a geographic region riddled with government corruption, not to mention Al Qaeda and the Taliban. Holbrooke is among those pushing for a larger military commitment in Afghanistan, but is he oblivious to the risk that the region could become a bottomless pit for American money and human lives? Will Holbrooke's tough talk with the Afghan government be enough to turn the situation around, or is it as irretrievable as our alliance in the 1960's with the corrupt government of South Vietnam? Regardless of how one comes down on the merits on these difficult questions, persons of differing views should agree that we want to have confidence in the good judgment and oversight capability of people in charge of our diplomatic and military efforts in Afghanistan. If Holbrooke cannot show better judgment this time around than he has in the past, he should step aside and allow someone else to do the job.

Background investigations for incoming officials may not be the Obama Administration's strong suit, but a related area in which the President has had more success is limiting conflicts of interest that incoming officials bring from the private sector. It is in this area that the President's Executive Order of January 21 tightened up the rules.[2] Among other things, the Order requires incoming Administration appointees[3] to sign a pledge that, for two years, they won't work on particular matters involving specific parties, including regulations and contracts that are "directly and substantially" related to their former clients or employers.[4] The Order imposes even stricter rules on incoming appointees who are registered lobbyists.[5] The Order recognizes that the revolving door into government is a serious problem and at least attempts to deal with it.[6]

There could, however, be problems with implementation of the Order. So many senior government officials come in from the private sector that this is a difficult area to regulate. If restrictions are too onerous, people from the private sector will not agree to serve. Indeed there is already controversy over how many waivers from the Executive Order will be granted, as well as over whether agency lawyers will interpret the Order narrowly to require recusals from some matters but not others.[7] If too many waivers are granted or the Order is interpreted too narrowly, its purpose will be compromised.

The President's Order also addresses the revolving door out of government and the excessive influence former government officials can exert on their agencies. For senior Administration officials, the Order lengthens the post-employment ban on "representing back" to their former agencies from one year to two years.[8] Administration appointees who leave to become lobbyists are required to promise not to lobby other Administration appointees for the remainder of the Administration.[9]

There are several difficulties with this approach. First, a pledge of this sort is difficult to enforce vis-à-vis former Administration officials after they leave the government. It lacks the teeth of the existing law (a one-year ban for senior officials, a two-year ban for very senior officials, and no additional restrictions for lobbyists) in 18 U.S.C. 207, which, although narrower in scope, is a criminal statute rather than a pledge. Second, if violations of the criminal statute 18 U.S.C. 207, which Richard Holbrooke was charged with violating, are not prosecuted as vigorously as they should be and are not considered impediments to future government appointments, it is difficult to envision government officials taking the pledge in the President's Order seriously. The President should have urged the Justice Department to step up enforcement of the existing law and should have categorically barred persons who violated the existing law from serving in his Administration. Third, the pledge will be meaningless if the President releases his appointees from the pledge by rescinding or amending the order at the end of his Administration, which is what President Clinton did with another similar order at the end of his administration. The President should make it clear that this will not happen, that the rule he announces now will remain the rule when his Administration draws to a close and his appointees seek opportunities outside the government. Persons who violate the pledge should not be welcomed back into any future administration.

The President deserves credit for taking unprecedented steps in the Executive Order of January 21, 2009 to limit lobbyists' influence on government and to address the more problematic aspects of the revolving door from the private sector in and out of government. It remains to be seen whether the President can stay the course or whether exceptions will swallow the rules. It also remains to be seen whether, despite the stricter rules the President has imposed, he can attract to the federal government people with private sector expertise whom the government needs.

The revolving door furthermore is not the only means by which lobbyists and other private sector interests influence government decisions. Partisan politics and campaign contributions are an even bigger factor. From this perspective, it is troubling that the President has retained the White House Office of Political Affairs (OPA). OPA was for much of the George W. Bush Administration run by Karl Rove. Senator John McCain promised in the 2008 presidential campaign to abolish the OPA and move most of its functions over to the Republican National Committee. The issue, however, received little attention and Senator Obama was not forced to match or even address this campaign promise. Under President Obama, OPA has been taken over by Patrick Gaspard, a labor union advisor from New York.

Political advisors have a long history in the White House. Beginning in the Reagan Administration, they worked within a separate OPA with its own head. A number of factors, including the so-called "permanent campaign" that began in the Clinton years and lasts all four years of a President's term, demand for campaign contributions, and the enhanced role of lobbyists and interest groups in elections, have drawn OPA into purely partisan politics not only for the President's reelection but for members of Congress.

The work of OPA staff members is twofold. First, they advise the President on the political viability of Administration policies. Second, on "personal" time, they moonlight for the President's political party—among other things, speaking at campaign events, coordinating strategy with candidates, and facilitating political work by other Administration officials.

The theory behind this dual role is that a beneficial synergism will result. Political work is not part of the official duties of White House staff members, but it puts them in contact with candidates, grass roots political organizations, and pollsters. Presumably, knowledge gained thereby informs official-capacity political advice to the President.

There are, however, ethical and practical problems with this arrangement.

The first problem is legality. The Hatch Act prohibits government officials from engaging in political activity using official titles or at government expense. Most government officials may not participate in political activity while on government property or during working hours. An exception, however, allows senior political appointees to do so provided they do not use their official titles or incur additional expense for the government.

This exception permits some people to do both official and political work in the same office, provided they purport to distinguish between the two. Numerous gadgets—BlackBerries, cell phones, computers, etc.—are thus provided by the RNC or the DNC to OPA staff and some other Administration officials. Modern technology makes it easier than it once was to coordinate with political campaigns. Calls coming from White House officials on DNC cell phones and emails sent on DNC BlackBerries are, legally, not coming from the White House at all. They are merely "personal capacity" communications by persons who happen to be White House staff.

These distinctions are more theoretical than real. In most Administrations OPA staff members use the same internal reporting structure to coordinate political activity that they use for official duties. When they make phone calls or send email, everyone knows where they work. When they speak at campaign events, everyone knows who they are. Calling partisan political activity by White House staff "personal" rather than "official" is a legal fiction.

The second problem is conflict of commitment. There is no way of knowing how much time is spent on politics instead of official duties because time records for senior political employees are not required. Presumably, records of reimbursements they receive from campaigns for travel expenses are filed with the FEC, but this information is difficult for the public to obtain. Little is known, for example, about how many trips are taken by OPA staff and who pays for them.

The third problem is conflict of interest. Many contacts made in partisan politics are with fundraisers and donors. The Hatch Act allows government employees to speak at fundraisers provided they do not explicitly ask for money (another legal distinction with little grounding in reality). White House staff and other Administration officials are highly sought-after speakers because they fill the room with paying customers.

These customers usually want something in return. Lobbyists are among the most frequent attendees (some fundraisers are hosted by lobbyists). Government officials learn at these events what contributors want. Official-capacity advice based on these views reflects a well-heeled segment of the President's political party, but does not necessarily encapsulate what is best for the country or even what is politically viable.

Concurrent political and official roles thus put government officials in an untenable position. Critics often blame OPA staff members for the resulting problems and claim things would be better if another political party controlled the White House. These problems, however, are inevitable.

Retaining the White House OPA can work for the Obama Administration, but an ethical quagmire will be inevitable unless the role of OPA changes. OPA staff, along with other White House staff and senior Administration officials, should not personally participate in partisan politics. The President should be assisted by a staff with undivided loyalties to the government and not beholden to the supporters of a political party.

One key area where the Office of Political Affairs is often involved, and sometimes clumsily involved, is personnel decisions. On the whole, this White House has received relatively little criticism for the type of blatant politicization of hiring and firing decisions that characterized the early days of the Clinton Administration, when Republican U.S. attorneys were fired en masse and even the White House Travel Office was a vehicle for political patronage (the most noted controversy over politicized firings in the George W. Bush Administration occurred in the second term with the U.S. attorney firings). There has been, however, at least one glaring exception—the firing by President Obama of a Republican inspector general in a manner that showed insensitivity to the Inspector General Reform Act of 2008, a statute that Senator Obama had sponsored in Congress in order to depoliticize hiring and firing of inspectors general. The fact that the firing of this Inspector General was delegated to, of all people, the chief White House ethics lawyer, made the episode even more discomforting.

Gerald Walpin had been appointed by President Bush, and confirmed by a Democratic-controlled Senate, to be the Inspector General for the Corporation for National and Community Service, which oversees AmeriCorps. In 2009, Walpin was a holdover from the Bush Administration and was widely known to be a conservative Republican. The Board of the Corporation complained to the White House in May 2009 that Walpin was ineffective.[10] Walpin had also recently completed an investigation of St. Hope Academy, a nonprofit founded by Kevin Johnson, now Mayor of Sacramento and a political ally of the President. Walpin had referred Johnson for prosecution; the Acting United States Attorney declined to prosecute and instead settled the case with St. Hope Academy, which was required to return hundreds of thousands of dollars to Americorps. The United States Attorney then wrote a letter to the Council of the Inspectors General on Integrity and Efficiency strongly criticizing Walpin for being one-sided and overzealous in the investigation. The Council, however, never got a chance to complete an investigation of the issues raised in the Acting U.S. Attorney's letter (the Council was established under the 2008 Reform Act, and its purpose includes investigating allegations against an IG and recommending appropriate action). The President fired Walpin first.

Norman Eisen, the chief White House ethics lawyer, called Walpin on June 10, 2009 and told him to resign within one hour or be fired. Walpin did not resign and was fired later that same day with 30 days paid leave prior to termination. His access to his office and to government email was cut off immediately. President Obama then sent a brief letter to Congress stating that Walpin had been fired because the President no longer had "confidence" in him. Members of Congress from both parties said this was an insufficient explanation and clamored for the meaningful report of the reasons for the firing contemplated by the 2008 Reform Act (the Act requires the President to report to both houses of Congress the reasons for firing or transferring an IG at least 30 days *before* a firing or transfer of an inspector general). A few days later, Eisen wrote a letter to three individual Senators reciting the criticism of Walpin's conduct by the Acting U.S. Attorney and also stating that Walpin had appeared "disoriented" and "confused" at a recent Americorps board meeting. Eisen also met with congressional staff persons to explain the firing. Some members of Congress were not satisfied with any of these explanations and wondered how Walpin could be both overzealous and one-sided and yet "disoriented" and "confused" (Walpin is 77 years old, but by all accounts from colleagues in and out of government, including former White House Counsel Bernard Nussbaum, he is still very sharp[11]). The fact that he was transferred to 30 days administrative leave rather than allowed at least 30 days to wrap up his responsibilities also appeared to be, at best, an effort to technically comply with the Reform Act while avoiding its intent.

Regardless of the merit of this firing, or lack thereof, the White House appeared oblivious to the fact that Congress had enacted a law specifically designed to avoid unexplained firings of inspectors general and that Congress had in that same statute demanded a meaningful opportunity to discuss with the President the prospective firing of an inspector general before the President made his final decision. Congress had also intended the Council of the Inspectors General, not the White House, to address in the first instance allegations of misconduct by inspectors general.

Furthermore, firing people does not belong in the White House ethics lawyer's job portfolio. The ethics lawyer's job is advising the President and White House staff on ethics, not making or implementing policy or personnel decisions, particularly decisions that appear to have a strong political component. Indeed, the ethics lawyer's job is to put the brakes on when a proposed White House action raises the appearance of impropriety, for example, by violating the spirit if not the letter of an act of Congress. When someone has to tell the political people in the White House that they cannot do something they want to do, that person is often the White House ethics lawyer. Much of this advice is private, and when the political people decide to do something that is arguably inappropriate, they should not ask the White House ethics lawyer to do it for them.

Here, the best advice would have been for the President to stay as far away from the Walpin situation as possible, until the Council finished its investigation and made a recommendation. Only in the most urgent of circumstances would the White House want to take immediate action. An allegedly overzealous—or alternatively "disoriented"—IG at Americorps doesn't come close. Also, if someone were going to take the ill-advised step of calling Walpin and pressuring him to resign, and then explaining this step to upset members of Congress, that person should not be the White House ethics lawyer.

Turning to broader concerns, if there is one aspect of the President's policies that is worrisome from a government ethics perspective, it is his acceleration of a trajectory already set by his predecessors toward expanding the size of government and the scope of government's responsibilities. Presidents Clinton and Bush did much the same, although there was sometimes talk of making government smaller, more responsive, and more efficient. More money is passing through the hands of government than ever before, and government is trying to solve problems in areas as diverse as homeland security, health care, bailouts of failing companies, and military support for struggling foreign governments. In some of these areas, government engagement and expenditure is needed, and in others not, policy issues that will not be discussed here. Regardless, expansion of government, particularly rapid expansion of government into new areas of engagement without sufficient attention to conflicts of interest and other ethics issues, can, and already has, come at the expense of government integrity.

As the United States most recently learned in Iraq, wars pose enormous risk to the ethics of government officials. Billions of dollars are spent, and conflicts of interest and other problems plague relationships between the United States government and its own civilian and military employees. Part of the problem is the number of outside entities we rely upon to do jobs we cannot do or don't want to do to achieve military and political objectives, including private companies such as Halliburton and Blackwater, nongovernmental organizations (NGOs), and foreign governments that purport to be our allies. This is nothing new. The American Revolution,[12] the Civil War, World War II, and just about every other war saw not only a rise in patriotism but private profiteering by persons eager for a share of the money the government spent on those wars. President Obama's most immediate engagement is Afghanistan, but that conflict could easily spill over into other countries in the region. The United States also is not finished in Iraq. Iran, Korea, a growing number of terrorism cells in Africa, and instability in Southeast Asia are also concerns. If the United States addresses these concerns unilaterally or as a principal protagonist, in addition to the fact that United States dollars and soldiers will be more at risk than those of other nations, there will probably be greater risk to the integrity of our government than when our country is at peace. Preparedness for conflicts of interest and other ethics problems should be a much greater part of military preparedness than it is currently. These problems, however, like other problems in war, are sometimes difficult to predict.

After expenditures on foreign engagements, the next most pressing concern is expenditures on bailing out private companies. Here also, the Obama Administration is making relatively minor adjustments to the interventionist approach that emerged in the last few months of the Bush Administration. Much of corporate America is apparently too big to fail, and the government won't let some companies fail. As I have explained

elsewhere in an essay on bailouts and government ethics, this role for government is inconsistent with fiduciary obligations government officials have in managing public funds.[13] The risk of politicized decisions, conflicts of interest, insider trading, and other ethics problems is acute. The United States may not be able to continue to have a revolving door between the private sector and top echelons of government—and benefit from the experience that it brings into government—if government officials will not only regulate entire industries but also pick winners and losers among particular companies. These problems can be mitigated to some extent with stricter ethics rules, more systematized approaches to bailouts, and other strategies for preparedness, but here also preparedness will only go so far. Government ethics, along with the economic system in general, would be better off if the United States could find alternatives to bailing out companies that fail.

Then there is health care. It would be naive to assume that restructuring such a massive portion of the American economy can be accomplished without conflicts of interest and other ethics problems for the government officials who determine who pays what and who gets paid what in the new plan. There may be other reasons to proceed with health care reform, but this part of the cost should not be underestimated. Mitigating conflicts of interest and other ethics problems is possible if they are honestly acknowledged by the Administration and Congress, but these problems cannot be eliminated. The President's plan is so general that much of the detail is being supplied by Congress (at this point, there are several different versions of a plan being proposed). Allowing Congress rather than the White House to fill in the details avoids one of the political pitfalls of the Clinton health plan that in 1993 was sent in a near "finished" state to a Congress which refused to enact it. President Obama's approach of giving Congress a freer hand, however, could give lobbyists the upper hand as they use their relationships with hundreds of members of Congress to exert influence over the final product.

I do not suggest here that the only means of achieving good ethics in government is to have no government. I do suggest that when government expands the scope and size of its responsibilities and commensurate expenditures, government ethics problems are likely to expand as well. This cost, as well as the other costs of government activism and intervention, needs to be taken into account when policy makers deliberate about what the responsibilities of government should be. Unbridled growth of government itself could be the biggest threat to government ethics that this President or any other will confront.

Endnotes

1 *See* Memorandum to Designated Agency Ethics Officials and Inspectors General, from Stephen D. Potts, Director of OGE, dated August 14, 2000 DO-00-029 (1999 Conflict of Interest Prosecution Survey, discussing at pages 3-4 the facts and resolution of United States v. Richard Holbrooke).

2 Exec. Order No. 13,490, Ethics Commitments by Executive Branch Personnel, January 21, 2009, available at http://www.whitehouse.gov/the_press_office/ExecutiveOrder-EthicsCommitments.

3 Appointee is defined in Section (2)(b) of the Order:

> "Appointee" shall include every full time, non career Presidential or Vice-Presidential appointee, non career appointee in the Senior Executive Service (or other SES type system), and appointee to a position that has been excepted from the competitive service by reason of being of a confidential or policymaking character (Schedule C and other positions excepted under comparable criteria) in an executive agency. It does not include any person appointed as a member of the Senior Foreign Service or solely as a uniformed service commissioned officer.

4 Paragraph 2 of the pledge, intended for "All Appointees Entering Government," reads: "I will not for a period of 2 years from the date of my appointment participate in any particular matter involving specific parties that is directly and substantially related to my former employer or former clients, including regulations and contracts." Particular matters involving specific parties are usually thought to include contracts, investigations, lawsuits, and other matters with identifiable parties, but not government regulations that affect an entire industry. The specific reference to "regulation" in this Executive Order, however, implies that its reach could be considerably broader.

5 Paragraph 3 of the pledge reads:

> If I was a registered lobbyist within the 2 years before the date of my appointment, in addition to abiding by the limitations of paragraph 2, I will not for a period of 2 years after the date of my appointment: (a) participate in any particular matter on which I lobbied within the 2 years before the date of my appointment; (b) participate in the specific issue area in which that particular matter falls; or (c) seek or accept employment with any executive agency that I lobbied within the 2 years before the date of my appointment.

6 The Order has been favorably received by commentators. Dennis Thompson, for example, has commentated favorably on the objectives of the Executive Order while recognizing that the Administration thus far still lacks a coordinated approach to the broader range of ethics problems in government that are not addressed in the Order. *See* Dennis F. Thompson, Obama's Ethics Agenda: The Challenge of Coordinated Change, The Forum, Vol. 7 : Iss. 1, Article 8 (2009 Berkeley Electronic Press).

7 *See* Kenneth P. Vogel, Grassley After White House Ethics Waivers, Politico, June 10, 2009 (Senator Chuck Grassley (R-IA) has demanded disclosure of waivers and recusals under the Executive Order).

8 Paragraph 4 of the pledge required under the Order states:

> If, upon my departure from the Government, I am covered by the post employment restrictions on communicating with employees of my former executive agency set forth in section 207(c) of title 18, United States Code, I agree that I will abide by those restrictions for a period of 2 years following the end of my appointment.

9 Paragraph 5 of the pledge reads: "In addition to abiding by the limitations of paragraph 4, I also agree, upon leaving Government service, not to lobby any covered executive branch official or non career Senior Executive Service appointee for the remainder of the Administration."

10 See Letter dated June 17 2009 from the Board of the Corporation for National and Community Service to Senator Charles Grassley (stating that Walpin was ineffective as Inspector General, but not providing further explanation as to how or why Walpin was ineffective, and stating that the Board supported the President's decision to remove Walpin). An agency inspector general is charged with oversight of all agency personnel, including its head and its board members, if any, making a request from an agency head or board to remove an inspector general an awkward request from the vantage point of the White House. Complicating matters further in this instance, the chairman and vice chairman of the board and another board member had only a few months before commented favorably on Walpin's performance; for whatever reason, they appeared to have changed their minds (the Board had recently objected to both Walpin's Johnson/St. Hope report and another report by Walpin finding waste in millions of dollars in grants by a recipient organization). The legislative history of the 2008 Reform Act reflects Congressional intent that inspectors general be independent of their agencies, an objective that is undermined if the White House does not scrutinize carefully—and demand specifics in—complaints about inspectors general coming from agencies. Absent a pressing crisis, it also would be more appropriate for the White House to wait to act until specific allegations are investigated thoroughly by the Council

for the Inspectors General.

11 *See* Letter dated June 23, 2009 from 145 lawyers from the New York bar, including the author of this essay, to Senator Joseph Lieberman and other Senators vouching for Walpin's abilities ("[W]e have never seen Mr. Walpin to be 'confused, disoriented [or] unable to answer questions.'").

12 *See* Richard W. Painter, *Ethics and Corruption in Business and Government: Lessons from the South Sea Bubble and the Bank of the United States* (2006 Maurice and Muriel Fulton Lecture in Legal History, published by the University of Chicago Law School) (discussing a fraudulent scheme to use the South Sea Company through 1720 to fund England's national debt from wars with Spain, and then the 1789 plan of Alexander Hamilton for the United States government to assume the Revolutionary War debt of the individual states, which led to profiteering by speculators using inside information about the plan to buy up state notes at a fraction of par value).

13 *See* Richard W. Painter, Bailouts: An Essay on Conflicts of Interest and Ethics When Government Pays the Tab (September 2009), posted on SSRN at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1470910.

Religious Liberties

Why the Supreme Court Has Fashioned Rules of Standing Unique to the Establishment Clause

*By Carl H. Esbeck**

The U.S. Supreme Court is quite vigilant in enforcing its justiciability rules concerning standing to sue. For over half a century, however, the Supreme Court has reduced the rigor of its standing rules when a claim is lodged under the Establishment Clause of the First Amendment. The Court famously did so with respect to federal taxpayer standing in the venerable case of *Flast v. Cohen*,[1] but in no instance other than claims invoking the Establishment Clause is federal or state taxpayer standing ever permitted.[2] Less well known is the reduced rigor with which the Court has applied its standing rules when it comes to a plaintiff's "unwanted exposure" to a religious symbol or other speech attributable to the government.

The Roberts Court narrowly construed its prior cases permitting taxpayer standing to challenge government payments for religious purposes in *Hein v. Freedom From Religion Foundation, Inc*.[3] Recently the Court granted certiorari in *Salazar v. Buono*,[4] a case which raises the question of the standing required of a plaintiff in an "unwanted exposure" lawsuit that seeks the removal of a Latin cross on federal property because it is alleged to be in violation of the Establishment Clause.

The Supreme Court's cases on "unwanted exposure" do not require religious coercion or other individualized harm as plaintiff's "injury in fact." Rather, the cases evince a willingness to find standing when a plaintiff's status naturally results in him or her being personally exposed to the government's unwanted religious expression or the plaintiff is forced to assume a special burden to avoid such exposure. The plaintiff in *Salazar v. Buono* lacks that status and, hence, will not likely be found to have standing unless the Court extends its precedents.

I. Statement of the Case

In 1934, the Veterans of Foreign Wars erected a Latin cross on a location known as Sunrise Rock in the Mojave Desert in southeastern California.[5] This was unauthorized by the federal government, which owned the property. The cross is a memorial to members of the armed forces who died in World War I. In 1994, the site where the cross is located became part of the Mojave National Preserve, which is administered by the National Park Service, Department of the Interior. The Mojave National Preserve consists of 1.6 million acres of federal land in the Mojave Desert of southeastern California.

The respondent, Frank Buono, filed this lawsuit in March 2001, seeking a declaration that the Latin cross on government land violated the Establishment Clause, as well as an injunction ordering the permanent removal of the cross. At the time suit was filed, Buono was a retired employee of the National Park Service residing in Oregon. He retired twelve years ago in 1997. When Buono was still employed by the Park Service, he was assigned to the Mojave Preserve from January 22, 1995 to December 10, 1995. It was during this period that Buono learned of the Latin cross and visited the site at Sunrise Rock. Buono first became troubled when there was a request to erect a Buddhist stupa[6] near the cross. When the request was denied, Buono believed it was wrong for the cross to remain while similar access was denied for the stupa. His objections later evolved and expanded. Although retired, Buono retains an active interest in the Mojave National Preserve and visits the Preserve two to four times per year.

Buono is a Roman Catholic and testified that he does not find a Latin cross religiously offensive. Rather, he is offended because the cross remains at Sunrise Rock but similar access is denied to displays such as the Buddhist stupa, and because the National Park Service fails to remove the cross, a symbol of Christianity, from government land. When visiting the Mojave National Preserve, Buono has taken to avoiding Sunrise Rock so as not to be re-exposed to the cross, such avoidance being an added burden because it means not using Cima Road. One can see the Latin cross from the highway where Cima Road passes by Sunrise Rock. Cima Road is the most convenient road for accessing other areas of interest within the Preserve.

The Supreme Court has developed a three-part requirement for standing to sue. The plaintiff must have suffered, or is immediately threatened with, a specific "injury in fact." There must be a causal link between the defendant's conduct and plaintiff's injury. And the plaintiff seeks a remedy of a type traditionally rendered by our courts of law or equity.

The lower federal courts held that Buono has personalized "injury in fact" such that he has standing to bring this claim alleging a continuing violation of the Establishment Clause. The federal district court wrote as follows:

> Buono is deeply offended by the cross display on public land in an area that is not open to others to put up whatever symbols they choose. A practicing Roman Catholic, Buono does not find the cross itself objectionable, but stated that the presence of the cross is objectionable to him as a religious symbol because it rests on federal land.[7]

First quoting with approval this passage by the district court, as well as taking note of Buono's avoidance of Cima Road, the Ninth Circuit Court of Appeals went on to observe that:

> Buono is, in other words, unable to "freely us[e]" the area of the Preserve around the cross because of the government's allegedly unconstitutional actions…. We have repeatedly held that inability to unreservedly use public land suffices as injury-in-fact.... Such inhibition constitutes "personal injury suffered... *as a consequence* of the alleged constitutional error," beyond simply "the psychological consequence presumably produced by observation of conduct with which one disagrees."[8]

* *Carl H. Esbeck is the R.B. Price Professor and Isabelle Wade & Paul C. Lyda Professor of Law, University of Missouri.*

Given Buono's testimony that as a Catholic he suffers no religious offense because of the cross, spiritual injury cannot be a basis for "unwanted exposure" standing. That leaves two other possibilities: (1) offense because others cannot erect their symbols near where the cross is located; or (2) offense that the cross, a Christian symbol, stands on government property in violation of the separation of church and state. From the Ninth Circuit's statement quoted above, the circuit court—while noting both offensives as Buono's claimed "injury in fact"—is relying principally on Buono's "unwanted exposure" to the continued presence of a Latin cross on government property. Moreover, notes the circuit panel, Buono found this offense sufficiently weighty that he has taken to avoiding Cima Road and thereby incurring additional travel burdens as he explores the Mojave National Preserve.

Buono lacks third-party standing to complain that others are denied access to Sunrise Rock so that they might erect their own symbols.[9] Thus, Buono's offense that others are denied their rights is a claim of "injury in fact" for the Buddhists who sought to erect a stupa some years back.[10]

The plaintiff's other claim of "injury in fact" is a bit more involved. Buono seeks only injunctive relief from an ongoing injury. He does not seek damages. That leaves Buono's alleged ongoing "injury in fact" as being either: (1) unwanted exposure to the cross because of the government's failure to meet its duty of church-state separation which requires, in his view, removal of the cross from government land; or (2) restricted use of Cima Road to avoid being re-exposed every time he observes the government's failure to remove the cross. The first allegation, however, is a claim of "injury in fact" when a strict separationist is offended by a church-state violation while observing a religious symbol on government land. That is like the claimed "injury" discussed and rejected in *Valley Forge Christian College v. Americans United*.[11] And the second allegation of "injury in fact" is one of restricted use of Cima Road *because of* Buono's offense that the government has failed to remove the cross. Thus the second alleged harm (avoiding offense) logically collapses into the first (being offended).

With respect to the alleged church-state violation observed by Respondents, Americans United for Separation of Church and State, Inc., et al., the *Valley Forge* Court held:

> Although respondents claim that the Constitution has been violated, they claim nothing else. They fail to identify any personal injury suffered by them as a consequence of the alleged constitutional error, other than the psychological consequence presumably produced by observation of conduct with which one disagrees.... It is evident that respondents are firmly committed to the constitutional principle of separation of church and State, but standing is not measured by the intensity of the litigant's interest or the fervor of his advocacy.[12]

Buono tries to circumvent this passage in *Valley Forge* by asserting he suffers a personal injury in that he does not use Cima Road to avoid re-exposure to the Latin cross.

That is not enough for Buono to secure standing, as the *Valley Forge* Court went on to explain. The Court distinguished the facts before it in *Valley Forge* from that of the parents and school-age children exposed to unwanted prayer and devotional Bible reading in *Abington School District v. Schempp*:[13]

> "The parties [in *Schempp*] are school children and their parents, who are directly affected by the laws and practices against which their complaints are directed." ... The plaintiffs in *Schempp* had standing, not because their complaint rested on the Establishment Clause—for as *Doremus* [*v. Board of Education*, 342 U.S. 429 (1952),] demonstrated, that is insufficient—but because impressionable schoolchildren were subject to unwelcome religious exercises or were forced to assume special burdens to avoid them. [Americans United, *et al.*] have alleged no comparable injury.[14]

The Supreme Court's "unwanted exposure" precedents require that a plaintiff's status naturally result in being personally exposed to offensive religious expression by the government or forced to assume special burdens to avoid such exposure. In *Schempp*, the claimants' natural circumstance of public school attendance was such that the students were brought into personal exposure to the unwelcomed prayer and biblical devotions or forced to assume special burdens to avoid them.

This sensible rule has developed to prevent standing by contrivance. Requiring "injury" so as to have standing to sue can easily be manufactured if all one has to do is travel several miles to the site of a religious symbol or other expression of the government's and personally observe it on one occasion. Thus, it makes sense that a plaintiff's status (e.g., student, legislator, local municipal citizen) must naturally bring him or her into personal contact with the offending expression.

Buono's ongoing claim is that he will suffer an offense cognizable under the Establishment Clause if he travels to observe the cross which he deems a church-state violation, or he is "forced to assume special burdens to avoid" being re-exposed to the church-state violation. However, Buono's status does not naturally subject him to personal exposure to the cross. Buono's request for injunctive relief means that he necessarily avers an ongoing violation of the Establishment Clause. But he is a retired employee of the National Park Service residing in Oregon. Buono's visits to the Preserve are totally at his own free will. It is not as if Buono is currently employed by the Park Service and his job duties require that, from time to time, he travel Cima Road past Sunrise Rock. Buono's path to standing is foreclosed by *Valley Forge*, as well as that Court's reliance on *Schempp* and *Doremus*.

II. In Cases Raising "Unwanted Exposure" to Religious Expression Attributable to the Government, Reduced-Rigor Standing Has Been Permitted Only Where the Plaintiff's Status Naturally Results In Personal Exposure to the Unwanted Religious Expression

There is a very close connection between "injury in fact" for purposes of standing and damages (or "harm") as a necessary element of every claim under the Establishment Clause and for which plaintiff seeks a remedy. Indeed, they usually have been treated as one and the same by the Supreme Court. Therefore, the standing question in this case puts at issue a crucial element for stating a claim under the Court's modern Establishment Clause.

As with taxpayer standing (discussed Part III, *infra*), the Supreme Court's "unwanted exposure" cases under the Establishment Clause have resulted in reduced-rigor rules with

respect to the "injury in fact" required for standing. However, this reduction in the rigor with which "injury" is assessed is a narrow exception[15]—same as it is with taxpayer standing.[16] Reduced rigor in the required "injury" has been permitted only in cases challenging religious symbols or other expression attributable to the government. And only then does the lesser "injury" suffice where the plaintiff's status naturally results in personal exposure to the unwanted religious expression, or the plaintiff is forced to assume a special burden to avoid re-exposure.

The Supreme Court's cases of "unwanted exposure" to government religious speech are not great in number—just sixteen. Moreover, in nearly all of these cases—just three exceptions—the plaintiff's standing was not challenged on appeal by the government and thus was not an issue argued by counsel and decided by the Court. This second line of cases, therefore, have less to teach us with respect to what the Court minimally requires to have the "injury in fact" required for standing to bring a case of "unwanted exposure" to religious speech by the government. In chronological order the cases are as follows:

1. *McCollum v. Board of Education*[17] invalidated a local school district's program allowing nearby churches to hold optional religion classes in public school classrooms during regular school hours. The plaintiff was a resident and taxpayer of the local school district, and "a parent whose child was then enrolled in the Champaign public schools."[18] Also relevant to plaintiff's subjection to the program to have standing to challenge it, the Court said:

> The operation of the State's compulsory education system thus assists and is integrated with the program of religious instruction carried on by separate religious sects. Pupils compelled by law to go to school for secular education are released in part from their legal duty upon the condition that they attend the religious classes.[19]

The government's challenge to plaintiff's standing was rejected without analysis in a single sentence: "A second ground for the motion to dismiss is that the appellant lacks standing to maintain the action, a ground which is also without merit. Coleman v. Miller, 307 U.S. 433, 443, 445, 464."[20] (*Coleman* addressed the jurisdiction of the Court to review actions by state legislators said to have ratified a proposed amendment to the federal Constitution.) Accordingly, we do not have an explanation by the Court with respect to what "injury in fact" is required to file a case of "unwanted exposure" to religious expression by the government.

2. *Doremus v. Board of Education*[21] challenged teacher-led devotional Bible reading in New Jersey public schools. However, the Court did not reach the merits. Some plaintiffs, claiming status as state taxpayers, were dismissed for lack of standing. And a parent of a student subjected to the religious exercise had sued, but his child had subsequently graduated and thus his claim was moot. Accordingly, the case is not an instance where the Court ruled on the "injury in fact," required of a plaintiff claiming "unwanted exposure" to religious speech attributable to the government.

3. *Engel v. Vitale*[22] was a challenge to a statewide program of daily classroom prayer in New York public schools. The plaintiffs were "parents of ten pupils... insisting that use of this official prayer in the public schools was contrary to the beliefs, religion, or religious practices of both themselves and their children."[23] The government did not challenge the standing of the plaintiffs. That is surprising because the objecting parents and their school-age children could obtain an opt-out from the prayer exercise.[24] So once again the case did not present an instance where the Court determined the "injury in fact" required of a plaintiff claiming "unwanted exposure" to religious speech attributable to the government.

The fact that the "observance on the part of the students is voluntary," however, did not escape the Court's notice.[25] The prayer being voluntary would make a difference under the Free Exercise Clause, explained the Court, where coercion is an essential element of the *prima facie* claim. But with respect to the Establishment Clause, coercion or compulsory exposure to the prayer need not be shown.[26] This is because the object of the modern Establishment Clause is to separate church and state so as to prevent injury to either or both, as opposed to being a rights-based claim with its object being to prohibit individual religious harm.[27] In Part III, *infra*, it will be shown how this relates to standing, and thus why the modern Court has fashioned "reduced rigor" standing rules only under the Establishment Clause.

4. *School District of Abington Township v. Schempp*[28] involved consolidated cases from Philadelphia and Baltimore, both challenging daily classroom prayer and devotional Bible reading in public schools. In both instances, the religious exercises were optional.[29] In the Philadelphia case, the plaintiffs were:

> Edward Lewis Schempp, his wife Sidney, and their children, Roger and Donna… members of the Unitarian Church in Germantown, Philadelphia, Pennsylvania, where they... regularly attend religious services.... The [two] children attend the Abington Senior High School, which is a public school operated by appellant district.[30]

Also, "Edward Schempp and the children testified as to specific religious doctrines purveyed by a literal reading of the Bible 'which were contrary to the religious beliefs which they held and to their familial teaching.'"[31]

In the Baltimore case, the plaintiffs were "Mrs. Madalyn Murray and her son, William J. Murray III, ... both professed atheists."[32] The "petition particularized the petitioners' atheistic beliefs and stated that the rule, as practiced, violated their rights 'in that it threatens their religious liberty by placing a premium on belief as against non-belief and subjects their freedom of conscience to the rule of the majority....'"[33]

The lack of plaintiffs' standing to challenge the religious practices under the Establishment Clause was raised as an issue by the government.[34] The Court reasoned in footnote 9 that the plaintiffs had standing as follows:

> [T]he requirements for standing to challenge state action under the Establishment Clause, unlike those relating to the Free Exercise Clause, do not include proof that particular religious freedom are infringed.... The parties here are school children and their parents, who are directly affected by the laws and practices

against which their complaints are directed. These interests surely suffice to give the parties standing to complain.

Thus, standing under the modern Establishment Clause is not only different, but the need for "injury in fact" is of lesser rigor. That much is clear. Footnote 9 cites as authority *McGowan*, *Engel*, and *Doremus*, but as we have seen in none of those cases did the government challenge the plaintiffs' standing to bring an "unwanted exposure" claim.

As in *Engel*, the *Schempp* Court explained its lack of concern that plaintiffs did not prove they were victims of the government's compulsion or coercion. Coercion is an element of a Free Exercise Clause claim which is rights-based, but compulsion is not required to state a claim under the Establishment Clause.[35] This is because the Establishment Clause is about policing the boundary between church and state. "[T]he Court found that the 'first and most immediate purpose [of the Establishment Clause] rested on the belief that a union of government and religion tends to destroy government and to degrade religion.'"[36] In Part III, *infra*, it will be shown how this relates to standing, and thus why the modern Court has fashioned "reduced rigor" standing rules only under the Establishment Clause.

5. *Chamberlin v. Dade County Board of Public Instruction*,[37] citing *Schempp*, summarily struck down prayer and devotional Bible reading in the Dade County, Florida public school district. The plaintiffs were parents of school-aged children enrolled in junior high and elementary schools in Dade County.[38] The plaintiffs' standing to raise an "unwanted exposure" claim was not challenged by the government in the Supreme Court, and thus we have no guidance on the needed "injury" from the Court.

6. *Stone v. Graham*[39] struck down a state law requiring the posting of the Ten Commandments in all public school classrooms. Plaintiffs described themselves "as a Quaker, a Unitarian, a non-believer, a mother of school age children and public school teacher, two children of compulsory school age attending public schools, a Jewish Rabbi, and as taxpayers."[40] The plaintiffs' standing to raise an "unwanted exposure" claim was not challenged by the government before the Supreme Court, and thus we have no guidance on the matter from the Court.

7. *Marsh v. Chambers*[41] upheld a state legislative practice of hiring a chaplain to offer a prayer at the beginning of each day when the legislature is in session. The plaintiff was simply described as "a member of the Nebraska Legislature and a taxpayer of Nebraska."[42] The Court also noted that the plaintiff "claiming injury by the practice is an adult, presumably not readily susceptible to 'religious indoctrination,' or peer pressure."[43] Although the government had challenged the plaintiff's standing in the circuit court,[44] it did not again press the issue before the Supreme Court.[45] Although conceded by the state, the Supreme Court nevertheless volunteered the following: "[W]e agree that Chambers, as a member of the legislature and as a taxpayer whose taxes are used to fund the chaplaincy, has standing to assert the claim."[46] Thus a person vested with the status of a legislator who is regularly in the legislative chamber when the offending prayer takes place is sufficient "injury in fact" to have standing in this "unwanted exposure" case.

8. *Lynch v. Donnelly*[47] upheld a municipal practice of displaying a nativity scene of Mary, Joseph, and the Christ child as part of a larger Christmas holiday scene in a park. The display was located in a private park in the heart of the shopping district.[48] The plaintiffs were described as Pawtucket, Rhode Island "residents and individual members of the Rhode Island affiliate of the American Civil Liberties Union, and the affiliate itself."[49] The Court's majority opinion does not discuss standing, thus it appears the government did not challenge plaintiffs' claimed "unwanted exposure" injury giving rise to standing.

In a now famous concurring opinion, Justice O'Connor first stated her "endorsement or disapproval test." Her test identifies an injury that is personal to certain plaintiffs that the Establishment Clause is said to prevent, namely that the Establishment Clause "prohibits government from making adherence to a religion relevant in any way to a person's standing in the political community."[50] Justice O'Connor goes on with what in her view is the nature of the "injury in fact":

> One is excessive entanglement with religious institutions, which may interfere with the independence of the institutions, give the institutions access to government or governmental powers not fully shared by nonadherents of the religion, and foster the creation of political constituencies defined along religious lines. The second and more direct infringement is government endorsement or disapproval of religion. Endorsement sends a message to nonadherents that they are outsiders, not full members of the political community, and an accompanying message to adherents that they are insiders, favored members of the political community. Disapproval sends the opposite message.[51]

A violation of the endorsement test always results in a plaintiff's "religious" injury because the test is contingent on "adherence [or nonadherence] to a religion." This "endorsement or disapproval" test has possibilities for identifying the personal religious injury that naturally flows from one's status as local citizen when the church-state matter at issue is "unwanted exposure" to a government's religious expression. But the injury must be religious, unlike that claimed by Buono. That said, it is not clear the extent to which a majority of the current Supreme Court embraces Justice O'Connor's test. The endorsement test would limit "unwanted exposure" standing to instances where there is religious injury. That is contrary to most of the Court's array of sixteen "unwanted exposure" cases collected here. Accordingly, there is no reason to limit "unwanted exposure" standing to instances of religious injury.

9. *Wallace v. Jaffree*[52] struck down a state law requiring that public schools begin the day with a moment of silence by students for prayer or meditation. The law was found to have a religious purpose.[53] The plaintiff challenging the law was a parent who sued on behalf of "three of his minor children; two of them were second-grade students and the third was then in kindergarten."[54] Plaintiff's standing to challenge the state law was not raised by the government. So once again we do not have the benefit of the Court's discussion of what minimal "injury" is required in an "unwanted exposure" claim.

10. *Edwards v. Aguillard*[55] struck down a state law requiring public schools to teach creationism whenever evolution is taught. The law was found to have a religious purpose.[56] The plaintiffs challenging the law "included parents of children attending Louisiana public schools, Louisiana teachers, and religious leaders."[57] The Court went on to observe:

> Families entrust public schools with the education of their children, but condition their trust on the understanding that the classroom will not purposely be used to advance religious views that may conflict with the private beliefs of the student and his or her family. Students in such institutions are impressionable and their attendance is involuntary. The State exerts great authority and coercive power through mandatory attendance requirements, and because of the students' emulation of teachers as role models and the children's susceptibility to peer pressure.[58]

Thus the "harm" to plaintiffs' school-age children was the natural consequences of their status as students in Louisiana schools.

Once again there was no challenge by the government before the Supreme Court to plaintiffs' standing to call into question the state law. So we can only infer the "injury" needed for standing in a case of "unwanted exposure" to government religious speech.

11. *County of Allegheny v. Greater Pittsburgh ACLU*[59] involved challenges to two local governmental displays during the December holiday season. The Court struck down a nativity scene inside the county courthouse, and upheld an outdoor display of a Menorah, Christmas tree, and liberty banner at a different location jointly operated by the city and county. The plaintiffs challenging both displays were "the Greater Pittsburgh Chapter of the American Civil Liberties Union and seven local residents" of the city and county.[60] Once again the government did not challenge the plaintiffs' standing before the Court.

12. *Lee v. Weisman*[61] struck down the practice of inviting clergy to offer prayers at public school commencement ceremonies. Attendance at the ceremony was voluntary, and no penalty attached to a student who did not attend.[62] The plaintiffs challenging the practice were "Daniel Weisman, in his individual capacity as a Providence taxpayer and as [father] of Deborah," a student now graduated from the middle school, and enrolled in the high school where a similar prayer arrangement was conducted at its commencement.[63] Plaintiffs' standing was discussed. The Court said:

> We find it unnecessary to address Daniel Weisman's taxpayer standing, for a live and justiciable controversy is before us. Deborah Weisman is enrolled as a student at Classical High School in Providence and from the record it appears likely, if not certain, that an invocation and benediction will be conducted at her high school graduation.[64]

Once again the voluntary nature of the ceremony—hence lack of compulsion—did not make a difference so long as the claim is brought under the Establishment Clause.

13. *Santa Fe Independent School District v. Doe*[65] struck down a public school process whereby a student is elected by fellow students to offer words of inspiration (with prayer as a likely choice) over the loudspeaker system before high school football games. The plaintiffs challenging the practice were "two sets of current or former students and their respective mothers. One family is Mormon and the other is Catholic."[66] The government did not challenge the standing of the plaintiffs to bring their claim under the Establishment Clause.

14. *Elk Grove Unified School District v. Newdow*[67] concerned a plaintiff who was denied standing to challenge the words "under God" in the Pledge of Allegiance recited by public school students, including his daughter, at the beginning of each school day. Although the pledge was optional, both the daughter and her mother, who held legal custody, wished to have the daughter recite the pledge. Standing was denied because the plaintiff, although the student's father, was a noncustodial parent having no say in the matter. Accordingly, *Newdow* does not discuss the "injury in fact" needed for standing by a plaintiff complaining of "unwanted exposure" to religious expression attributable to the government.

15. *Van Orden v. Perry*[68] upheld the constitutionality of a Ten Commandments monument, one of several monuments on display on the grounds outside the State of Texas Capitol. The plaintiff challenging the monument was described as follows:

> Thomas Van Orden is a native Texan and a resident of Austin. At one time he was a licensed lawyer, having graduated from Southern Methodist Law School. Van Orden testified that, since 1995, he has encountered the Ten Commandments monument during his frequent visits to the Capitol grounds. His visits are typically for the purpose of using the law library in the Supreme Court building, which is located just northwest of the Capitol building.
>
> Forty years after the monument's erection and six years after Van Orden began to encounter the monument frequently, he sued….[69]

As one trained as a lawyer but without a law office or library of his own, as well as a citizen of Austin, it was natural that he took advantage of the free use of the law library near the Capitol. The government did not challenge Van Orden's standing before the Court.

16. *McCreary County v. ACLU of Kentucky*[70] struck down the Ten Commandments placed in display cases, along with other historical documents, in two county courthouses in the State of Kentucky. The plaintiffs challenging both displays were all too briefly described as "American Civil Liberties Union of Kentucky, et al."[71] The Court also explained that in both counties "the hallway display was 'readily visible to... county citizens who use the courthouse to conduct their civic business, to obtain or renew driver's licenses and permits, to register cars, to local taxes, and to register to vote.'"[72] A lower court opinion explains that in addition to the ACLU of Kentucky, the plaintiffs were Lawrence Durham and Paul Lee.[73] From the context it is apparent that Durham and Lee are residents of the county. The lower court said the ACLU had organizational standing because it "has members in Pulaski County who would have standing for the same reason that the named plaintiffs have standing."[74] And the government suggested in its briefs that "the Ten Commandments were posted in order to teach Pulaski County residents about American religious

history and the foundations of the modern state."[75] Although before the district court the government challenged plaintiffs' standing because they lacked the necessary "injury in fact,"[76] having lost the issue at the trial level the government did not raise the standing question before the Supreme Court. One can infer from *McCreary County* that a county citizen who has to visit the site of the offending religious message in order to do necessary legal transactions with the county government has the status and personal "unwanted exposure" so as to have "injury in fact" for purposes of standing.

It is remarkable that in only three out of sixteen cases has plaintiffs' standing been challenged before the U.S. Supreme Court on the basis that there was no "injury in fact" due to "unwanted exposure" to the government's religious speech. The three cases are *Schempp*, *Marsh*, and *Weisman*. These three cases involve plaintiffs who are parents and their school-age children, and a legislator. The rule to draw from *Schempp*, *Marsh*, and *Weisman*, and to a lesser degree the other thirteen cases where lack of standing might have been raised but was not, is that the "injury" required in an "unwanted exposure" case is that the offended plaintiff's status in life must have brought him or her into personal contact with the government's religious symbol or other expression.[77] Following this rule will prevent parties who would contrive their exposure "injury" by going out of their way to travel to the site of a religious symbol and observe it merely to acquire standing. Buono has no such status such that he has "injury in fact" endowing him with a "case" or "controversy" for which he has standing to sue.

III. Why the Court Has Permitted Reduced-Rigor Standing in Only Two Instances, Both Involving Claims under the Establishment Clause

In circumstances very different than the one before the Supreme Court, a claimant under the Establishment Clause can have individualized "injury in fact" that meets all of the normal requirements for standing. These harms run from economic loss, to inability to qualify for public office, to restrictions on academic inquiry.[78] But in each of the six cases set out in the footnote, plaintiffs had conventional "injury in fact" and thus met the usual "case" or "controversy" requirements for standing. That is not so with respect to cases involving "unwanted exposure" to religious symbols or other speech fairly attributable to the government. Only in two types of cases—taxpayer and "unwanted exposure" claims—has the Court applied a reduced-rigor test for "injury in fact" so as to ease the path to reaching the merits of a claim under the Establishment Clause. Why is that so?

The Court's modern view of the Establishment Clause was instituted sixty-two years ago with its decision in *Everson v. Board of Education of the Township of Ewing*.[79] Because both the Free Exercise and Establishment Clauses are pro-religious freedom,[80] the question arose early with respect to how the two Clauses were to be distinguished. The Court's answer came soon in *Engel v. Vitale*[81] and was reaffirmed a year later in *School District of Abington Township v. Schempp*.[82] As the *Engel* Court said:

> Although these two clauses may in certain instances overlap, they forbid two quite different kinds of governmental encroachment upon religious freedom. The Establishment Clause, unlike the Free Exercise Clause, does not depend upon any showing of direct governmental compulsion and is violated by the enactment of laws which establish any official religion whether those laws operate directly to coerce nonobserving individuals or not.[83]

The Court goes on to explain that the reason that coercion is not a required element of a no-establishment claim is that the Clause is first and foremost about the separation of church and state.[84] Church-state separation is a relationship between two centers of authority. This is not due to any hostility to religion but for the protection of both the freedom of the church and to prevent division within the body politic when government takes sides on explicitly religious questions. Disestablishment deregulated religion, thus protecting both church and state. Individual liberties are protected by the Establishment Clause only as a consequence of keeping these two authorities in right order relative to each other, and sometimes the individual liberties protected are not religious, e.g., economic liberty, access to public office, freedom of academic inquiry, etc.[85]

It thus developed in the Supreme Court that the Free Exercise Clause was confined to addressing those situations where religious practice or observance had come under state coercion. Without evidence of coercion, either standing was denied (consider the discussion in Part II, *supra*, in *Engel* and *Schempp*) or the free exercise claim failed on the merits.[86] The Free Exercise Clause is thus a rights-based claim; it runs in favor of religious individuals and faith groups they form.[87]

The Establishment Clause operates quite differently—all the while retaining its character as pro-religious freedom. The Establishment Clause works to limit the power of government. In that sense, it operates much like a structural clause.[88] Many an individual claimant need not show personal religious harm to win a claim under the Establishment Clause.[89] Indeed, in two lines of cases the claimant does not need to show personalized injury at all—taxpayer and "unwanted exposure" cases. This came about because—unlike free exercise which is rights-based—the Court's modern Establishment Clause is about separation of church and state. When church and state are not rightly ordered, the harm or damage might be other than religious. As this Court said in *McGowan v. Maryland:*[90]

> If the purpose of the "establishment" clause was only to insure protection for the "free exercise" of religion, then what we have said above concerning appellants' standing to raise the "free exercise" contention would appear to be true here. However, the writing of Madison, who was the First Amendment's architect, demonstrate that the establishment of a religion was equally feared because of its tendencies to political tyranny and subversion of civil authority.[91]

Such oppression often resulted in injury other than religious harm (*McGowan* was economic), indeed it can result in instances where no one has individualized injury and hence no one has conventional standing to sue. This is called a "generalized grievance."[92]

In this regard, the modern Supreme Court's work via the Establishment Clause to keep rightly ordered church and state causes the no-establishment principle to operate in many

respects like the structural clauses of the Constitution which separate the powers of the three federal branches. And just as some violations of separation of powers can occur with no one personally harmed, a "generalized grievance" can and does occur where there is a colorable violation of the modern Establishment Clause but no one with individualized harm. The first such case appeared before the Court in *Flast v. Cohen*,[93] and the Court responded by permitting limited federal taxpayer standing. Stated differently, the surrogate of taxpayer as plaintiff with "injury in fact" permitted the Court to reach the merits of some no-establishment claims that would otherwise be nonjusticiable because no one had individuated injury to acquire standing.[94]

But as the plurality in *Hein v. Freedom From Religion Foundation, Inc.*,[95] recently said, *Flast* inadequately acknowledged—even when limited as it was to claims under the Establishment Clause—the distortion wrought to the doctrine of separation of powers.[96] So *Flast*, while still good law, has not been expanded.

Flast is not the only line of cases where the modern Court reduced the normal rigor of standing when it comes to the Establishment Clause. The other line is where plaintiffs claim injury due to "unwanted exposure" to religious speech but who did not suffer the coercion or compulsion that would normally be associated with the individualized injury required for standing. Early on, as we saw in Part II, *supra*, the most common case was public school students exposed to religion classes, prayer, and biblical devotions, but the exercise was optional. The Court's response was to reduce the rigor of the required "injury in fact" by stating that coercion was not an element of a claim under the Establishment Clause. Like *Flast*, this necessarily required a trade off. With respect to the Court's co-ordinate branches, reducing the rigor of standing was at the expense of the doctrine of separation of powers. With respect to the States, reduced rigor standing was at the expense of federalism. In either instance, reducing the "injury" needed for standing permitted the Court to reach the merits of an Establishment Clause claim that would otherwise be outside the Court's subject matter jurisdiction. Like *Flast*, however, the reduction in the rigor of normal standing requirements was narrow: only where the offended plaintiff's status naturally caused him or her to personally come into "unwanted exposure" to the government's religious expression was standing permitted.

Buono's status does not fit within the limits of the Supreme Court's narrow exception with respect to its "unwanted exposure" cases. He has no responsibilities as a local citizen, such as in *McCreary County*, to frequent the site at Sunrise Rock. He holds no status as a student or student's parent, such as in *McCollum* or *Schempp*, which results in his presence at the site of the Latin cross, nor is he a legislator needing to be present in chambers to do his job as in *Marsh*. Assuming Buono has paid the admission fee to enter the Mojave National Preserve, certainly he has a legal right to be present at Sunrise Rock. But his presence is entirely by his free and unrestrained choice. Such a circumstance is no different than a citizen of India, who as a resident alien with a five-year visa to reside and work in Massachusetts, takes a vacation to Southeast California and pays the admission fee to enter the Mojave Preserve and happens to spot the Latin cross out of the windshield of his automobile as he drives by Sunrise Rock. This is one of those instances where if Buono has Article III standing to sue, then the entire population of people within the jurisdiction of the United States has standing to sue upon a single automobile ride along Cima Road. None of the Court's sixteen cases set out in Part II, *supra*, is nearly so expansive.

CONCLUSION

The Supreme Court will have to expand its law with respect to "unwanted exposure" cases to find Frank Buono has standing to sue. Just the opposite inclination was demonstrated by the Roberts Court in *Hein*, and there is no obvious reason that has changed. *Hein* reaffirmed federal taxpayer standing, as originally announced in *Flast*, when the no-establishment principle was at risk because of congressional appropriation legislation. The plurality in *Hein* was right to do so. At the same time the *Hein* plurality was correct to not expand taxpayer standing into the myriad of discretionary decisions by officials in the Executive Branch. *Flast* represented a tolerably small compromise to separation of powers, in return for the Supreme Court taking its rightful role as a co-equal branch with Congress in the duty to police the boundary between church and state. The plaintiffs in *Hein*, on the other hand, were asking for the Court to toss overboard the doctrine of separation of powers.[97] Frank Buono's assertion of standing in this "unwanted exposure" case is far more like the plaintiffs in *Hein* than in *Flast*.

Further, should the Supreme Court dismiss Frank Buono's complaint for lack of standing there will be no need to resolve the merits of Buono's difficult no-establishment claim involving prickly issues of congressional motive.[98] Generally the Court would welcome the opportunity to not extend itself and resolve a difficult constitutional question on the merits when the matter can so sensibly be disposed of on jurisdictional grounds.

Endnotes

1 392 U.S. 83 (1968).

2 *See, e.g.,* DaimlerChrysler Corp. v. Cuno, 547 U.S. 332, 347-49 (2006) (denying taxpayer standing in complaint alleging a violation of the Dormant Commerce Clause).

3 551 U.S. 587 (2007) (plurality opinion). *See* Carl H. Esbeck, *What the* Hein *Decision Can Tell Us About the Roberts Court and the Establishment Clause*, 78 Miss. L. J. 199 (2008).

4 129 S. Ct. 1313 (Feb. 23, 2009) (No. 08-472), *cert. granted,* Buono v. Kempthorne, 527 F.3d 758 (9th Cir. 2008).

5 The facts recited in this Statement of the Case are from uncontested findings of the district court. *See* Buono v. Norton, 212 F. Supp.2d 1202, 1204-07 (C.D. Cal. 2002).

6 A stupa is a mound-like structure containing Buddhist relics.

7 212 F. Supp.2d at 1207.

8 Buono v. Norton, 371 F.3d 543, 547 (9th Cir. 2004) (citations omitted). Subsequent proceedings involving the merits of the Establishment Clause claim appear at *Buono v. Norton*, 364 F. Supp.2d 1175 (C.D. Cal. 2005), and *Buono v. Kempthorne*, 527 F.3d 758 (9th Cir. 2008). The government did not continue to challenge Frank Buono's standing in these later proceedings because the district and circuit courts had already ruled adversely on the matter. The government did not thereby waive its objection to Buono's

standing. Standing is derived from the "case" or "controversy" requirement in U.S. Const. Art. III, § 2, cl. 1. That provision defines the subject matter jurisdiction of the federal courts. Because it is structural rather than rights-based, objections to a federal court's lack of subject matter jurisdiction can never be waived. *See* Rule 12(h)(3), Federal Rules of Civil Procedure.

9 *See, e.g.*, Warth v. Seldin, 422 U.S. 490, 510 (1975) (summarizing those limited instances when third-party standing is permitted, none of which remotely apply here).

10 Such an averment states a free speech claim that equal access is being denied to a limited public forum, not a claim under the Establishment Clause about religious speech attributable to the government. *Cf.* City of Pleasant Grove, Utah v. Summum, 129 S. Ct. 1125 (2009).

11 454 U.S. 464, 485 (1982) (holding plaintiffs lack standing to challenge government plans to give surplus land to religious college).

12 *Id*. at 485-86.

13 374 U.S. 203 (1963).

14 *Valley Forge*, 454 U.S. at 487 n. 22.

15 It is a narrow exception because federal courts are of limited jurisdiction per U.S. Const. Art. III, § 2, cl. 1. And, when Article III jurisdiction is pushed to its outer reaches, then separation of powers is necessarily implicated.

16 Hein v. Freedom From Religion Foundation, Inc., 551 U.S. 587, ___, 127 S. Ct. 2553, 2564 (2007) (plurality opinion) ("In *Flast*, the Court carved out a narrow exception to the general constitutional prohibition against taxpayer standing.").

17 333 U.S. 203 (1948).

18 *Id*. at 205.

19 *Id*. at 209.

20 *Id*. at 206

21 342 U.S. 429 (1952).

22 370 U.S. 421 (1962).

23 *Id*. at 423.

24 *Id*. 423 n. 2.

25 *Id*. at 430.

26 *Id*.

27 *Id*. at 430-32.

28 374 U.S. 203 (1963).

29 *Id*. at 224-25.

30 *Id*. at 206.

31 *Id*. at 208.

32 *Id*. at 211.

33 *Id*. at 212.

34 *Id*. at 224 n. 9.

35 *Id*. at 221, 223.

36 *Id*. at 221.

37 377 U.S. 402 (1964) (per curiam).

38 *See* Chamberlin v. Dade County Board of Public Instruction, 160 So.2d 97, 98 (Fla. 1964).

39 449 U.S. 39 (1980) (per curiam).

40 *See* Stone v. Graham, 599 S.W.2d 157, 159 (Ky. 1980).

41 463 U.S. 783 (1983).

42 *Id*. at 785.

43 *Id*. at 792.

44 *Id*. at 785.

45 *Id*. at 786, n. 4.

46 *Id*.

47 465 U.S. 668 (1984).

48 *Id*. at 671.

49 *Id*. at 671.

50 *Id*. at 687.

51 *Id*. at 687-88 (emphasis added).

52 472 U.S. 38 (1985).

53 *Id*. at 56-60.

54 *Id*. at 42.

55 482 U.S. 578 (1987).

56 *Id*. at 586-94.

57 *Id*. at 581.

58 *Id*. at 584.

59 492 U.S. 573 (1989) (plurality opinion in part).

60 *Id*. at 587.

61 505 U.S. 577 (1992).

62 *Id*. at 583.

63 *Id*. at 584.

64 *Id*.

65 530 U.S. 290 (2000).

66 *Id*. at 294.

67 542 U.S. 1 (2004).

68 545 U.S. 677 (2005) (plurality opinion).

69 *Id*. at 682.

70 545 U.S. 844 (2005).

71 *Id*. at 852.

72 *Id*. (citation omitted).

73 *See* American Civil Liberties Union of Kentucky v. Pulaski County, 96 F. Supp.2d 691 (E.D. Ky. 2000).

74 *Id*. at 694.

75 *Id*. at 698.

76 *Id*. at 694-95.

77 Frequency or regularity of exposure to the offending religious expression attributable to the government in some instances may be evidence that the plaintiff naturally has the status required to have "unwanted exposure" standing. Thus, for example, one would expect that a citizen of a municipality with a Holiday Christmas display to have more frequent exposure to the offending Christmas display on the lawn of the city hall, than say the frequency of exposure of one who lives five hundred miles away. Yet regularity of exposure is not a substitute for the needed status because frequency of exposure can easily be contrived. On the other hand, while personal exposure is always required, regularity of the exposure may be quite limited where a plaintiff is mandated by legal duty to incur exposure to the offending religious symbol or other speech. One example of such duty is where plaintiff's job responsibilities require exposure to the offending religious speech. Another example is where plaintiff's duty, albeit not regular, is as a county citizen called as a juror or witness in a lawsuit and the religious symbol must be passed just once or twice as plaintiff enters the courthouse. *See, e.g.,* Books v. Elkhart County, 401 F.3d 857, 862 (7th Cir. 2005).

78 Consider the department store in *Estate of Thornton v. Caldor, Inc.*, 472 U.S. 703, 707 (1985) (increased employment regulation resulting in economic harm), the tavern in *Larkin v. Grendel's Den, Inc.*, 459 U.S. 116, 118 (1982) (denial of a liquor license resulting in economic harm), the public school teacher desirous of expanding the science curriculum in *Epperson v. Arkansas*, 393 U.S. 97, 100 (1968) (hindrances to academic inquiry resulting in criminal charges and loss of job as "injury"), the forced taking of a theistic oath by a

freethinking atheist in *Torcaso v. Watkins*, 367 U.S. 488 (1961) (inability to qualify for public office as "injury"), shuttering one's business on Sunday in *McGowan v. Maryland*, 366 U.S. 420, 422, 430-31(1961) (economic harm to retail stores and criminal fines imposed on their employees), and closing one's retail store on Sunday in *Two Guys from Harrison-Allentown, Inc. v. McGinley*, 366 U.S. 582, 592 (1961) (lost business as economic harm).

79 330 U.S. 1 (1947). Not only did *Everson* incorporate the Establishment Clause making it applicable to state and local governments, but *Everson* looked to the history of disestablishment in the new American states, especially Virginia, for the principles behind the meaning of the Clause. *Id*. at 11-15. *See* Carl H. Esbeck, *The 60th Anniversary of the* Everson *Decision and America's Church-State Proposition*, 23 J. of Law & Religion 15 (2007-08).

80 Just like the First Amendment is pro-freedom of speech and pro-freedom of the press, the First Amendment is also pro-freedom of religion. Now, being pro-freedom of religious is markedly different from being pro-religion. The latter is prohibited by the modern Establishment Clause, thereby maintaining the requisite government neutrality. But the First Amendment is pro-religious freedom. Moreover, this is as true of the Establishment Clause as it is of the Free Exercise Clause. While commonplace to some, others will be surprised to have the Establishment Clause portrayed as pro-religious freedom. This is to say that the separation of church and state, properly conceived, is far more about protecting religious freedom than it is about furthering modernity's project to confine religion.

81 370 U.S. 421 (1962).

82 374 U.S. 203, 221, 223 (1963). *Accord* Lee v. Weisman, 505 U.S. 577 (1992).

83 *Id*. at 430.

84 *Id*. at 425-36.

85 *See* the cases collected in note 78, *supra*.

86 *See* Harris v. McRae, 448 U.S. 297, 320 (1980) (free exercise claim requires compulsion of religious belief); Tilton v. Richardson, 403 U.S. 672, 689 (1973) (plurality opinion) (coercion required to state free-exercise claim); Board of Education v. Allen, 392 U.S. 236, 248-49 (1968) (same).

87 Of course, to state a claim that involved coercion with respect to a religious practice did not mean that every such claim would be successful. And with the decision in *Employment Division v. Smith*, 494 U.S. 872 (1990), fewer free exercise claims do succeed.

88 *See* Carl H. Esbeck, *The Establishment Clause as a Structural Restraint on Governmental Power*, 84 Iowa L. Rev. 1 (1998). A more summary statement of the evidence that this Court has applied the Establishment Clause as if it were structural in nature, as well as how such a view explains not only this Court's special standing rules with respect to no-establishment but other validations as well, *see* Carl H. Esbeck, *The Establishment Clause as a Structural Restraint: Validations and Ramifications*, 18 J. of L. & Politics 445, 453-71 (2002).

89 *See* the cases collected in note 78, *supra*.

90 366 U.S. 420 (1961).

91 *Id*. at 430.

92 A famous trio of "generalized grievances" are represented by *Schlesinger v. Reservists Committee to Stop the War*, 418 U.S. 208 (1974) (colorable violation of structural clause in the Constitution, but no one with individualized injury and hence no one with standing), *United States v. Richardson*, 418 U.S. 166 (1974) (same), and *Ex parte Levitt*, 302 U.S. 633 (1937) (per curiam) (same).

93 392 U.S. 83 (1968).

94 There is speculation in *Flast*, repeated in later cases, to the effect that federal taxpayers have a personal right of conscience to not have the taxes they pay be appropriated by Congress to religious organizations, even when the appropriations are for secular purposes. James Madison's *Memorial and Remonstrance*, circulated in Virginia in the summer and fall of 1785, is cited as the origin of this assertion. However, Patrick Henry's bill, successfully opposed by Madison during Virginia's disestablishment struggles in 1784 and 1785, was legislation to impose a religious assessment for the payment of the salaries of Christian clergy. The assessment—or religious tax—was an earmarked tax. The money went into a special trust fund held by the state for the sole purpose of clergy salaries. Indeed, each taxpayer even got to designate which clergyman was to receive his or her tax payment. It was not a general tax the money of which went into the general U.S. Treasury, and from which at a later time Congress would appropriate money for all sorts of matters from B-2 bombers to price supports for dairy farmers. In the instance of an earmarked tax like Henry's, there is a direct causal link between taxpayer and the amount received by his or her clergyman. Causation being necessary to link the act of paying one's taxes under coercion to one's violation of religiously informed conscience. With most federal appropriations today, such as the Elementary and Secondary Education Act of 1965 at issue in *Flast*, there is no such causal link any more than a particular taxpayer is linked to dairy support payments. Justice Harlan, dissenting in *Flast*, saw the fiction in the personal conscience claim right from the start. *Flast*, 393 U.S. at 128-29. For a full account of the historical details, *see* Carl H. Esbeck, *Protestant Dissent and the Virginia Disestablishment, 1776 – 1786*, 7 Geo. J. L. & Pub. Pol'y 51, 79-81, 89-92 (2009).

95 551 U.S. 587 (2007) (plurality opinion).

96 *Id*. at __, 127 S. Ct. at 2569-70.

97 Carl H. Esbeck, *What the* Hein *Decision Can Tell Us About the Roberts Court and the Establishment Clause*, 78 Miss. L. J. 199, 217-20 (2008).

98 Question Presented No. 2 in the grant of certiorari in *Salazar v. Buono* reads: "Whether, even assuming respondent has standing, the court of appeals erred in refusing to give effect to the Act of Congress providing for the transfer of the land to private hands."

Smith, *Stormans*, and the Future of Free Exercise: Applying the Free Exercise Clause to Targeted Laws of General Applicability

*By Mark L. Rienzi**

Suppose that a new Christian church in town announced that it will have a special focus on the biblical story of Jesus turning water into wine. Accordingly, rather than the single sip of wine commonly consumed at other Christian services, worshippers at the new church drink several glasses of wine as part of the service, in imitation of the wedding guests at Cana.

Unfortunately, after church, the drunken worshipers spill into the streets and are generally a loud and raucous group. Neighbors complain about being woken up early by their noise. Drunk driving arrests, previously unheard of on Sunday mornings, begin to rise.

In an effort to eliminate the noise, traffic, and other problems caused by the churchgoers, the town outlaws the consumption of alcohol on Sunday mornings. The law is neutral and generally applicable—it outlaws all consumption, and applies to all drinkers, religious or not. The law is not passed out of any animus toward the new church, but rather as a good faith attempt to protect the rest of the population from the impact of the drunken faithful.

This hypothetical highlights an important blind spot in the way most courts approach Free Exercise cases. Does the Free Exercise Clause extend to situations where the legislature deliberately targets a religious practice, but does so for neutral reasons and is willing to extend the ban to people who happen to engage in the same practice for non-religious reasons? While one can imagine reasonable arguments on both sides about the constitutionality of the Sunday morning alcohol ban, it seems absurd to say that the Free Exercise Clause is not part of the equation. Yet under the First Amendment analysis presently employed by many courts, that result is entirely likely.

I. Two Types of Free Exercise Cases

Since the Supreme Court's 1990 decision in *Employment Division v. Smith*,[1] virtually all free exercise cases have been grouped into one of two categories. On one hand is the rare case where a law treats religious conduct more harshly than if the same conduct were engaged in for non-religious purposes. For example, in *Church of the Lukumi Babalu Aye, Inc. v. City of Hialeah*, the Court invalidated a law that outlawed ritual animal sacrifices because "certain religions" may propose to engage in such practices.[2] The law also included a series of exemptions making clear that the only animal killings addressed would be the ritual sacrifices conducted as part of the Santeria religion.[3] The Court found that such religious discrimination in the text or operation of a statute is subjected to strict scrutiny and therefore only permitted when narrowly tailored to a compelling state interest.[4] Accordingly, such outright discrimination against religion is virtually never permissible; as Justice Souter observed at the time, it is also exceedingly rare.[5]

On the other hand are laws that do not make express religious distinctions in their terms or operation.[6] For example, the ban on use of peyote at issue in *Smith* outlawed all peyote use, whether engaged in for religious reasons or not. Such "neutral and generally applicable" laws are not subject to free exercise analysis at all, and instead are permissible unless they violate some other constitutional principle.[7]

When viewed through the *Smith/Lukumi* dichotomy, the Free Exercise Clause appears to retain relevance only in the rare cases where the law treats religious conduct *worse* than the same conduct engaged in for secular reasons. Accordingly, except for rare cases of clear religious discrimination or animus, virtually all laws challenged on federal free exercise grounds are found to be "neutral and generally applicable," and therefore exempt from free exercise analysis under *Smith*.[8]

II. Targeted Laws of General Applicability

A third possibility exists. A legislature seeking to outlaw a particular religious conduct—perhaps for perfectly neutral reasons—may enact a law that is designed to target that religious conduct, but is nevertheless facially neutral and generally applicable.

In some cases, this targeting of religious conduct may be the result of animus—the lawmakers want to eliminate the religious conduct *because* it is religious. Courts generally appear able to address this type of targeting under *Lukumi*, particularly where the lawmakers are open about their anti-religion motivation.[9]

In many other cases, however, the targeting will occur without any particular animus toward religion at all. That is, the targeting of the religious conduct will occur not *because of* the religious nature of the conduct, but for some other reason (such as public safety). These laws still clearly target religious conduct—indeed, their very goal is to eliminate that conduct—but will usually apply whether the conduct is engaged in for religious reasons or any other reason.

The Sunday morning alcohol ban described above is this type of targeted law of general applicability. The law plainly targets religious conduct—its very purpose is to rein in the drunken faithful from the new church—but it is nevertheless neutral (i.e., the legislature was not restricting the drinking *because* it was religious, but because it caused other problems) and generally applicable (i.e., it applied to other drinkers as well, not just those from the new church).

To further illustrate this type of targeting, suppose a town has an influx of Orthodox Jewish and Seventh Day Adventist business owners who, for religious reasons, keep their shops closed on Saturdays. Many shoppers stop coming downtown on Saturdays, because so many businesses are closed. Citing the harm to the entire downtown business community from the Saturday closings, the town council enacts a law requiring all

* *Assistant Professor, The Catholic University of America, Columbus School of Law.* *Thanks to Kristen Morgan for research assistance.*

businesses to remain open on Saturdays. The law is passed for economic reasons, and not out of animus toward religion. The law is also generally applicable—although it was designed to address the Saturday closings caused by religious conduct, it also would prohibit Saturday closings for non-religious reasons.

The Saturday opening law is a targeted law of general applicability. It is *targeted*, because religious conduct was causing a particular problem (harm to the downtown economy) and the object of the law was to address that harm by stopping the religious conduct. Yet the law is also *neutral* (i.e., there is no evidence of animus and the closings are not targeted *because* of religion but for other reasons) and *generally applicable*, in that it would apply to the same problematic conduct if it happened to be engaged in for reasons other than religion (for example, if the owner is on vacation).

The Saturday opening law can be contrasted with the law at issue in *Smith*. In *Smith*, the challenged law was a general ban on drug use.[10] While the law had an incidental effect on religious users of peyote, the primary purpose of the drug ban was not to eliminate the religious use of peyote, but, presumably, to promote public safety. *Smith* would have involved targeting if the legislature had passed the drug law because of problems it had with religious uses of peyote. In contrast, while the Saturday opening law was not the result of religious animus and would apply generally, it was enacted in response to religious conduct, and its main goal is to eliminate that conduct.

This type of targeting of religious conduct is, of course, as burdensome on the religious actors as if the legislature had said "we want to get rid of the religious store owners, so let's make them keep their stores open on Saturdays." In either case, the lawmaker has identified religious conduct as problematic and set out to prohibit that conduct in a way that will force the store owners to either violate their beliefs or lose their stores. Yet, if the legislature betrays no animus, and if the legislature is willing to ban the same conduct in the rare situation where it occurs without religious motivation, courts generally excuse this type of targeting of religious conduct from free exercise analysis entirely.

If recent cases are any guide, the Saturday opening law would be analyzed by most courts as follows. First, the court would determine that, unlike the law at issue in *Lukumi*, the Saturday opening law is neutral. The court would analyze whether the text contains any indicators that it is designed to restrict only religiously motivated conduct, or to restrict that conduct because of the religious motivation. Courts would also look at the operation of the statute to make sure that, in operation, the law does not treat religious conduct differently from the same conduct engaged in for non-religious purposes.[11]

In determining neutrality, some courts would stop with the text and operation of the law, perhaps mindful of Justice Scalia's argument in his *Lukumi* concurrence that "it is virtually impossible to determine the singular 'motive' of a collective legislative body."[12] These courts would deem the law neutral and, if also generally applicable to non-religious closings, exempt from Free Exercise analysis under *Smith.*[13]

Most courts, however, would look into the law's history for any evidence of bad intent, namely that the law targeted the Saturday closings *because of* their religious motivation.[14] Finding a neutral reason for the law—namely protection of the downtown economy—the courts would likely deem the law "neutral and generally applicable," and therefore constitutional under *Smith*.[15] Most courts would not consider the fact that the law is targeted, in that the entire point of the law was to address a problem caused by religious conduct by prohibiting that conduct.

III. Example: The Ninth Circuit's Plan B Decision

The Ninth Circuit's recent decision in *Stormans, Inc. v. Selecky* provides a clear example of lawmakers targeting religious conduct through the use of generally applicable laws, and a court failing to even apply the Free Exercise Clause to analyze the challenged law.[16]

Stormans concerned a Washington Board of Pharmacy regulation adopted in response to certain pharmacists refusing to dispense a drug known as Plan B.[17] Plan B is often referred to as the "morning after pill" or as an "emergency contraceptive" because it can be taken after sexual intercourse to prevent pregnancy.[18] Both the drug's manufacturer and the FDA acknowledge that Plan B can work by preventing an already-fertilized egg from implanting in the uterus.[19] For this reason, some pharmacists object on religious grounds to selling or dispensing Plan B, because they believe they would be participating in terminating an already-started human life.[20]

In 2005, the state Pharmacy Board began getting calls asking its position about pharmacies and pharmacists who refused on religious grounds to sell Plan B.[21] The Board convened meetings with Washington State Pharmacy Association, Planned Parenthood and other interested parties to discuss whether and under what circumstances a pharmacist should be permitted to refuse to fill a prescription on religious grounds.[22] Ultimately, the Board adopted a draft rule that allowed a pharmacist to refuse to dispense a medication for religious reasons, but required that no pharmacist or pharmacy obstruct a patient's effort to obtain lawfully prescribed drugs.[23]

The immediate reaction to the Board's draft rule confirmed that the focus was on religious refusals. The same day, Washington's Governor Christine Gregoire sent a letter to the Chairman of the Pharmacy Board "stating her strong opposition to the draft rule."[24] The Governor emphasized that "no one should be denied appropriate prescription drugs based on the personal, religious or moral objection of individual pharmacists."[25] In a related press conference, the Governor stated that she could remove the entire Board with the legislature's consent but "she would prefer not to take such a drastic step."[26] The Governor then sent the Board a different draft rule, which required all pharmacies to dispense all lawfully prescribed drugs, and prevented pharmacists from refusing to dispense for religious reasons.[27] Perhaps not surprisingly, the Governor's statement prompted a change of heart, and the Board voted unanimously in favor of the Governor's proposal.[28] The Board then enacted a rule requiring pharmacies to dispense all prescribed medications, regardless of religious objections, and essentially requiring double staffing if any pharmacy wished to accommodate a pharmacist's religious objections.[29]

Shortly thereafter, the plaintiff pharmacy and pharmacists filed suit, arguing that the rule violated, *inter alia*, the federal Free Exercise Clause.[30] The pharmacy, asserting the free exercise rights of its owners, alleged that forcing it to sell Plan B would force its owners to choose between violating the law and violating their religion.[31] The pharmacists alleged that they were forced to take lower paying, less desirable jobs as a result of the State's requirement.[32]

The district court found that, although the law purported to apply to refusals to sell any drug for any reason—i.e., it appeared to be generally applicable—"[f]rom the very beginning of this issue, the focus of the debate [was] on Plan B and on religious objection to dispensing that drug."[33] The court found that the "overriding objective of the subject regulations was, to the degree possible, to eliminate moral and religious objections from the business of dispensing medication."[34] Under *Lukumi*, the court found that plaintiffs were likely to succeed on the merits of their Free Exercise challenge, and entered a preliminary injunction against enforcement of the Rule.[35]

Defendants appealed and last month the Ninth Circuit reversed. The majority held that, in deciding whether a law is subject to free exercise analysis, it would not consider "the legislative history of the law—its historical background, the events leading up to its adoption, and its legislative or administrative history, including contemporaneous statements made by members of the decisionmaking body."[36] Instead, looking solely at the text of the rule, the court concluded that it was neutral and generally applicable, because it applied to refusals to sell drugs other than Plan B, and because it would restrict even refusals to sell Plan B for reasons other than religious objection.[37] Accordingly, despite evidence that the entire point of the rule was to address refusals to sell Plan B on religious grounds, the Court found that the Free Exercise Clause did not even apply under *Smith*.[38]

The Ninth Circuit's decision is surely wrong on the issue of whether the history of the law can be considered in the *Smith* analysis. There is some agreement among the circuit courts that the *Smith*/*Lukumi* analysis requires consideration of these facts.[39] Indeed, the Supreme Court itself appears to consider such facts both in the free exercise context and in its Establishment Clause cases.[40]

The Ninth Circuit's refusal to look at the context of the law forced it to miss what should have been an obvious example of targeting. There seems to be little doubt that the focus of the entire rulemaking process was on religious objections to selling Plan B and how to solve the alleged problems caused by such religious objections.[41] Indeed, the very press release announcing the rules acknowledged they were "sparked by complaints that some pharmacists and pharmacies refused to fill prescriptions for emergency contraceptives—also known as morning after pills or Plan B."[42]

Nevertheless, the Ninth Circuit permitted the rule to completely avoid free exercise analysis because the rule was generally applicable.

Had the Ninth Circuit considered the context of the rule, it would have seen clear evidence that the law targeted religious conduct, and that its "general applicability" was extremely dubious. In fact, over the twelve years preceding the rule, there was no evidence of any problem with refusals to sell any drug other than Plan B.[43] Thus the rulemakers essentially banned something—refusals to sell drugs other than Plan B—that rarely or never happened. And in exchange for regulating something that rarely or never happened, the lawmakers were able to directly target and outlaw the religious behavior to which they objected—religious refusals to sell Plan B—and avoid free exercise analysis entirely.

IV. *Smith* and the Proper Analysis of Targeted Laws

As the *Stormans* case demonstrates, an over-reading of the *Smith* decision can lead courts to fail to even apply the Free Exercise Clause in the cases for which it is most appropriate, namely, when lawmakers deliberately set out to prohibit religious conduct. A proper reading of *Smith*, however, requires application of strict scrutiny in such cases.

First, the shorthand version of *Smith* commonly used by courts—that "neutral and generally applicable" laws are immune from free exercise attack—is incorrect, or at least incomplete. *Smith* involved an Oregon law prohibiting possession of certain controlled substances, including peyote.[44] The plaintiffs were dismissed from their jobs at a private drug rehabilitation center when they admitted to having used peyote in sacramental ceremonies at their Native American Church.[45] Plaintiffs were then denied unemployment benefits because they were found to be ineligible for having engaged in work-related misconduct for violating the controlled substance law.[46]

Oregon's controlled substances law was clearly not targeted at religious conduct. When enacted, the law had absolutely nothing to do with religion or religious objectors, but was presumably focused on general public safety. As the Court made clear, it was not "specifically directed" at the plaintiffs.[47] Moreover, the law's impact in outlawing certain religious conduct was not its central purpose. The legislature did not first determine that the religious use of peyote was causing problems and then pass a general law to address that problem. Rather the restriction on religious use of peyote was "merely the incidental effect of a generally applicable and otherwise valid provision."[48]

Thus, the *Smith* opinion should not be read to lower the level of scrutiny for laws that target a particular religious conduct. Unlike the law at issue in *Smith*, a targeted law would be "specifically directed" at religious conduct, and would restrict that conduct as the direct or purposeful effect of the law, rather than an "incidental" one.[49] This is true in the rare case of a legislature that acts out of true religious animus, as well as in the more common situation in which the lawmakers simply wish to eliminate the religious conduct for some reason that has nothing to do with religion, as in our hypothetical Saturday openings law or the Sunday morning alcohol ban. Accordingly, *Smith* should not be read to excuse such targeted laws from free exercise analysis. To the contrary, *Smith*'s emphasis on the lack of targeting, and on the *incidental* nature of the burden to a wholly unrelated purpose, makes clear that targeted laws should still receive strict scrutiny.

Even if *Smith* had left open the question of targeted but generally applicable laws, common sense requires that such laws be subject to scrutiny under the Free Exercise Clause.

In a pluralistic society, "free exercise"—i.e., that which the Free Exercise Clause is intended to protect—necessarily means different things to different people. While a Catholic's free exercise may include attendance at Sunday Mass and consumption of sacramental wine, a Santerian's may include animal sacrifice, a Native American's may include use of peyote, and a Jew's may include observing a Saturday Sabbath. In this respect, the free exercise of religion is actually particularly *susceptible* to targeting through laws of general applicability—members of other religions will be largely unaffected by a "general" ban on animal sacrifice or consumption of wine before noon on Sundays. Thus where lawmakers are clearly targeting particular religious conduct, they should not escape all First Amendment scrutiny simply by using a generally applicable rule to do it.

A law that deliberately targets religious conduct should also be subject to scrutiny because, by definition, the asserted government interests were presumably not strong enough to have previously resulted in a general law. In *Smith*, for example, one cannot doubt the legitimacy of the state's asserted interest in regulating hallucinogenic drugs. The state had asserted that interest, and passed laws to further that interest, having nothing to do with the plaintiffs' religious use of peyote. *Smith* may well have had a different outcome if that law had been "specifically directed" at stopping Native American religious use of peyote as its direct and not "incidental" effect, even if the legislature had been willing to ban non-religious uses in the process as well.

Likewise, in *Stormans*, despite the well-known facts that no pharmacy can stock all drugs, and that pharmacies from time to time will not be able to fill a particular prescription, Washington did not seek to mandate filling of prescriptions until it was responding to religious refusals to dispense Plan B. Courts should heavily scrutinize the arguments of lawmakers who had never before regulated a particular behavior, but who, when presented with religious exercise of that behavior, suddenly assert that neutral reasons are sufficient to foreclose free exercise. Not all such cases will lead to invalidation—there may be compelling interests at stake that, for some reason, were never previously recognized—but it makes no sense to say that free exercise analysis should not even be applied.

Finally, strict scrutiny is appropriate precisely because in many targeting cases the legislature will not be acting out of anti-religious animus, but will simply be engaging in ordinary political balancing. The fundamental point of the Free Exercise Clause is that religion is not supposed to be merely equivalent to all other interests that a lawmaking body balances, but merits special protection. If and when a lawmaking body considers religiously motivated conduct and then decides that other factors are of sufficient importance to outlaw that conduct, it would seem particularly odd for the Free Exercise Clause to be totally inapplicable simply because the end result of that balancing process is a generally applicable law.

V. Conclusion

One negative effect of the *Smith* decision is that lawmakers (and some courts) read that decision to expressly authorize the deliberate targeting and criminalization of religiously motivated conduct, so long as the targeting is achieved by a facially neutral and generally applicable law, and as long as the legislature can offer neutral reasons for opposing the conduct. In this sense, *Smith* has created the roadmap for legislatures that wish to prohibit a particular religious activity and avoid constitutional scrutiny. As the *Stormans* case indicates, when given a roadmap on how to achieve policy goals while avoiding constitutional scrutiny, at least some lawmakers will follow the map.[50]

The *Smith* decision has long been subject to intense criticism, with critics arguing that it left the Free Exercise Clause essentially a dead letter.[51] Even if the Court is loath to reconsider *Smith*, the *Stormans* case presents an ideal opportunity for the Court to confirm that *Smith* is not a roadmap for lawmakers to regulate religious conduct. Rather, the Court can make clear that *Smith* was simply addressing how courts should analyze pre-existing, generally applicable laws that were not enacted to prohibit religious conduct, and that laws that specifically target religious conduct—regardless of the motives of the legislature—are still within reach of the Free Exercise Clause.

Endnotes

1 494 U.S. 872 (1990).

2 Church of the Lukumi Babalu Aye, Inc. v. City of Hialeah, 508 U.S. 520, 535 (1993).

3 *Id.* at 527-28.

4 *Id.* at 546-47.

5 *Id.* at 564 ("In being so readily susceptible to resolution by applying the Free Exercise Clause's 'fundamental nonpersecution principle' this is far from a representative free-exercise case.") (Souter, J., concurring).

6 *Id.* at 535 ("Apart from the text, the effect of a law in its real operation is strong evidence of its object.").

7 Employment Division v. Smith, 494 U.S. 872, 879 (1990) ("Subsequent decisions have consistently held that the right of free exercise does not relieve an individual of the obligation to comply with a 'valid and neutral law of general applicability on the ground that the law proscribes (or prescribes) conduct that his religion prescribes (or proscribes).'"(citation omitted)); *id.* at 881 ("The only decisions in which we have held that the First Amendment bars application of a neutral, generally applicable law to religiously motivated action have involved not the Free Exercise Clause alone, but the Free Exercise Clause in conjunction with other constitutional protections, such as freedom of speech and of the press.").

8 As a result, plaintiffs challenging government-imposed burdens on religion over the past two decades have largely looked to sources beyond the First Amendment for their strongest protection. In response to *Smith*, Congress, state legislatures, and state courts acted to provide greater protections for religion. Congress enacted the Religious Freedom Restoration Act ("RFRA"), 42 U.S.C. §§ 2000bb-2000bb-4 (2006), which, though held invalid as applied to state actions, Boerne v. Flores, 521 U.S. 507, 507 (1997), still provides for strict scrutiny (i.e., a law must be narrowly tailored to serve a compelling government interest) as the standard when the federal government imposes substantial burdens on an individual's religion, Gonzales v. O Centro Espirita Beneficente Uniao de Vegetal, 546 U.S. 418, 424 (2006). Many state legislatures enacted state RFRA laws to provide similar protections from burdens on religion imposed by state and local governments. *See, e.g.,* Fla. Stat. §§761.01-761.05 (2008); 775 Ill. Comp. Stat 35/1-35/99 (2008); Tex. Civ. Prac. & Rem. Code Ann. §§ 110.001-110.012 (Vernon 2008). Many state courts also rejected *Smith* as a matter of state constitutional law, preferring instead to retain the compelling interest test as previously articulated in *Wisconsin v. Yoder* and *Sherbert v. Verner*. *See, e.g.,* Attorney General v. Desilets, 636 N.E.2d 233, 235-36 (Mass. 1994).

9 *See, e.g.*, Blackhawk v. Pennsylvania, 381 F.3d 202 (3rd Cir. 2002) (finding law failed general applicability test); Midrash Sephardi, Inc. v. Town of Surfside

366 F.3d 1214, 1232-33 (11th Cir. 2004) ("a closer look at § 90-152 reveals that Surfside improperly targeted religious assemblies and violated Free Exercise requirements of neutrality and general applicability."); *id.* at 1234 ("Under Lukumi, it is unnecessary to identify an invidious intent in enacting a law-only Justices Kennedy and Stevens attached significance to evidence of the lawmakers' subjective motivation.").

10 494 U.S. at 874.

11 *See, e.g.,* Lighthouse Institute for Evangelism, Inc. v. City of Long Branch 510 F.3d 253, 275 (3rd Cir. 2007) ("The Plan is clearly neutral; there is no evidence that it was developed with the aim of infringing on religious practices"); Grace United Methodist Church v. City of Cheyenne, 451 F.3d 643, 649-50 (10th Cir. 2006) ("A law is neutral so long as its object is something other than the infringement or restriction of religious practices.... In the instant case, there is no dispute that the zoning ordinance at issue is neutral on its face.").

12 *Lukumi*, 508 U.S. at 558 (Scalia, J., concurring in part and concurring in the judgment); *see*, *e.g.*, Vision Church v. Village of Long Grove 226 F.R.D. 323, 326 (N.D. Ill. 2005) ("For the above stated reasons, the court takes as its starting point the principle that it ought not, generally speaking, facilitate inquiries into the motivations of legislators." (referring to *Lukumi*)).

13 *See, e.g.,* San Jose Christian College v. City of Morgan Hill, 360 F.3d 1024, 1031-32 (9th Cir. 2004); Miller v. Reed, 176 F.3d 1202, 1206-08 (9th Cir. 1999).

14 *See*, *e.g.*, Locke v. Davey, 540 U.S. 712, 725 (2004) ("In short, we find neither in the history or text of Article I, § 11, of the Washington Constitution, nor in the operation of the Promise Scholarship Program, anything that suggests animus toward religion."); *Lukumi*, 508 U.S. at 540 ("Relevant evidence includes, among other things, the historical background of the decision under challenge, the specific series of events leading to the enactment or official policy in question, and the legislative or administrative history, including contemporaneous statements made by members of the decisionmaking body.") (opinion of Kennedy, J, and Stevens, J.); Cottonwood Christian Center v. Cypress Redevelopment Agency, 218 F.Supp.2d 1203, 1225 (C.D. Cal. 2002) ("Here, there is significant circumstantial evidence of a discriminatory intent.").

15 *See*, *e.g.*, Olsen v. Mukasey, 541 F.3d 827, 832 (8th Cir. 2008) ("Absent evidence of an 'intent to regulate religious worship,' a law is a neutral law of general applicability" (citing Cornerstone Bible Church v. City of Hastings, 948 F.2d 464, 472 (8th Cir.1991).).

16 571 F.3d 960 (9th Cir. 2009).

17 *Id.* at 964-65.

18 Stormans, Inc. v. Selecky, 524 F. Supp. 2d 1245, 1248 (W.D. Wash. 2007).

19 *Stormans*, 571 F.3d at 968; *see also http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm109795.htm* ("If fertilization does occur, Plan B may prevent a fertilized egg from attaching to the womb (implantation)"); *http://www.planbonestep.com/plan-b-prescribers/how-plan-b-works.aspx* ("Plan B® One-Step works... by: ... [a]ltering the endometrium, which may inhibit implantation.").

20 *Stormans*, 524 F. Supp. at 1248-49.

21 *Id.* at 1250.

22 *Id.*

23 *Id.* at 1251.

24 *Id.*

25 *Id.*

26 *Id.*

27 *Id.*

28 *Id.*

29 *Stormans*, 571 F.3d at 966-67.

30 *Id.* at 967.

31 *Id.*

32 *Id.*

33 *Stormans*, 524 F. Supp. 2d at 1260.

34 *Id.* at 1259.

35 *Id.* at 1266.

36 *Stormans*, 571 F.3d at 982.

37 *Id.* at 983-84.

38 *Id.* at 987.

39 *See*, *e.g.*, Eulitt *ex rel*. Eulitt v. Maine, Dept. of Educ., 386 F.3d 344, 356 (1st Cir. 2004) ("[T]he legislative history clearly indicates Maine's reasons for excluding religious schools from education plans."); Children's Healthcare Is a Legal Duty, Inc. v. Min De Parle, 212 F.3d 1084, 1090 (8th Cir. 2000) ("To determine whether strict scrutiny applies, a court looks at a law's text, legislative history, and real operation" (citing *Lukumi*)); *see also* Midrash Sephardi, Inc. v. Town of Surfside, 366 F.3d 1214, 1226 (11th Cir. 2004) ("Although the legislative history of a statute is relevant to the process of statutory interpretation, we do not resort to legislative history to cloud a statutory text that is clear."). *But see* St. John's United Church of Christ v. City of Chicago, 502 F.3d 616, 633 (7th Cir. 2007) ("[W]e do not run the risk of selective use of statements in legislative history that might not reflect the intent of the legislature.").

40 *See*, *e.g.*, McCreary County v. American Civil Liberties Union of Ky., 545 U.S. 844, 862 (2005) (Establishment Clause); Locke v. Davey, 540 U.S. 712, 725 (2004) ("[W]e find neither in the history or text of Article I, § 11, of the Washington Constitution, nor in the operation of the Promise Scholarship Program, anything that suggests animus toward religion."); Santa Fe Ind. Sch. Dist. v. Doe, 530 U.S. 290, 308 (2000) (Establishment Clause).

41 *See*, *e.g.*, Stormans, Inc. v. Selecky, 524 F. Supp. 2d 1245, 1259 (W.D. Wash. 2007) (quoting letter from Washington State Human Rights Council stating that "the drug at the center of this issue is Plan B.").

42 *Id.*

43 *Id.* at 1261 ("A review of complaints referred to the Board from 1995 to 2007 does not indicate a problem with access to HIV-related medications, or any other medications for that matter.").

44 Employment Division v. Smith, 494 U.S. 872, 874 (1990).

45 *Id.*

46 *Id.*

47 *Id.* at 878 ("They contend that their religious motivation for using peyote places them beyond the reach of a criminal law that is not specifically directed at their religious practice.").

48 *Id.*

49 *See*, *e.g.*, McTernan v. City Of York, 564 F.3d 636, 647 (3rd Cir. 2009) (citing *Smith* and *Lukumi*); Cornerstone Christian Schools v. University Interscholastic League, 563 F.3d 127, 135 (5th Cir. 2009) (citing *Smith* and *Lukumi*); Colorado Christian University v. Weaver, 534 F.3d 1245, 1257 (10th Cir. 2008) (citing *Smith*).

50 A similar problem occurs in the free speech context, where cases such as *Ward v. Rock Against Racism*, 491 U.S. 781, 791 (1989), and *Clark v. Community for Creative Non-Violence*, 468 U.S. 288 (1984), established that legislatures could abridge free speech rights so long as they did so through content-neutral time, place, and manner regulations. Given this roadmap, legislatures have shown considerable creativity in regulating very particular speech through time, place, and manner regulations. The city of Seattle, for example, wished to regulate protesters at the 1999 World Trade Organization conference. Rather than proceed by injunction (which would have faced heightened scrutiny under *Madsen v. Women's Health Ctr., Inc.,* 521 U.S. 753 (1994)) or by a law that specifically identified the WTO protests as the problem, Seattle instead proceeded by a mayoral emergency order that outlawed all speakers for the exact time and location of the WTO conference. Because, in theory, protesters might have shown up at that exact time and location to talk about other issues, the order was treated as neutral by the courts. *See* Menotti v. City of Seattle, 409 F.3d 1113 (9th Cir. 2005).

Even more egregiously, Massachusetts has used the "time, place, and manner" test to create a law that obviously targets abortion protesters without,

so far, facing strict scrutiny. Massachusetts has enacted a statute banning speech only (a) within 35 feet of an abortion clinic, and (b) while the clinic is open. Because these restrictions technically relate to the "place" and the "time" of the speech and because in theory someone could show up at an abortion clinic to protest about issues other than abortion, courts have, quite remarkably, deemed the law neutral. *See* McCullen v. Coakley, 571 F.3d 167 (1st Cir. 2009).

As in the Free Exercise context, these legislatures are taking a test announced by the Court when it was evaluating a non-targeted law—i.e., a law that was passed with no intention of regulating a specific group of speakers or religious observers—and using that test as a roadmap for how to target disfavored speakers or conduct while avoiding constitutional scrutiny.

51 Church of the Lukumi Babalu Aye, Inc. v. City of Hialeah, 508 U.S. 520, 577 (1993) (Souter, J., concurring) ("'Neutral, generally applicable' laws, drafted as they are from the perspective of the non-adherent, have the unavoidable potential of putting the believer to a choice between God and government."); Michael W. McConnell, *Free Exercise Revisionism and the* Smith *Decision*, 57 U. Chi. L. Rev. 1109, 1153 (1990) ("Religious exercise is no longer to be treated as a preferred freedom; so long as it is treated no worse than commercial or other secular activity, religion can ask no more.").

Book Reviews

Federalism: Political Identity and Tragic Compromise

By Malcolm M. Feeley and Edward Rubin

*Reviewed by George D. Brown**

In *Federalism: Political Identity and Tragic Compromise*, Professor Malcolm Feeley and Dean Edward Rubin continue their assault on American federalism and its defenders. Readers of *Engage* will be familiar with many of the themes, which were set forth in the authors' widely cited earlier article in the UCLA Law Review: "Federalism: Some Notes on a National Neurosis." Their book—a slim (153 pages of text and 47 pages of footnotes), but formidable volume—updates their arguments, takes them a step further, and makes an important contribution to the literature of federalism.

The book is divided into two parts: the first develops a general theory of federalism, particularly its preconditions and its utility. The second part applies this theory to the United States. The authors define federalism as a governmental system "that grants federal autonomy to geographical subdivisions or subunits." The subunits must have a reserved domain and the power to assert their jurisdictional rights against the central government. Federalism should be distinguished from other organizational forms such as consociation, decentralization, and local democracy. The key variables within any nation that lead to adoption of a federal system are geography and sharp normative variations among the populations of the subunits.

> The basic reason that nations adopt a federal system or maintain a federal regime that was adopted in a prior era… is to resolve conflicts among citizens that arise from the disjunction between their geographically based sense of political identity and the actual or potential geographical organization of their polity.

Federalism can be useful in expanding "the range of psychopolitical resources available for the creation of a political regime," but it is best viewed as "an alternative to dissolution, civil war, or other manifestations of a basic unwillingness of the people in some geographic area within the nation to live under the central government."

This is a core point. Feeley and Rubin see federalism as a sub-optimal compromise designed for only a few situations. They drive home the point by utilizing such adjectives as "tragic" and "grim" to describe federal solutions. For them, uniformity is good, and differences are bad, at least when it comes to a nation's ability to achieve the optimal result of a unified political identity. For the authors, wisdom resides at the center. Thus, real citizen participation could be more directly achieved through the national government's "hiring community organizers," and "funding local organizations."

These observations about participation are part of a key step in Feeley and Rubin's overall argument: federalism has no independent value as an organizing principle for nations, other than "grim" ones like avoiding civil war. Thus, they are compelled to knock down such staples of federalism justification as the values of interjurisdictional competition and the role of experimentation, most frequently evoked in Justice Brandeis' laboratory metaphor. For example, Feeley and Rubin's response to the latter justification is largely to deny that experimentation happens, and to insist that when it does happen the phenomenon could just as easily be the "happy incident" of managerial decentralization. Moreover, effective experimentation requires goal-setting by the central authority. "[C]entralization is necessary not only to initiate the experimental process but also to implement the results of that process in any reasonably effective fashion."

Fundamental to the authors' dismissal of experimentation as a value of federalism is the view that in a nation norms come from the center. "Normative variation" is to be avoided. The notion that quasi-independent polities acting through their own political processes might contribute to the dialogue over what basic values are seems impossible. At this point they trot out the example of slavery to show what experimentation can lead to.

Besides, states don't experiment anyway according to Feeley and Rubin. Drawing on economic theory, they conclude that "individual subunits will have no incentive to invest in experiments that involve any substantive or political risk; they will instead prefer to be free riders and wait for other subunits to generate them. This will, of course, produce relatively few experiments." Indeed, the authors suggest that states "must be forced or encouraged to [experiment] by the central authority."

As this review is being written, the national political process is involved in a major debate over health care. Prominent in that debate is an assessment of the results of the Massachusetts Health Care Initiative. The New York Times recently engaged in an extensive review of "The Massachusetts Model" and its central role in the national debate. The Times examined the state program's "growing pains and glitches," in particular, their financial consequences. For purposes of the federalism debate, the point is not whether the plan "works," but the fact that the state was willing to undertake it. State political actors saw political gain in undertaking an experiment of the sort Feeley and Rubin declare does not happen. The state political processes worked—whether or not the plan does—in large part because they are viable and meaningful.

Clearly, one's evaluation of Feeley and Rubin's negative view of federalism in general depends heavily on whether one accepts their view of it as an organizing principle of limited value, valid in only a few extreme situations. However, as the discussion of experimentation suggests, their real target is American federalism. Independent justifications, such as the laboratory thesis, largely stand or fall based on how they play out in the United States. Indeed, a fair amount of discussion

* *George D. Brown is the Robert F. Drinan, S.J. Professor of Law at Boston College Law School, where he specializes in National Security Law. His latest article is "Counter-Counter-Terrorism Via Lawsuit"—The Bivens Impasse, 82 Southern California Law Review 841 (2009).*

of the American experience takes place in the first part of the book. For example, several pages are devoted to critiquing Justice O'Connor's opinion in *Gregory v. Ashcroft*.

By the end of the first part, the reader will have probably made up his or her mind both on federalism in general and on American federalism in particular. Thus, I will focus briefly on only three aspects of the second half: how federalism developed in the United States; whether it does or should play a role as a contemporary organizing principle; and why there are those who defend it. As for why we have federalism, the authors view particular documents, e.g., the Constitution, as less important than "basic questions of political identity." Thus, "the Constitution should be regarded not as a definite determination of the relationship between the national government and the states but, rather, as one event, albeit an important one, in the four-hundred year evolution of political identity among a group of people whose outer boundaries had been autocratically defined."

Looking both at the key documents and at questions of political identity, the authors find ambiguity in some of the former—particularly the Constitution—and a growing sense of a national identity in the latter area. The main issue which kept federalism alive was slavery. "The slave states could only protect themselves through slavery…" Indeed, "[b]y the 1850s, this was the only function federalism served."

What is federalism's role today? Rubin and Feeley insist that it has none. Pre-existing centripetal, nationalizing forces have accelerated sharply since the Civil War—itself "a defeat for the federalist conception of the United States"—and twentieth century phenomena such as the New Deal. There are no profound regional attitudinal features that would make federalism attractive. "The United States, despite its size, its ethnic diversity, and its self-image as a vast and variegated nation, is in fact a heavily homogenized culture with high levels of normative consensus." The absence of the criteria that, occasionally, justify a nation's recourse to a federalist structure means that "federalism no longer serves any purpose in the United States." Following this logic to its conclusion, Feeley and Rubin declare flatly that "the United States is no longer a federal regime."

The last statement is not literally true, of course. The Constitution still contains the Tenth Amendment—suggesting that some powers are reserved to the states—and the list of enumerated national powers—suggesting that some powers are not granted. The authors offer a partial justification for the statement based on the evolution of the system as reflected in Supreme Court doctrine. Over the years, the Court has permitted the enumerated powers—particularly the Commerce Clause—to develop into the equivalent of a national police power. Cases like *United States v. Lopez* and *United States v. Morrison* are outliers, decisions possible only because a national consensus on the underlying normative issues had not yet formed.

One can, of course, defend *Lopez* and *Morrison*, and the constitutional vision they represent, on the ground of original intent or by citing the obvious fact that no constitutional amendment has declared the end of the federal system or the conferral of a general police power on the national government. It is important, however, to understand the crucial role that the first, theoretical part of the book plays in bringing the authors to this point. Suppose they are right that federalism is a suboptimal organizational compromise that is valuable in a limited number of situations, and that those situations long ago ceased to exist in the United States. Perhaps American federalism would become obsolete, existing on paper but not anywhere else, including Supreme Court decisions. That is why it is important for those who disagree with the claims in the second half of the book to take issue with those in the first half—to defend, for example, the concept of experimentation as one that takes on particular importance in sub-national polities where the government has real power and where citizens perceive its processes as worth participating in. It is essential to argue that federalism furthers important values because of the status it grants to sub-national units, and that it merits a broader role than the highly limited one to which Feeley and Rubin consign it—a suboptimal compromise "to resolve conflicts among citizens that arise from the disjunction between their geographically based sense of political identity and the actual or potential geographic organization of their polity."

Since the authors reject any broader role for federalism—and notions that it advances a range of values—they are forced to come to grips with the third question: why does anyone defend it? At first blush they seem to offer the vision of benighted, albeit benign, individuals who act based on "nostalgia-drive sentiments, the bromides of high school civics, and conceptual confusion." The theme of nostalgia is repeated—the current Supreme Court is described as "particularly nostalgia-driven"—but a darker explanation is easy to discern. As noted, the authors contend that by the 1850s, preserving slavery "was the only function federalism served." Despite the Civil War and emancipation, many Southern states preserved a high degree of segregation which Feeley and Rubin label "apartheid." The authors hammer home the point that with formal slavery gone, "states rights meant, in essence, the right of Southern states to preserve apartheid." Thus, in a slight modification, or updating, of their earlier statement, the authors identify slavery and apartheid as "the only rationales for federalism for the past 150 years or so."

This sounds more than a bit like labeling those who support federalism as racists, especially since the analysis on the next page turns to island territories and suggests that they are kept in a form of federalism because they have "overwhelmingly nonwhite populations." In fact, "[t]he motivation for this continued reliance on federalism… is the same as the motivation for the continued reliance on federalism within the United States in the years preceding World War II—namely, racism."

Perhaps this is just an example of what the book jacket calls "bold argument… certain to provoke controversy." Still, Feeley and Rubin do exhibit a tendency to denigrate those who disagree with them. Shortly after a slash-and-burn analysis of Justice O'Connor's opinion in *Gregory v. Ashcroft*—for example, she advances "pseudoarguments"—they state that "[i]t seems difficult to imagine that any American, even a Supreme Court Justice, is so parochial as to be unaware that such unitary regimes as England, Sweden, Denmark, and the Netherlands have met the highest standards of political participation and human

rights protection." One doubts that the gratuitous reference to Supreme Court Justices encompasses Ruth Bader Ginsberg or Stephen Breyer.

Feeley and Rubin also display a singularly off-putting pretentiousness, seeking, it would seem, to beat the reader into a submissive acknowledgement that people who know so much must be right. Thus, by the bottom of page ten, the following authors and thinkers have been cited or invoked (here given in order of appearance, by last name, unless otherwise indicated): Eleazar, Riker, McKay, Etzioni, Sandel, Dryzek, J. Cohen, Habermas, Lijphart, Dahl, M. Weber, Arendt, Schutz, Siddens, Touraine, Rawls, Descartes, Locke, Kant, Husserl, Heidegger, Hegel, A. Cohen, Saint Augustine, Anderson, E. Weber, Miller, Oommen, and Smith. In one paragraph (!), the authors draw lessons from ancient Athens, Norman Sicily, the second-century Roman Empire, the early Tang dynasty, the Umayyd caliphate, the Carolingian Empire, and "premodern empires—such as the Abbasid and Ottoman in the Middle East, the Mauryan and Gupta in India, and the Nara-Heian in Japan."

Jargon rears its head, almost to the point of parody. We learn that "[i]dentity can be understood as the self's interpretation of itself. This would be true for the Cartesian, Kantian, Husserlian, and Heideggerian self, although it would have different ontological significance in each case." Things are more complicated, however. Some philosophers "urge that the self develop an identity as an independent, morally responsible agent." Others "argue that this is impossible in the ordinary course of life, where socially constructed conceptions of identity prevail, conceptions that can only be escaped if the self sheds its identity through either a transcendental *epoché* or a reconnection with the essence of *Dassein*."

All in all, the book's mixture of condescension and pretension can be annoying at times, but should not deter the reader from exploring the arguments against American federalism. Feeley and Rubin have made an important contribution to the dialogue about it. The viewpoint they represent is not about to go away. Neither is federalism.

Regulation by Litigation

By Andrew P. Morriss, Bruce Yandle, and Andrew Dorchak

*Reviewed by Margaret A. Little**

Regulation by Litigation, an innovation in American political theory, has descended upon the American polity, hardly noticed by its citizenry and arguably even less understood by its elected political representatives, the mainstream press, and most legal or political analysts.

What a fascinating story it is, this business of regulation by litigation, using litigation and the courts to achieve and enforce regulatory regimes against entire industries without having to go through the expense, uncertainty, or trouble of securing legislative or rule-making authority for such regulation. And a business it most certainly is—when wielded by private lawyers, it is the most lucrative new field of practice in the legal market purchasable by a law license and friends in high places.

Three scholars, Andrew Morriss, Bruce Yandle, and Andrew Dorchak, have undertaken a painstaking dissection of regulation by litigation by examining three case studies—1.) the EPA's 1998 suit against heavy-duty diesel engine manufacturers, 2.) asbestos and silica dust private mass tort litigation, and 3.) state and private sponsored lawsuits against the tobacco industry.

The book begins with a comprehensive discussion of the academic legal and economic theories and constructs underlying the regulation by litigation approach such as *public choice theory*—the use of economic analysis to explain political decisions—and its unfailing dark companion, *rational ignorance*—to assist the reader in understanding both the origin of this species of regulation, its taxonomy, and its surprising ability to transcend legal and constitutional prohibitions, to say nothing of public outcry. Though a bit of a slog for the general reader, the walk through the theoretical constructs—*public interest theory, capture theory/rent-seeking, special interest theory, political wealth extraction*, and the delightfully and quite accurately named *bootleggers-and-Baptists theory*—is well worth it to equip an informed citizen with the tools to understand how such a lucrative and often lawless phenomenon could arise and flourish. But the devil, as always, is in the details. The empirical case studies shorn of theory best illustrate the dark matter that makes up this constitutionally and legally flawed model of regulation. A brief synopsis of the facts of each case study follows to assist in enlightening the reader—and the public.

* *Margaret A. Little is a Stratford, Connecticut attorney in the private practice of commercial litigation and appeals in state and federal court. A graduate of Yale College and Yale Law School, Ms. Little clerked for the Hon. Ralph K. Winter, United States Court of Appeals for the Second Circuit.*

Heavy Duty Diesel Engine Litigation

In 1998, despite a full array of laws and regulations already in place that regulate the manufacture of heavy duty diesel engines, the EPA sued all U.S. heavy-duty diesel engine manufacturers, alleging that they illegally used electronic engine controllers as "defeat devices" to frustrate existing emission standards. This suit was brought against a background of Clean Air Act Amendments under which the EPA used negotiated rulemaking to create "non-compliance penalties" in 1985. That arrangement permitted the sale of engines that failed to meet emissions targets and, under a regulatory deal negotiated by the EPA, the California Air Resources Board ("CARB"), and the engine manufacturers, also set future regulatory and emission-reduction targets through model year 2004. The EPA also tacitly knew of and permitted the use of the controllers which became the subject of a government lawsuit. By means of this 1998 suit, settled just five months later in October 1998, the EPA obtained regulatory concessions from the manufacturers, including an agreement to "pull forward" the model year 2004 standards to October 2002 and the payment of hundreds of millions of dollars in fines. The EPA then returned to traditional notice-and-comment rule making for emission standards for future model years.

Two features of this unprecedented litigation by the EPA draw particular note from the authors. The pull-forward provisions permitted the EPA to impose the 2004 standards two years earlier than they could have done through rulemaking because of rule-making's lead-time requirements. Second, the EPA did not require that the engines meet U.S. emissions standards at all points during operation, instead requiring that they meet the less stringent European standards. A third factor, state regulatory action, also came into play when CARB imposed even more stringent standards, causing the Engine Manufacturers' Association to bring suit in the Sacramento Superior Court. That court held that the more stringent state regulations were "unlawful, unconstitutional, void, invalid and beyond the scope of [CARB's] authority," a finding arguably equally applicable to the federal regulatory suit.

The authors show how the concentration of diesel engine manufacture in just four companies and the lack of vertical integration, among other factors, made the industry vulnerable to state and federal regulatory overreach, such as this litigation, followed by a swift settlement negotiated amongst the few players. The authors' research and interviews of EPA staffers reveal that the 1998 suit was motivated by a complex web of political and career advancing strategies, Vice-President Al Gore's presidential campaign, internal regulatory rivalries, EPA litigators' ignorance of the EPA's tacit approval of the engine controllers, and genuine outrage of the litigators and top policymakers at the earlier failure to secure compliance with the spirit of the regulations over the years. Litigation solved all of these problems for the EPA, permitting it to advance the 2004 standards by two years, avoid notice and comment delays and industry challenges, and lock in these regulatory changes at the end of an administration through settlements. Such settlements are considered politically untouchable, unlike regulatory tightening at the end of an administration, which can be and often is reversed by the new administration.

Unlike rule-making adopted through public processes where all concerned can be heard, the regulation by litigation model adopted by the EPA conferred benefits on environmental groups and state regulators, imposed the costs amongst all consumers (given that transportation affects the price of most products), and insulated the EPA's settlement-negotiated standards from changes in the executive branch. It also led to a boom and bust cycle that "devastated the diesel engine manufacturers during the last quarter of 2002 and the first quarter of 2003" and under which over half a million new diesel trucks with model year 2004 standards were added to the roads from 2005 to 2006. This addition of trucks worked against improvement of air standards, which the EPA is supposed to be working with the industry to promote. (Lawful rule making is not exempt from these boom and bust effects. The authors note that a boom and bust cycle also accompanied the tightened 2007 standards, but at least that process observed due process for the concerned parties.)

Because the settlement applied only to domestic manufacturers, foreign producers could and did increase sales in the U.S. market, and because purchasers in the domestic markets could and did prebuy in large quantities in 2002, the EPA's 1998 litigation and settlement actually significantly increased the population of dirty engines on American roads. For a period of time, air quality deteriorated and the cost of moving goods increased, the latter due both to the pre-buy market distortions and the massive fines imposed by the EPA. The authors conclude that on a net basis, the episode was highly costly for the US economy, denied due process protections otherwise available to affected parties through rule-making, and conferred no benefit on anyone—including the environment itself—other than to strengthen the already powerful hand of state and federal regulators.

Dust Litigation

Silica

Dust litigation is a term applied to mass tort cases involving both asbestos and silica-related airborne dust in the workplace. The authors provide a comprehensive study of three phases of such litigation—silica cases in the 1930's, asbestos cases after 1973, and modern silica litigation—and thereby reach conclusions about how such mass torts can shape and misshape our legal and economic landscape. They show that the Depression-era cases led to the adoption by the states of a comprehensive worker's compensation regime that had the advantage of creating an important repeat-player with an incentive to prevent such occupational diseases—the insurance companies. While not a perfect system for anyone, since worker's compensation is costly and limitations are evident in both compensation and proof of causation, the system offered a comprehensive solution to address and diminish widespread occupational disease while reducing costs by shifting the disputes out of court.

The Asbestos Behemoth

Decades of amassed cases of asbestos litigation have relaxed substantive constraints in tort, procedural, insurance, and bankruptcy law in order to facilitate recovery in these burdensome and never-ending cases, leading to some unintended consequences, such as giving incentives to both the plaintiff's and defendant's bar to cooperate in the perpetuation of such claims, thereby reducing the adversarial role upon which the strength of claims are tested under American law. The procedural and substantive loosening of rules under the staggering load of such claims has led to a species of litigation dominated by a small group of firms with the financial and intellectual capital to finance, develop and settle cases in short order: the asbestos plaintiff's bar. By overwhelming the courts and the defendants, this bar has freed itself from supervision of its fee and settlement practices, lowered costs of entry and proof by plaintiffs, funded a powerful lobby to influence courts and legislatures, put the lawyers in charge of the cases, and converted the process into a repeat player game with few checks on plaintiffs' counsel. The bankruptcy auction aspect of these cases has long been noted. The loss of knowledge and control by the courts is stunningly evidenced by the Maryland Court of Appeals' admission that it could not determine how many claims were before it when hearing the appeal of a mass-consolidated case. Essentially, through their domination of the legal and bankruptcy processes, a small number of plaintiff's firms act as private and lavishly compensated regulators in a costly quasi-administrative compensation system of questionable efficiency and fairness.

The Silicosis Sham

The flaws of the asbestos model were recently dramatically exposed when a medically sophisticated judge, federal judge Janis Jack, was assigned the third wave studied here: the silicosis multi-district litigation cases. Given a rare opportunity to examine the big picture, Judge Jack uncovered a compelling pattern of fraud on the part of the repeat player lawyers and doctors heavily invested in the claims. The authors conclude that the massing up of such claims allows the private plaintiff's bar to exert bet-the-company bankruptcy power that sustains and perpetuates wealth transfers to these private and unaccountable regulators. The plaintiff's bar, unlike public regulators, do not have even an ostensible connection to the public interest and accordingly pursue their economic interests that impose huge costs with no balancing of interests in play.

Silicosis litigation also shows the critical role that ignorance plays in these regulatory schemes. Data on the extent and severity of silicosis are based mostly on estimates and conjecture. Like the breast implant litigation, where billions of dollars passed from bankrupted industries into the hands of the trial bar—and some modestly compensated plaintiffs—the silicosis cases show that when a court's ability to discern good claims from bad and good science from bad is compromised or non-existent, massed-up claims can lead to indefensible wealth transfers. It was only the fortuitous assignment of this third wave of silicosis claims to a smart, scientifically literate, analytical, and pattern-recognizing judge that led to the exposure of systematic fraud by these would-be regulators.

Tobacco Litigation

No more compelling example of the dangers of regulation by litigation can be found than the Master Settlement Agreement ("MSA") entered into in November of 1998 by the state attorneys general and the four major tobacco companies.[1] The MSA created a cartel protecting the big four in exchange for gargantuan off-budget funding of state governments, payments to the private plaintiff's tobacco bar of billions of dollars of fees in perpetuity, permitting the states to impose the largest tax increase in the history of the product without paying the political price of an overtly legislated tax.

The assemblage of interest groups that led to this grim state of affairs includes politically ambitious and politically unaccountable state attorneys general, their cronies in the trial bar, fresh from the infusions of capital generated by the asbestos cases, public health leaders and groups both in and outside of state and federal government, private class action lawyers, the tobacco companies, and eventually state and federal legislators urged to bring this hydra-headed litigation to an end. The state attorneys general took on a constitutionally troubling function of both the legislative and executive branches when they set the equivalent of a tax on future sales and imposed a regulatory regime on all existing and future tobacco producers. By using the vehicle of litigation and mass settlement, these actors have pursued an intentional strategy of obscuring the provisions of the settlement and avoiding public debate on these important questions.

The book painstakingly shows how, by avoiding constraints imposed by constitutions and statutes, "bootleggers and Baptists can manipulate the regulatory environment to their mutual benefit and to the detriment of the public at large." In the case of the MSA, state officials without authority either to levy taxes or regulate health created a national regulatory scheme for cigarette sales that violates state and federal antitrust laws. The authors note the "dubious constitutionality" of the scheme under the federal compact clause, and further make the critical point that the MSA rode roughshod over the checks and balances of state government.

What Next?

The authors suggest a number of important ways in which such legally dubious initiatives could be prevented in the future, while noting that "[n]o global solution suggests itself, and we must fall back on a call for transparency." Regulation by litigation does indeed have a future, a well capitalized one fueled by the financing available from earlier successful litigation regulatory regimes. The book makes a compelling case that the dark and essential twin of public choice theory, rational ignorance, has prevailed far too long. The Supreme Courts of Pennsylvania and California are wrestling with the recurring legal and constitutional infirmities of similar schemes which will continue to recur as long as they are so lavishly rewarded and studiously ignored by the mainstream press and the public at large. The authors call for a "thoughtful conversation" about regulation-by-litigation. It is long overdue.

This book represents a comprehensive and thorough treatment of the subject, strong on conceptual analysis and empirical evidence. The most difficult aspect of bringing this

problem to public attention is that, as the authors note, "the judicial process, while public in name, is private in essence." Regulation by litigation arises when the executive and legislative functions collapse into the hands of a regulator, an organized sector of the bar, or a confederacy of ambitious state attorneys general. The public debate the authors call for needs first and foremost to bring the constitutional imperatives of the separation of powers to bear upon this debate so that we may be a government of laws, not of the very flawed and enormously enriched politicians and attorneys that make up the players in this costly and lawless game. Thoughtful men and women need to remember the fundamental principles of separation of powers and ask how such corrupt and lawless schemes could arise and flourish virtually unnoticed in a free and open society.

Endnotes

1 By way of full disclosure, the reviewer of this book represented Philip Morris in the Connecticut suit brought by its Attorney General against the tobacco companies and has also published prior scholarship critical of the litigation, scholarship that was cited by the authors to this book.

Practice Group Publications Committees

Administrative Law & Regulation
Christian Vergonis, *Chairman*
Daren Bakst, *Vice-Chairman*

Civil Rights
Brian T. Fitzpatrick, *Chairman*

Corporations, Securities & Antitrust
Michael Fransella, *Chairman*
Carrie Darling, Robert T. Miller
& Robert Witwer, *Vice Chairmen*

Criminal Law & Procedure
Ronald J. Rychlak, *Chairman*
William A. Hall, Jr. & Peter Thomson, *Vice Chairmen*

Environmental Law & Property Rights
Donald Kochan, *Chairman*

Federalism & Separation of Powers
Ilya Somin, *Chairman*
Geoffrey W. Hymans, *Vice Chairman*

Financial Services & E-Commerce
Charles M. Miller, *Chairman*

Free Speech & Election Law
Reid Alan Cox, *Chairman*
Francis J. Menton, Jr., *Vice Chairman*

Intellectual Property
Mark Schultz, *Chairman*
Steven M. Tepp & Adam Mossoff, *Vice Chairmen*

International & National Security Law
Nathan Sales, *Chairman*
Soraya Rudofsky, *Vice Chairman*

Labor & Employment Law
James A. Haynes, *Chairman*
Raymond J. LaJeunesse, Jr., *Vice Chairman*

Litigation
Margaret A. Little & Christopher C. Wang
Co-Chairmen

Professional Responsibility & Legal Education
John J. Park, Jr., *Chairman*

Religious Liberties
William Saunders, *Chairman*

Telecommunications & Electronic Media
Raymond L. Gifford, *Chairman*

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

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

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

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