HARVARD
Journal on Legislation

Volume 43, Number 1 **Winter 2006**

Articles

Recent Developments

ARTICLE

THE CONSUMER FRAUD CLASS ACTION: REINING IN ABUSE BY REQUIRING PLAINTIFFS TO ALLEGE RELIANCE AS AN ESSENTIAL ELEMENT

SHEILA B. SCHEUERMAN*

This Article argues that the recent rise in consumer fraud class action lawsuits is tied to the concomitant failure of many state courts to require reliance during class certification. In particular, it contends that the lack of a reliance requirement creates incentives for plaintiffs' attorneys to bring consumer fraud class action suits without ever alleging that the consumers relied on, and hence that they were damaged by, the alleged misrepresentation, all in the hopes of forcing a settlement. The Article also provides a detailed history of the FTC, the subsequent rise in state consumer fraud statutes, and the early failure of both federal and state government agencies to adequately pursue violations of these laws. It then asserts that this failure and the subsequent rise of public law tort theory led state courts to slowly chip away at the element of reliance in a misguided attempt to provide adequate deterrence. However, now that both the FTC and state attorneys general enforce these consumer protection laws with more vigor, the Article concludes that requiring reliance for the resolution of private suits, while not requiring it in cases of public enforcement, creates the correct balance of individual justice and deterrence.

If you buy a jar of jam labeled "Simply 100% Fruit," do you really expect the jam to contain nothing but fruit? Apparently, today's consumer class action plaintiffs or, more accurately, their lawyers do. In *Smith v. J.M. Smucker Co.*,[1] the plaintiffs filed a class action consumer fraud lawsuit in Illinois state court, alleging that Smucker's "Simply 100% Fruit"

---

* Honorable Abraham L. Freedman Fellow and Lecturer in Law, Temple University School of Law. I would like to thank Richard K. Greenstein, Anthony J. Franze, Christopher J. Robinette, and Byron G. Stier for their helpful comments on earlier drafts of this Article. In addition, I would like to thank Gloria Lee for her excellent research assistance. Errors and omissions are mine alone.

[1] No. 03CH08522 (Ill. Cir. Ct. filed May 16, 2003), *noted in* J.M. Smucker Co. v. Rudge, 877 So. 2d 820, 821 (Fla. Dist. Ct. App. 2004).

jams do not contain 100% fruit.[2] Twenty-one identical "100% Fruit" class actions have been filed against Smucker's in twelve other states.[3]

Increasingly, plaintiffs' lawyers[4] are using consumer fraud statutes to pursue class actions based on manufacturers' alleged misrepresentations[5] about their products. By themselves, these lawsuits are not troubling. But when the consumers themselves have never relied on a manufacturer's misrepresentation, have never independently sought redress, and likely will never receive meaningful benefit from a suit (although their lawyers stand to make millions of dollars), these class actions become more akin to corporate blackmail than to consumer protection.

What prompted this trend? A significant factor in the rise of consumer fraud class action suits[6] is the emerging practice of allowing these claims to proceed through the process of class certification without any allegation of reliance—the traditional causal element of a common law misrepresentation claim that requires an injured party to allege that the manufacturer's misrepresentation induced the consumer to purchase the

---

[2] *Rudge*, 877 So. 2d at 821 (describing allegations of claim in *Smith*). Notably, the ingredients label on the Smucker's strawberry "Simply 100% Fruit" product indicates that the jam is, in fact, made entirely from fruit products: fruit syrup, strawberries, lemon juice concentrate, fruit pectin, red grape juice concentrate, and natural flavors. Howard Fischer, *Smucker's Mislabels Its Spread, Suit Claims*, ARIZ. DAILY STAR, July 24, 2004, at D, *available at* 2004 WLNR 11612584. Other Smucker's products, by contrast, contain high fructose corn syrup, corn syrup, sugar, and citric acid (in addition to strawberries and fruit pectin). *Id.*

[3] *See Rudge*, 877 So. 2d at 821 & n.1 (discussing Florida and Illinois cases and noting eighteen state class actions in eleven states); Fischer, *supra* note 2 (noting three additional suits in Arizona, California, and Wisconsin).

[4] Commentators have noted that class action suits often are not initiated by an injured party seeking redress but rather are created by lawyers. Whereas the typical lawsuit begins with a client seeking representation from a particular attorney, class action "attorneys use regulatory, media, and other electronic databases to identify instances of possible corporate wrongdoing." DEBORAH R. HENSLER ET AL., CLASS ACTION DILEMMAS: PURSUING PUBLIC GOALS FOR PRIVATE GAIN 72 (2000). Thus, these suits generally are seen as "lawyer-driven," not "client-driven." *See, e.g.*, John C. Coffee, Jr., *Class Action Accountability: Reconciling Exit, Voice, and Loyalty in Representative Litigation*, 100 COLUM. L. REV. 370, 384 (2000) ("[I]n the class action, the class representative is usually a token figure, with the class counsel being the real party in interest."); Victor E. Schwartz et al., *Federal Courts Should Decide Interstate Class Actions: A Call for Federal Class Action Diversity Jurisdiction Reform*, 37 HARV. J. ON LEGIS. 483, 492 (2000) ("[M]any [class actions] arise simply as a result of the creativity of entrepreneurial contingency fee lawyers."). *Cf.* HENSLER ET AL., *supra*, at 402 (noting that although plaintiff class action attorneys play a critical role in driving class action litigation, consumers, regulators, journalists, and ordinary lawyers also play a part).

[5] As used in this Article, "misrepresentation" refers to all methods of conveying untrue information to consumers that might induce a consumer to buy a product, including mislabeling, false advertising, and other deceptive sales promotion techniques. *See* RESTATEMENT (SECOND) OF TORTS § 525 cmt. b (1977) (stating that "misrepresentation" denotes not only "words spoken or written but also any other conduct that amounts to an assertion not in accordance with the truth"); JOHN MICKLEBURGH, CONSUMER PROTECTION 171 (1979) (noting that a "misrepresentation" means "an untrue statement of fact, made by one party to the contract (the 'misrepresentor') to the other (the 'misrepresentee'), before or at the time of contracting").

[6] *See infra* notes 18–20 and accompanying text.

product.[7] Because "a host of individual factors could have influenced a class member's decision to purchase the product," reliance-causation presents an individual issue for each class member.[8] The individualized nature of establishing reliance-causation makes class certification under Federal Rule 23(b)(3)[9] more difficult.[10] Perhaps to facilitate class actions, however, courts recently have softened the underlying substantive law, eliminating reliance as a required element.[11] This reduced standard, in turn, has helped propel the growing trend of consumer fraud class action lawsuits in which the plaintiffs never relied on an alleged misrepresentation—a trend that has affected numerous industries, including cigarette manufacturers,[12] fast food companies,[13] gasoline producers,[14] and the telecommunications sector.[15] Yet, despite recent billion-dollar verdicts[16] and million-dollar settlements,[17] and the fact that roughly one-third of class

---

[7] *See* discussion *infra* Part III.B.

[8] Hazelhurst v. Brita Prods. Co., 744 N.Y.S.2d 31, 33 (N.Y. App. Div. 2002).

[9] Federal Rule of Civil Procedure 23(b)(3) requires that "questions of law or fact common to the members of the class predominate over any questions affecting only individual members." In addition, the plaintiff also must show that "the class action is superior to other available methods for the fair and efficient adjudication of the controversy." FED. R. CIV. P. 23(b)(3). Finally, a class must satisfy the four threshold requirements of Rule 23(a): (1) numerosity, (2) commonality, (3) typicality, and (4) adequacy of representation. FED. R. CIV. P. 23(a). Most state systems employ similar requirements for class certification. 4 ALBA CONTE & HERBERT B. NEWBERG, NEWBERG ON CLASS ACTIONS § 13:1, at 399 (4th ed. 2002) (noting that Federal Rule of Civil Procedure 23 is the "most prevalent model" for state class action rules). Most importantly, the majority of states follow the federal predominance and commonality requirements. *See id.* §§ 13:9, 13:10, 13:16, at 404–06, 410–13 (noting that although specific language varies, state class action rules require commonality).

[10] *Cf., e.g., Hazelhurst*, 744 N.Y.S.2d at 33 (refusing to certify class under New York statute requiring individual reliance).

[11] *See* discussion *infra* Part III.B.2. Some scholars have argued that class certification is appropriate even where reliance is a required element, contending that reliance can be proven statistically or based on an objective standard. *E.g.*, Samuel Issacharoff, *The Vexing Problem of Reliance in Consumer Class Actions*, 74 TUL. L. REV. 1633, 1654 (2000) (arguing for an objective standard of reliance). Such evidentiary questions, however, are beyond the scope of this Article.

[12] *E.g.*, Aspinall v. Philip Morris Cos., 813 N.E.2d 476 (Mass. 2004).

[13] *E.g.*, Pelman v. McDonald's Corp., No. 02 Civ. 7821 (RWS), 2003 WL 22052778 (S.D.N.Y. Sept. 3, 2003), *vacated in part*, 396 F.3d 508 (2d Cir. 2005); *cf.* Michelle Morgante, *Mother Sues Cereal Makers for Recent Low-Sugar Claims*, CHARLESTON GAZETTE (W. Va.), Mar. 29, 2005, at 5C, *available at* 2005 WLNR 4922627 (discussing newly filed class action against cereal manufacturers based on allegations that the "low sugar" labeling misleadingly suggests that the cereals are healthier).

[14] *E.g.*, Weinberg v. Sun Co., 777 A.2d 442 (Pa. 2001).

[15] *E.g.*, Davis v. Powertel, Inc., 776 So. 2d 971 (Fla. Dist. Ct. App. 2000).

[16] *E.g.*, Price v. Philip Morris Inc., No. 00-L-112, 2003 WL 22597608 (Ill. Cir. Ct. Mar. 21, 2003) (awarding $10.1 billion judgment based on alleged misrepresentation of the "healthiness" of "light" cigarettes). This case is pending on appeal before the Illinois Supreme Court. *See* Supreme Court of Illinois, Docket, September Term of 2005, at 3 (2005), http://www.state.il.us/court/SupremeCourt/Docket/2005/Pdf/0905_Docket.pdf (listing Price v. Philip Morris Inc., No. 96236 on advisement docket of Sept. 2005 term) (last visited Nov. 4, 2005).

[17] *See, e.g.*, Bart Jansen, *Senate OKs Curb on Class Action: The Bill Aims to Steer Class-Action Lawsuits, Such as One Involving Poland Spring Water, into Federal Courts*,

actions brought against business defendants each year involve consumer claims,[18] these cases have received little attention in the tort reform debate. Indeed, during the 1995–1996 period, "[c]onsumer cases accounted for half of all reported state judicial decisions in class actions against business defendants,"[19] and within this category, "fraud cases comprised the largest fraction of reported federal judicial decisions."[20]

Recent attempts at reform, such as the Class Action Fairness Act of 2005,[21] which principally addresses the appropriate forum for class litigation and imposes limits on types of settlements, will do little to stem the tide of such suits. Before the Class Action Fairness Act became law, for instance, Yale law professor George Priest explained to President Bush that the bill was "not going to solve the problem."[22] Rather, the solution, according to Priest, required tighter application of the liability standards underlying a proposed class.[23] This Article attempts to fill the gap identified by Professor Priest by offering substantive guidance on how to fix the underlying liability rules in misrepresentation class actions: courts should treat fraud like fraud and require plaintiffs to allege "reliance" as an essential element of a consumer misrepresentation case.

Part I of this Article describes the new "misrepresentation" action and explains why state consumer fraud statutes have become attractive class action vehicles. Part II examines the origins of the misrepresenta-

---

PORTLAND PRESS HERALD (Me.), Feb. 11, 2005, at A1, *available at* 2005 WLNR 1944456 (noting settlement of $9.35 million in class action against Poland Spring water company based on advertising that product was "spring water" when water was pumped from spring source).

[18] *See* HENSLER ET AL., *supra* note 4, at 53–54. In 2000, the RAND Institute for Civil Justice ("ICJ") published a study of class action litigation. *Id.* Given the absence of any comprehensive class action database, the ICJ used LEXIS/NEXIS to develop its data. *Id.* at 52. The ICJ surveyed three sources for the 1995–96 period, including reported judicial decisions. Consumer cases represented thirty-five percent of reported judicial decisions. *Id.* at 54 fig.3.2. Within this category of consumer claims, roughly one-third were fraud-related, which encompassed deceptive sales practices, false advertising, and deceptive labeling. *Id.* at 55.

[19] *Id.* at 57 & 57 fig.3.5; *see also* FTC, CONSUMER FRAUD IN THE UNITED STATES: AN FTC SURVEY ES-2 (2004) (finding that nearly 25 million Americans were victims of consumer fraud in 2003), *available at* http://www.ftc.gov/reports/consumerfraud/040805confraud.pdf.

[20] HENSLER ET AL., *supra* note 4, at 57 & 57 fig.3.5; *see also* James M. Underwood, *Rationality, Multiplicity & Legitimacy: Federalization of the Interstate Class Action*, 46 S. TEX. L. REV. 391, 402 (2004) ("[A]ll things considered, the consumer class action will further evolve into the single most widespread tool for the class action."); John H. Beisner & Jessica Davidson Miller, *They're Making a Federal Case Out of It . . . In State Court*, 25 HARV. J. L. & PUB. POL'Y. 143, 156 (noting 340% increase in federal class actions and 1315% increase in state class actions over past decade).

[21] Class Action Fairness Act of 2005, 28 U.S.C.A. §§ 1453, 1711–1715 (West 2005).

[22] Press Release, White House, *President Discusses Lawsuit Abuse at White House Economy Conference* (Dec. 15, 2004), *available at* http://www.whitehouse.gov/news/releases/2004/12/20041215-11.html.

[23] *Id.* (stating that "the most fundamental reforms have to come from the courts. It's the courts that created this problem and it has to come from the courts in redefining liability rules.").

tion class action and describes how the forces that drove the creation of the consumer protection laws in the 1960s still serve as the backdrop for the current interpretation of the private damages action.

Part III examines modern consumer fraud suits brought by the government and contrasts the standards applicable to government suits with those applicable to private actions. It explains how courts have embraced the public purpose of a government suit—deterrence and punishment—in interpreting private consumer fraud statutes and have abandoned the traditional tort requirement of reliance-causation.

Part IV describes the "public tort law" theory that has contributed, in large part, to the abandonment of reliance-causation by state courts. It contrasts "public law" theory with the traditional understanding of the tort system as a means of providing redress and shows how public law theory provides an interpretative foundation for understanding the relaxation of reliance-causation requirements in misrepresentation class action suits.

Finally, Part V argues that the historical forces that led to the creation of the consumer class action—and the "public law" approach to these statutes—should no longer provide the interpretative framework for misrepresentation cases. Requiring reliance for private suits achieves the proper balance of public and private resources: allowing government agencies to seek restitution and injunctive relief where there is no consumer reliance and letting private litigants seek damages where reliance provides a causal connection between the defendant's conduct and the injury. Part V concludes that reinstating the traditional reliance requirement is an appropriate and simple fix that would restore the balance between public enforcement and private litigation.

I. THE NEW CLASS ACTION: STATUTORY MISREPRESENTATION CASES

Pelman v. McDonald's Corp.[24] is typical of the new misrepresentation cases in which the plaintiffs claim to have been defrauded but did not rely on any specific misrepresentation by the defendant. In *Pelman*, the plaintiffs filed a putative class action against McDonald's, alleging that McDonald's misrepresented that its products were nutritious and could be consumed as part of a healthy lifestyle on a daily basis.[25]

The plaintiffs were minor children whose parents purchased McDonald's for them three to five times a week.[26] The suit did not allege that the parents of these children actually relied on any false advertisement by

---

[24] No. 02 Civ. 7821 (RWS), 2003 WL 22052778 (S.D.N.Y. Sept. 3, 2003), *rev'd in part*, 396 F.3d 508 (2d Cir. 2005).

[25] *Id.* at *2. The plaintiffs based these allegations on a variety of McDonald's advertisements, such as an ad describing McDonald's beef as "nutritious" and "leaner than you think." *Id.*

[26] *Pelman*, 396 F.3d at 510.

McDonald's, nor could they even point to a specific false advertisement.[27] Rather, the plaintiffs claimed that their own "misconceptions" about the healthiness of a McDonald's diet generally resulted from McDonald's "long-term deceptive campaign."[28] Plaintiffs brought suit under two of New York's consumer fraud statutes: New York General Business Law Section 349[29] and New York General Business Law Section 350.[30]

The district court found that only the Section 350 claim required actual reliance,[31] even though both statutes used identical causation language.[32] The district court held that the plaintiffs' vague allegations of reliance on a "long-term deceptive campaign" did not satisfy Section 350's reliance requirement.[33] Rather, the court concluded that the plaintiffs must claim that they saw the allegedly false advertisement and "relied to their detriment" on the specific advertisement.[34]

Although the district court dismissed the suit in part for lack of causation,[35] the Second Circuit reversed, finding that the bare assertions in the complaint were sufficient to state a claim.[36] To date, plaintiffs still have not identified the particular advertisements that deceived them or even demonstrated that they ever saw or heard any particular advertisement.[37]

Under common law theories, a consumer would not be able to pursue a claim against McDonald's unless she had justifiably relied on the manufacturer's misrepresentations.[38] The reliance requirement would en-

---

[27] *Pelman*, 2003 WL 22052778, at *7.

[28] *Id.*

[29] Section 349-a prohibits "[d]eceptive acts or practices in the conduct of any business, trade or commerce or in the furnishing of any service in this state." N.Y. GEN. BUS. LAW § 349-a (McKinney 2004).

[30] Section 350 prohibits false advertising. N.Y. GEN. BUS. LAW § 350 (McKinney 2004).

[31] *Pelman*, 2003 WL 22052778, at *7.

[32] Section 349-h allows "any person who has been injured by reason of any violation of this section" to bring a damages action. N.Y. GEN. BUS. LAW § 349-h (McKinney 2004). Likewise, Section 350(e) provides that "[a]ny person who has been injured by reason of any violation of [this] section" may bring a damages action. *Id.* at § 350-e.

[33] *Pelman*, 2003 WL 22052778, at *8.

[34] *Id.*

[35] *Id.* at *11–*12, *14.

[36] *Pelman*, 396 F.3d at 511–12.

[37] *See* Plaintiffs' Memorandum of Law in Opposition to Defendant's Motion for a More Definite Statement, Pelman v. McDonald's Corp., No. 02 Civ. 7821 (RWS), 2005 WL 1276744 (S.D.N.Y. May 22, 2005) (arguing that plaintiffs are not required to identify specific advertisements); Reply Memorandum of Law in Support of Defendant's Motion for a More Definite Statement, *Pelman*, No. 02 Civ. 7821 (RWS), 2005 WL 1276745 (S.D.N.Y. May 4, 2005) (arguing that plaintiffs have failed to identify particular McDonald's advertisements that caused them injury).

[38] To state a claim for common law fraud, a plaintiff must allege a false representation by the defendant; the defendant's knowledge or belief that the representation is false; the defendant's intent to induce the plaintiff to act or refrain from action in reliance upon the misrepresentation; the plaintiff's justifiable reliance upon the misrepresentation in taking action or refraining from it; and damage to the plaintiff resulting from such reliance. W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS § 105, at 728 (5th ed. 1984); *see also* RESTATEMENT (SECOND) OF TORTS § 525 (1977) (stating that misrepresen-

sure that the defendant's misrepresentation had, in fact, caused the alleged harm.

Consumer class actions like the *Pelman* suit, however, have pushed the law in a new direction that dilutes the reliance requirement. The *Pelman* case is just one high-profile example. Other cases include claims by smokers alleging that "light" cigarettes were deceptively labeled;[39] claims by cell phone consumers that the manufacturer misrepresented the phone's coverage area;[40] and claims by parents that baby food was not "pure and natural," as advertised.[41] Not one of these cases required the class members to plead that the manufacturer's allegedly false statement made any difference in their decision to buy the product. Thus, many courts have abandoned reliance—a crucial link between a defendant's misrepresentation and a plaintiff's injury—and have thereby significantly reduced the showing necessary to certify a case as a class action.[42]

Although some courts ultimately require proof of reliance to establish causation at trial,[43] certification places significant pressure on a defendant. Given the enormous amount of money at stake, certification becomes "the decisive point in a class action. Following certification, class actions often head straight down the settlement path because of the very high cost for everybody concerned, courts, defendants, plaintiffs of litigating a class action . . . ."[44] Regardless of the amount of actual damages, if any,

---

tation claim requires "pecuniary loss" and "justifiable reliance upon the misrepresentation").

[39] *E.g.*, Aspinall v. Philip Morris Cos., 813 N.E.2d 476 (Mass. 2004). *Aspinall* was a misrepresentation class action on behalf of Massachusetts purchasers of Marlboro Lights cigarettes. *Id.* at 485. The plaintiffs alleged that tobacco manufacturer Philip Morris deceptively marketed these cigarettes as "light," creating the impression that these cigarettes were "healthier" than other cigarettes. *Id.* at 480–82. On interlocutory appeal, the Massachusetts Supreme Court held that the plaintiffs were not required to prove that they relied on the label "light" when deciding to purchase that particular brand of cigarettes. *Id.* at 487–89. The court reached this conclusion even though the court found that, to be deceptive, an advertisement must induce consumers to "act differently from the way they otherwise would have acted (i.e., to entice a reasonable consumer to purchase the product)." *Id.* at 488.

[40] *E.g.*, Davis v. Powertel, Inc., 776 So. 2d 971 (Fla. Dist. Ct. App. 2000). In *Davis*, the Florida District Court of Appeal approved a misrepresentation class action based on a cell phone manufacturer's failure to disclose that the phone had been modified and would work only with its own wireless system. *Id.* at 972, 974–75. The court found that reliance on the manufacturer's statement was unnecessary. *Id.* at 974–75.

[41] *E.g.*, Tylka v. Gerber Products Co., 178 F.R.D. 493 (N.D. Ill. 1998). Because reliance was not required under Illinois' consumer fraud act, *id.* at 499, even those consumers who still bought Gerber baby food despite knowledge of the allegedly false advertising could be part of the class. *See id.* at 498, 502.

[42] *See, e.g.*, *Aspinall*, 813 N.E.2d at 486–87 (finding certification of consumer class action was warranted where reliance was not required). *But see* Philip Morris Inc. v. Angeletti, 752 A.2d 200, 239–40 (Md. 2000) (reversing certification of consumer class action based on individual issues of reliance).

[43] *E.g.*, Group Health Plan Inc. v. Philip Morris Inc., 621 N.W.2d 2, 13 (Minn. 2001); *see also* discussion *infra* Part III.B.1.

[44] Bruce Hoffman, Remarks, Panel 7: Class Actions as an Alternative to Regulation: The Unique Challenges Presented by Multiple Enforcers and Follow-On Lawsuits, 18 GEO.

many states allow recovery of a statutory minimum ranging from $100 to $2,000 per plaintiff.[45] Moreover, many statutes authorize multiple or punitive damages,[46] and four states even *require* an award of treble damages to a victorious plaintiff.[47]

Thus, certification can create enormous pressure on a defendant to settle, regardless of the merits of a case. As Judge Posner has explained, certification of a class action, even one without merit, forces defendants "to stake their companies on the outcome of a single jury trial, or be forced by fear of the risk of bankruptcy to settle even if they have no legal liability . . . ."[48] Judge Posner is not alone in his view. The Judiciary Committee of the United States House of Representatives recently concluded that a corporation faces enormous pressure to settle a case once a class is certified, even when the case lacks merit:

J. LEGAL ETHICS 1311, 1329 (2005) (panel discussion statement of Bruce Hoffman, then Deputy Director of the FTC's Bureau of Competition). A glaring example of the coercive effect of certification can be found under federal law. In Shaw v. Toshiba Am. Info. Sys., Inc., 91 F. Supp. 2d 926, 928–29 (E.D. Tex. 1999), the plaintiffs alleged that the defendant's computers would corrupt data under certain conditions and sought relief under the Computer Fraud and Abuse Act, 18 U.S.C. § 1030 (2000). The court certified a class consisting not only of current owners of Toshiba computers or even future owners, but also of "potential purchasers." *Id.* at 938. Incredibly, the court found, "it is not necessary for someone to actually own a defective computer in order to experience continuing, adverse effects from it." *Id.* As noted in the Congressional Record, not a single customer had reported any problem to Toshiba regarding this alleged defect. 149 CONG. REC. S12423 (daily ed. Oct. 3, 2003) (statement of Sen. Sessions). But now that the class was certified, Toshiba faced potential liability of $10 billion and felt forced to settle. *Id.* Thus, following certification, Toshiba settled the case for $2.1 billion. Shaw v. Toshiba Am. Info. Sys., Inc., 91 F. Supp. 2d 942, 953, 959 (E.D. Tex. 2000) (approving settlement class, though settlement class was limited to current owners). On top of the $2.1 billion award, Toshiba also agreed to pay $147.5 million in attorneys' fees to the plaintiffs. *Id.* at 961.

[45] *E.g.*, ALA. CODE § 8-19-10(a)(1) (LexisNexis 2002) (statutory minimum of $100); D.C. CODE ANN. § 28-3905(k)(1) (LexisNexis 2001) ($1,500); UTAH CODE ANN. § 13-11-19 (2001) ($2,000). Massachusetts has the lowest statutory minimum at $25. MASS. GEN. LAWS ANN. ch. 93A, § 9(3) (West 2005).

[46] Several states require a multiple award where the defendant willfully engaged in a deceptive act. *E.g.*, COLO. REV. STAT. § 6-1-113(2) (2004); LA. REV. STAT. ANN. § 51:1409 (2003); MASS. GEN. LAWS ANN. ch. 93A, § 9 (West 2005); N.H. REV. STAT. ANN. § 358-A:10 (1995); S.C. CODE ANN. § 39-5-140 (1985). Other states leave the award of multiple or punitive damages to the court's discretion. *E.g.*, ALA. CODE § 8-19-10(a)(2) (LexisNexis 2002); CONN. GEN. STAT. ANN. § 42-110g (West 2000 & Supp. 2005); D.C. CODE ANN. § 28-3905(k)(1) (LexisNexis 2001); IDAHO CODE ANN. § 48-608(1) (1997); 815 ILL. COMP. STAT. ANN. 505/10a (West 1999 & Supp. 2005); KY. REV. STAT. ANN. § 367.220 (West 2002); MO. ANN. STAT. § 407.025 (West 2001); MONT. CODE ANN. § 30-14-133 (2003); OR. REV. STAT. ANN. § 646.638 (West 1988 & Supp. 1998); R.I. GEN. LAWS § 6-13.1-5.2 (2001); TENN. CODE ANN. § 47-18-109 (2001).

[47] Hawaii, New Jersey, and North Carolina mandate treble damages by statute, while Texas imposes treble damages by judicial interpretation. HAW. REV. STAT. ANN. § 480-13 (LexisNexis 2002 & Supp. 2004); N.J. STAT. ANN. § 56:8-19 (West 2001); N.C. GEN. STAT. § 75-16 (2003); Woods v. Littleton, 554 S.W.2d 662, 668–69 (Tex. 1977) (finding that treble damages are mandatory even though statutory language suggests discretionary standard).

[48] *In re* Rhone-Poulenc Rorer, Inc., 51 F.3d 1293, 1299 (7th Cir. 1995).

> [T]he perverse result [is] that companies that have committed no wrong find it necessary to pay ransom to plaintiffs' lawyers because the risk of attempting to vindicate their rights through trial simply cannot be justified to their shareholders. Too frequently, corporate decisionmakers are confronted with the implacable arithmetic of the class action: even a meritless case with only a 5% chance of success at trial must be settled if the complaint claims hundreds of millions of dollars in damages.[49]

Moreover, these settlements do not necessarily benefit consumers. Many take the form of a "coupon settlement," where the consumer plaintiffs receive coupons from the defendant-manufacturer.[50] For example, a class sued bottled water company Poland Spring, alleging that it misrepresented that its product was "spring water."[51] In the settlement, class members simply received coupons for more bottled water—a product the class purportedly did not want in the first place.[52] The plaintiffs' lawyers meanwhile received a $1.35 million fee.[53] Moreover, even where a consumer class action settles with a monetary award to the plaintiffs, few individual plaintiffs will submit the necessary claims forms and ultimately share in these proceeds.[54]

The ease of certification and its coercive settlement pressure are not the only reasons plaintiffs' counsel have chosen to pursue these consumer protection claims. In many states, a prevailing plaintiff automatically recovers attorneys' fees.[55] Probably the most notorious attorneys' fee award

---

[49] H.R. REP. No. 106-320, at 8 (1999); *accord* S. REP. No. 109-14, at 20–21 (2005), *as reprinted in* 2005 U.S.C.C.A.N. 3, 21.

[50] For a thorough analysis of coupon settlements, see Christopher R. Leslie, *A Market-Based Approach to Coupon Settlements in Antitrust and Consumer Class Action Litigation*, 49 UCLA L. REV. 991 (2002) [hereinafter Leslie, *A Market-Based Approach to Coupon Settlements*]. The Class Action Fairness Act now provides some limits on coupon settlements. *See infra* notes 302–304 and accompanying text.

[51] Jansen, *supra* note 17; *accord* Marguerite Higgins, *Class Members Get Little in Suits, Lawyers' Fees Spur Legislation*, WASH. TIMES (D.C.), July 8, 2004, at A12, *available at* 2004 WLNR 811926.

[52] Jansen, *supra* note 17.

[53] *Id.*

[54] *See* Gail Hillebrand & Daniel Torrence, *Claims Procedures in Large Consumer Class Actions and Equitable Distribution of Benefits*, 28 SANTA CLARA L. REV. 747, 747 (1988) (noting that settlement "claims procedures are ill-suited to consumer class actions in which the class size is very large and the amount of damages per class member is relatively small. These cases are characterized by very low claims rates.").

[55] At least seventeen states automatically award attorneys' fees to a prevailing plaintiff. ALA. CODE § 8-19-10(a) (LexisNexis 2002); COLO. REV. STAT. § 6-1-113(2) (2004); GA. CODE ANN. § 10-1-399(d) (2000); HAW. REV. STAT. ANN. § 480-13(b) (LexisNexis 2002 & Supp. 2004); IDAHO CODE ANN. § 48-608(4) (1997); LA. REV. STAT. ANN. § 51:1409 (2003); ME. REV. STAT. ANN. tit. 5, § 213 (2002); MASS. GEN. LAWS ANN. ch. 93A, § 9(4) (West 2005); NEV. REV. STAT. ANN. § 41.600 (LexisNexis 2002); N.H. REV. STAT. ANN. § 358-A:10 (1995); N.J. STAT. ANN. § 56:8-19 (West 2001); N.M. STAT. ANN. § 57-12-10 (LexisNexis 2000); OKLA. STAT. ANN. tit. 15, § 761.1 (West 1993 & Supp. 2005); S.C. CODE ANN. § 39-5-140(a) (1985); TEX. BUS. & COM. CODE ANN. § 17.50(d) (Vernon

in recent years stands at $1.75 *billion*, awarded in *Price v. Philip Morris, Inc.*, a misrepresentation class action based on the alleged false impression created by labeling cigarettes "light."[56]

By not requiring reliance, courts have fueled class action abuse, providing additional incentives to bring claims that stand to benefit lawyers far more than consumers. The obvious question, then, is have courts been inclined to interpret these statutes as eliminating any reliance requirement? The answer lies in the origins of consumer fraud statutes and a misapplication of public law theory.

II. ORIGINS OF THE CONSUMER FRAUD CLASS ACTION

The enlargement of consumer remedies for false statements hardly has followed a straight path. Rather, various factors have converged in a piecemeal and disjointed fashion to facilitate these claims. Many of these trends first developed in the 1960s when the consumer protection movement reemerged.[57] Decades later, these changes still form the backdrop for modern interpretation of consumer fraud statutes.[58] Two main events

---

2002); WIS. STAT. ANN. § 100.20(5) (West 2004 & Supp. 2004); WYO. STAT. ANN. § 40-12-108(b) (2005). In addition, six states award attorneys' fees where the defendant acted willfully. DEL. CODE ANN. tit. 6, § 2533 (1999); GA. CODE ANN. § 10-1-373 (2000); HAW. REV. STAT. ANN. § 481A-4(b) (LexisNexis 2002); ME. REV. STAT. ANN. tit. 10, § 1213 (1997); NEB. REV. STAT. § 87-303(b) (1999); OKLA. STAT. ANN. tit. 78, § 54 (West 2002). Finally, at least eighteen states allow attorneys' fees at the court's discretion. CONN. GEN. STAT. ANN. § 42-110g(d) (West 2000 & Supp. 2005); FLA. STAT. ANN. § 501.211 (West 2002); IND. CODE ANN. § 24-5-0.5-4(a) (West 1980 & Supp. 2004); KAN. STAT. ANN. § 50-634(e) (1994); KY. REV. STAT. ANN. § 367.220(3) (West 2002); MD. CODE ANN., COM. LAW § 13-408(b) (LexisNexis 2000 & Supp. 2003); MINN. STAT. ANN. § 325D.45 (West 2004 & Supp. 2005); MO. ANN. STAT. § 407.025 (West 2001); MONT. CODE ANN. § 30-14-133(3) (2003); N.Y. GEN. BUS. LAW § 349(h), 350-e (McKinney 2004); N.C. GEN. STAT. § 75-16.1 (2003); OHIO REV. CODE ANN. § 1345.09(F) (LexisNexis 2002); OKLA. STAT. ANN. tit. 78, § 54 (West 2002); OR. REV. STAT. ANN. § 646.638(3) (West 1988 & Supp. 1998); 73 PA. STAT. ANN. § 201-9.2(a) (West 1993 & Supp. 2005); R.I. GEN. LAWS § 6-13.1-5.2(d) (2001); TENN. CODE ANN. § 47-18-109(e)(1) (2001); UTAH CODE ANN. § 13-11-19(5) (2001).

[56] No. 00-L-112, 2003 WL 22597608, at *29 (Ill. Cir. Ct. Mar. 21, 2003) (awarding attorney's fees in the amount of twenty-five percent of a $7.1005 billion compensatory award), *appeal docketed*, No. 96236 (Ill. argued Nov. 10, 2004).

[57] The consumer movement did not originate in the 1960s. Rather, the movement began thirty years earlier in the 1930s; proposals to create a federal "consumer counsel" were even floated in the late 1920s. STANLEY MORGANSTERN, LEGAL PROTECTION FOR THE CONSUMER 5 (Irving J. Sloan ed., 2d ed. 1978); *see also* Richard C. Schragger, *The Anti-Chain Store Movement, Localist Ideology, and the Remnants of the Progressive Constitution, 1920–1940*, 90 IOWA L. REV. 1011, 1070–76 (2005) (describing growth of consumer protection movement in the 1930s). However, in March 1962, President John F. Kennedy issued a presidential message to Congress addressing consumer issues. MORGANSTERN, *supra*, at 5. The next year, President Kennedy adopted a "Consumer Bill of Rights," which caused a flurry of new federal legislation and efforts to obtain stronger enforcement. PRACTISING LAW INST., CONSUMER PROTECTION 17 (1972). At the same time, the 1960s saw the rise of Ralph Nader and his consumerism movement. *See* discussion *infra* notes 76–78 and accompanying text.

[58] *See* discussion *infra* notes 196–200 and accompanying text.

during the 1960s produced the modern misrepresentation class action: (1) the harsh criticism resulting from the failure of federal government enforcement, and (2) the corresponding rise of state consumer fraud statutes.

A. Failings of the Public Enforcement

The Federal Trade Commission ("FTC")[59] was created in 1914.[60] Originally, however, the FTC was aimed at curbing the monopoly power of big business, not protecting consumers.[61] Indeed, the original Federal Trade Commission Act ("FTC Act") banned only "unfair methods of competition,"[62] which required that the "unfair methods" injure the business of a competitor. Thus, the original FTC Act did not cover false statements that affected only the public.[63] Although a few early FTC cases did involve deceptive advertising[64] or labeling,[65] these early cases also usually included an injury to competition, not just an injury to the consuming public.[66] As the Third Circuit pointed out, the FTC was "helpless" to remedy deceptive marketing where all members of an industry used the same deceptive practices or where competition did not exist for a particular product.[67]

---

[59] The FTC is headed by a panel of five Commissioners. Federal Trade Commission for the Consumer, Commissioners, http://www.ftc.gov/bios/commissioners.htm (last visited Nov. 4, 2005). The President, with the advice and consent of the Senate, appoints each Commissioner for a seven-year term. *Id.* The President also selects one Commissioner to serve as Chairman. *Id.* No more than three Commissioners may be from the same political party. *Id.*

[60] Act of Sept. 26, 1914, c. 311, 38 Stat. 717 (codified as amended at 15 U.S.C. § 45 (2000)).

[61] *See* FTC v. Raladam Co., 283 U.S. 643, 647 (1931) (stating object of FTC was to stop unfair competition); Marshall A. Leaffer & Michael H. Lipson, *Consumer Actions Against Unfair or Deceptive Acts or Practices: The Private Uses of Federal Trade Commission Jurisprudence*, 48 GEO. WASH. L. REV. 521, 524 (1980) (noting original antitrust purpose of FTC Act); William A. Lovett, *State Deceptive Trade Practice Legislation*, 46 TUL. L. REV. 724, 728 n.8 (1971) [hereinafter Lovett, *State Deceptive Trade Practice Legislation*] (noting FTC Act was intended to complement the Clayton Act).

[62] *Raladam Co.*, 283 U.S. at 644 n.1 (emphasis added).

[63] *Id.* at 649–51.

[64] *See, e.g.*, Sears, Roebuck & Co. v. FTC, 258 F. 307 (7th Cir. 1919) (sugar advertisements).

[65] *See, e.g.*, FTC v. Winsted Hosiery Co., 258 F. 483 (1922) (labels on woolen clothing).

[66] *See Winsted Hosiery Co.*, 258 U.S. at 493 (explaining that mislabeled goods diverted trade from honest manufacturers); *Sears, Roebuck & Co.*, 258 F. at 309 (discussing how Sears's advertisements suggested that competitors were "unfair dealers in sugar").

[67] Pep Boys—Manny, Moe & Jack, Inc. v. FTC, 122 F.2d 158, 161 (3d Cir. 1941); *see also* H.R. REP. No. 75-1613, at 3 (1937) (noting that the FTC was "powerless to act for consumer's protection" where all competitors engage in same unfair method or where no competition existed).

In 1938,[68] Congress finally gave the FTC the authority to protect consumers from "unfair and deceptive trade practices."[69] Indeed, Congress intended the FTC Act[70] to reach "every case from that of inadvertent or uninformed advertising to that of the most subtle as well as the most vicious types of advertisement."[71]

Although the FTC was armed with the authority to proscribe false advertising that injured the public,[72] the FTC did little to stop manufacturer misrepresentations. At the end of the 1960s, two scathing reports—one by a group of students led by Ralph Nader[73] and the other by the American Bar Association[74]—ruthlessly criticized the FTC's performance.[75]

In 1968, Ralph Nader recruited a group of law students who spent the summer investigating the FTC.[76] The final report, issued in January

---

[68] In 1938, Congress superseded *Raladam* with the Wheeler-Lea Amendment, which added a prohibition against "unfair or deceptive [trade] acts or practices." Wheeler-Lea Act of March 21, 1938, Pub. L. No. 75-447, 52 Stat. 111 (1939) (codified as amended at 15 U.S.C. § 45 (2000)); *see also* FTC v. Sperry & Hutchinson Co., 405 U.S. 233, 244 (1972) (noting that Wheeler-Lea Amendment reversed *Raladam*); United States v. J.B. Williams Co., No. 73-1624, 1973 WL 3183, at *8 (2d Cir. May 2, 1974) (discussing history of Wheeler-Lea Amendment). Congress first attempted to remedy the *Raladam* decision in 1935, but the bill died in the Senate. *J. B. Williams Co.*, 1973 WL 3183, at *8. It took Congress three years to pass the proposed changes. *Id.* For a compilation of the Wheeler-Lea Amendment's legislative history, see CHARLES WESLEY DUNN, WHEELER-LEA ACT: A STATEMENT OF ITS LEGISLATIVE RECORD (1938).

[69] 15 U.S.C. § 45 (2000). For a discussion of the Federal Trade Commission Act's legislative history from 1914 to 1938, see generally Holloway v. Bristol-Myers Corp., 485 F.2d 986, 990–96 (D.C. Cir. 1973).

[70] 15 U.S.C. § 45.

[71] H.R. REP. NO. 75-1613, at 5 (1937).

[72] *See Pep Boys*, 122 F.2d at 160 (finding procedure in FTC Act is "prescribed in the public interest"); *see also Sperry & Hutchinson Co.*, 405 U.S. at 239–44 (holding FTC has authority to proscribe unfair or deceptive practices regardless of any effect on competition).

[73] EDWARD F. COX, ROBERT C. FELLMETH & JOHN E. SCHULZ, "THE NADER REPORT" ON THE FEDERAL TRADE COMMISSION (1969) [hereinafter NADER REPORT].

[74] *Report of the ABA Commission to Study the Federal Trade Commission* [July–Sept.], Antitrust & Trade Reg. Rep. (BNA Spec. Supp.) No. 427 (Sept. 16, 1969) [hereinafter *ABA Report*]. In a separate statement attached to the ABA Report, then-Professor Richard Posner criticized the FTC's very existence and argued that the FTC served a "useful purpose" only in a "bare handful of cases." *Id.* at 112. Professor Posner urged "greater reliance on market processes and on the system of judicial rights and remedies" as a better alternative to the FTC. *Id.* at 118. Professor Posner later expanded his critique in a law review article, Richard A. Posner, *The Federal Trade Commission*, 37 U. CHI. L. REV. 47 (1969) [hereinafter Posner, *The Federal Trade Commission*], and a book, RICHARD A. POSNER, REGULATION OF ADVERTISING BY THE FTC (1973) [hereinafter POSNER, REGULATION OF ADVERTISING].

[75] *See infra* notes 76–81 and accompanying text. Although catalysts for great change, the *ABA Report* and the NADER REPORT were not the first reports to evaluate and criticize the FTC's performance. *ABA Report*, *supra* note 74, at 9. As early as the 1940s, critics charged that the FTC failed to prioritize its goals and focus on interests of public importance. *See id.* at 9–11 (discussing early criticisms of the FTC).

[76] NADER REPORT, *supra* note 73, at 3. Dubbed "Nader's Raiders" by the press, the group was comprised of six volunteer law students or recent law school graduates from Harvard and Yale and an architecture student from Princeton. *Id.* at 2. The students gathered information by conducting interviews of FTC employees and reviewing internal FTC

1969,[77] colorfully concluded that the FTC was "a self-parody of bureaucracy, fat with cronyism, torpid through an inbreeding unusual even for Washington, manipulated by the agents of commercial predators, [and] impervious to governmental and citizen monitoring."[78]

The government's immediate response to the Nader Report was the commission of the ABA Report.[79] Though more polite, the ABA Report was no less critical. The ABA charged that the FTC's consumer protection efforts were "inadequate" and "piecemeal."[80] The ABA found that the FTC "was preoccupied with technical labeling and advertising practices of the most inconsequential sort."[81]

As an illustrative example of the FTC's failure to address consumer fraud, and specifically false advertising, both the Nader Report and the ABA Report discussed the much publicized Geritol investigation of the 1960s.[82] In the 1950s, J. B. Williams Company manufactured Geritol, a vitamin and mineral supplement, and advertised the product as a remedy for fatigue and tiredness.[83] After more than three years of investigation, the FTC issued a complaint in December 1962, alleging that the statements misrepresented Geritol's efficacy.[84] But, not until 1965—nearly three years later—did the FTC finally order the company to stop making these statements.[85] Even though this order was affirmed by the Court of Appeals,[86] Geritol's TV advertisements changed little.[87] In an unusual proceeding, the FTC held public hearings in 1968 to determine whether Geritol's current advertising campaign complied with the 1965 order.[88] Following this hearing, the FTC concluded that the new Geritol advertisements "not only failed to comply with the order, but . . . are no less objectionable" than the original banned advertisements.[89] Instead of seeking civil penalties, however, the FTC simply ordered J. B. Williams to file another report.[90] Thus, "almost 10 years to the day after the beginning of the investigation,

---

documents. *Id.* at 6–7.

[77] *Id.* at 10.

[78] *Id.* at vii.

[79] *Id.* at xiii. Newly elected President Richard M. Nixon asked the American Bar Association to review "the 'present efforts of the Federal Trade Commission in the field of consumer protection, in its enforcement of the antitrust laws, and of the allocation of its resources between these two areas.'" *ABA Report*, *supra* note 74, at 4.

[80] *ABA Report*, *supra* note 74, at 37.

[81] *Id.* at 2.

[82] NADER REPORT, *supra* note 73, at 65–67; *ABA Report*, *supra* note 74, at 43–44.

[83] NADER REPORT, *supra* note 73, at 66; *ABA Report*, *supra* note 74, at 43.

[84] *ABA Report*, *supra* note 74, at 43.

[85] *ABA Report*, *supra* note 74, at 43; *see also* J.B. Williams Co. v. FTC, 381 F.2d 884, 886–87 (6th Cir. 1967) (discussing procedural history of Geritol case).

[86] *J.B. Williams Co.*, 381 F.2d at 891 (affirming order with slight modification).

[87] NADER REPORT, *supra* note 73, at 66; *ABA Report*, *supra* note 74, at 43.

[88] NADER REPORT, *supra* note 73, at 66; *ABA Report*, *supra* note 74, at 43.

[89] NADER REPORT, *supra* note 73, at 66 (quoting FTC News Release, Dec. 13, 1968) (internal quotations omitted); *accord ABA Report*, *supra* note 74, at 43.

[90] NADER REPORT, *supra* note 73, at 66; *ABA Report*, *supra* note 74, at 43.

the FTC found that certain of Geritol's commercials *still* violated the cease and desist order, but again it did not seek civil enforcement penalties."[91]

In short, the two reports highlighted the FTC's "utter lack of effectiveness."[92] As then-Professor Posner described the 1960s FTC, "[t]he Commission is rudderless; poorly managed and poorly staffed; obsessed with trivia; politicized; all in all, inefficient and incompetent."[93]

B. Rise of State Consumer Fraud Statutes

Not coincidentally, at the same time the FTC was being thrashed for its ineffectiveness, state legislatures enacted a tidal wave of consumer protection legislation. Three separate "model statute" movements emerged beginning in the 1960s and resulted in the enactment of consumer protection legislation in all fifty states.[94] First, the National Conference of Commissioners on Uniform State Laws ("NCCUSL")[95] proposed the Uniform Deceptive Trade Practices Act ("UDTPA").[96] Second, the FTC and the Committee on Legislation of the Council of State Governments proposed the Unfair Trade Practices and Consumer Protection Law ("UTP/CPL").[97] Finally, the NCCUSL proposed a third consumer fraud statute: the Uniform Consumer Sales Practices Act ("UCSPA").[98] Thus, by the mid-1970s, every state would have a consumer fraud statute that allowed private claims for damages.[99]

---

[91] *ABA Report*, *supra* note 74, at 43–44; *accord* NADER REPORT, *supra* note 73, at 67. Finally, in November 1969, the FTC certified its findings to the Attorney General. United States v. J.B. Williams Co., 498 F.2d 414, 418 (2d Cir. 1974). Five months later, the Attorney General filed suit against J.B. Williams, seeking a $1 million penalty against J.B. Williams and its advertising company. *Id.* The district court granted summary judgment in favor of the Government. United States v. J.B. Williams Co., 354 F. Supp. 521, 553 (S.D.N.Y. 1973). On appeal, however, the Second Circuit reversed on all but two counts, holding that the case presented triable issues of fact. *J. B. Williams Co.*, 498 F.2d at 421.

[92] NADER REPORT, *supra* note 73, at 95; *accord ABA Report*, *supra* note 74, at 12 (noting ineffective planning and coordination of activities within the agency).

[93] Posner, *The Federal Trade Commission*, *supra* note 74, at 47; *accord* POSNER, REGULATION OF ADVERTISING, *supra* note 74, at 21–23.

[94] *See infra* Parts II.B.1–3.

[95] The National Conference of Commissioners on Uniform State Laws is a non-profit association created in 1892 "to promote uniformity in state laws on all subjects where uniformity is deemed desirable and practicable." HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS AND PROCEEDINGS OF THE ANNUAL CONFERENCE MEETING IN ITS SEVENTY-THIRD YEAR 308 (1964) (citation omitted).

[96] *See infra* Part II.B.1.

[97] *See infra* Part II.B.2.

[98] *See infra* Part II.B.3.

[99] *See infra* notes 134 & 138 and accompanying text.

1. False Advertising Statutes and the UDTPA

The idea for state legislation started in 1964, when the FTC proposed that states enact false advertising statutes, modeled after New York's 1963 statute.[100] The FTC sought to continue its non-enforcement policy by shifting responsibility to "the lowest practicable level of government."[101] The next year, the Council of State Governments drafted a uniform false advertising statute.[102]

At the same time, in 1964, the NCCUSL proposed its own model consumer act: the Uniform Deceptive Trade Practices Act.[103] Although the UDTPA authorized private causes of actions, relief was limited to injunctions.[104] Five states adopted the 1964 version.[105] The FTC, however, believed that the 1964 UDTPA fell short in two respects.[106] First, the UDTPA only authorized private causes of action.[107] The FTC believed that a public official, such as the state's attorney general, should have authority to institute proceedings.[108] Second, the UDTPA contained not only an itemized list of prohibited practices, but also a "catch-all provision," applicable to conduct that "similarly creates a likelihood of confusion."[109] The

---

[100] Letter from Paul Rand Dixon, Chairman, FTC, to William D. Carey, Executive Assistant Director, Bureau of the Budget, Executive Office of the President (Apr. 14, 1966) (on file with author); *see also* N.Y. GEN. BUS. LAW § 350 (McKinney 2004) (prohibiting false advertising).

[101] Letter from Paul Rand Dixon, *supra* note 100.

[102] COUNCIL OF STATE GOV'TS, SUGGESTED STATE LEGISLATION 156–57 (1964).

[103] UNIFORM DECEPTIVE TRADE PRACTICES ACT (1964) (cited in HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS AND PROCEEDINGS OF THE ANNUAL CONFERENCE MEETING IN ITS SEVENTY-FIFTH YEAR 253–64 (1964)). The UDTPA was designed "to bring state [consumer protection] law up to date by removing undue restrictions on the common-law action for deceptive trade practices." *Id.* at 253. For an examination of the 1964 UDTPA, see Richard F. Dole, Jr., *Merchant and Consumer Protection: The Uniform Deceptive Trade Practices Act*, 76 YALE L.J. 485 (1967).

[104] UNIFORM DECEPTIVE TRADE PRACTICES ACT § 3 (1964) (cited in HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS AND PROCEEDINGS OF THE ANNUAL CONFERENCE MEETING IN ITS SEVENTY-FIFTH YEAR 262 (1964)).

[105] *See* UNIFORM DECEPTIVE TRADE PRACTICES ACT (1966) (cited in HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS AND PROCEEDINGS OF THE ANNUAL CONFERENCE MEETING IN ITS SEVENTY-FIFTH YEAR 299 (1966)). These states were Connecticut, Delaware, Idaho, Illinois, and Oklahoma. *Id.*; *accord* Dole, *supra* note 103, at 485 & n.4. Determining which model a state followed poses some difficulty as the states enacted—and revised—multiple "model" statutes. Accordingly, some states are categorized "twice" for following more than one model.

[106] Attachment to Letter from Paul Rand Dixon, Chairman, FTC, to William D. Carey, Executive Assistant Director, Bureau of the Budget, Executive Office of the President 4–5 (Apr. 14, 1966) (on file with author) [hereinafter Letter from FTC].

[107] UNIFORM DECEPTIVE TRADE PRACTICES ACT § 3(a) (1964) (cited in HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS AND PROCEEDINGS OF THE ANNUAL CONFERENCE MEETING IN ITS SEVENTY-FIFTH YEAR 262 (1964)) (authorizing private actions for injunctive relief).

[108] Letter from FTC, *supra* note 106, at 5.

[109] UNIFORM DECEPTIVE TRADE PRACTICES ACT § 2(a)(12) (1964) (cited in HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS AND PROCEEDINGS OF THE ANNUAL CONFERENCE MEETING IN ITS SEVENTY-FIFTH YEAR 262 (1964)).

FTC believed that the use of the word "similar" left doubt as to the scope of the catch-all provision.[110]

The NCCUSL revised the UDTPA in 1966 but did not address the FTC's concerns.[111] The revised UDTPA directed that the prevailing party be awarded costs and allowed the court to award attorneys' fees to the prevailing party.[112] Another four states adopted the 1966 version.[113] In addition, the UDTPA provides the foundation for consumer fraud statutes in another six states: Colorado, Minnesota, Nebraska, Nevada, New Mexico, and Oregon.[114]

2. *Unfair Trade Practices and Consumer Protection Law*

In 1970, the FTC and the Committee on Suggested State Legislation of the Council of State Governments issued their model statute: the Unfair Trade Practices and Consumer Protection Law.[115] The model UTP/CPL included three alternative versions, giving the states options concerning

---

[110] Letter from FTC, *supra* note 106, at 5.

[111] UNIFORM DECEPTIVE TRADE PRACTICES ACT (1966) (cited in HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS AND PROCEEDINGS OF THE ANNUAL CONFERENCE MEETING IN ITS SEVENTY-FIFTH YEAR 299 (1966)) (drafting history of the revised UDTPA does not reveal why the NCCUSL did not address the FTC's concerns).

[112] *Compare* UNIFORM DECEPTIVE TRADE PRACTICES ACT § 3(b) (1964) (cited in HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS AND PROCEEDINGS OF THE ANNUAL CONFERENCE MEETING IN ITS SEVENTY-FIFTH YEAR 263 (1964)) (permitting court to award attorneys' fees only in "exceptional cases" and stating costs "may" be assessed), *with* UNIFORM DECEPTIVE TRADE PRACTICES ACT § 3(b) (1966) (cited in HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS AND PROCEEDINGS OF THE ANNUAL CONFERENCE MEETING IN ITS SEVENTY-FIFTH YEAR 312 (1966)) (permitting court to award attorneys' fees against the plaintiff where action is "groundless," against the defendant where the action was willful, and stating costs "shall" be allowed).

[113] *See* JONATHAN SHELDON & CAROLYN L. CARTER, UNFAIR AND DECEPTIVE ACTS AND PRACTICES 133 & n.174 (6th ed. 2004) (noting that the states adopting the 1966 UDTPA were Georgia, Hawaii, Maine, and Ohio).

[114] *Id.* at 133 & nn.175–76. In 2000, however, the Commissioners withdrew the UDTPA as "obsolete." *Id.*

[115] UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 141–52 (1969)). A model UTP/CPL was initially published in SUGGESTED STATE LEGISLATION for 1967 and subsequently adopted by ten states: Arizona, Kansas, Maryland, Massachusetts, Missouri, New Mexico, Pennsylvania, Rhode Island, Texas, and Vermont. *Id.* at 141–42. This proposal limited coverage to eleven specific kinds of deceptive practice and any others that "similarly" created a likelihood of confusion or misunderstanding. COUNCIL OF STATE GOV'TS, SUGGESTED STATE LEGISLATION, § 1(d), at A-73 (1966). Other states, however, had enacted laws co-extensive with Section 5(a)(1) of the FTC Act, which prohibited all unfair methods of competition and unfair or deceptive acts or practices in trade or commerce. *See id.* As a result, the FTC and others suggested that states should have options in considering the adoption of the UTP/CPL to meet differing state requirements. *Id.* Proposed changes were printed in SUGGESTED STATE LEGISLATION for 1969. *Id.* The final 1970 version of the UTP/CPL was developed jointly by the FTC and the Committee on Suggested State Legislation, and incorporated these changes, as well as several additional modifications. *See id.*

which trade practices to prohibit.[116] For the first time, a model act allowed a private cause of action for damages[117] and authorized class actions where the deceptive practice "caused similar injury to numerous other persons similarly situated."[118] The UTP/CPL also provided for a minimum statutory damages award of $200, regardless of the amount of actual damages.[119]

The states acted quickly. By 1973, forty-four states had enacted consumer protection legislation.[120] Fourteen states adopted the first version of the UTP/CPL.[121] This alternative followed Section 5 of the FTC Act and broadly prohibited all "[u]nfair methods of competition and unfair or deceptive acts or practices in the conduct of any trade of commerce."[122] Thirteen states originally followed the second UTP/CPL alternative,[123] which prohibited all "[f]alse, misleading, or deceptive acts or practices in the conduct of any trade or commerce."[124] Finally, the third version listed twelve specific prohibited practices,[125] plus a "catch-all provision" encom-

---

[116] *Id.* at 146; *see also id.* at 142 (discussing three alternatives); Leaffer & Lipson, *supra* note 61, at 521 n.2.

[117] UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW § 8(a) (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 148 (1969)). Notably, the FTC's original 1966 proposal did not allow a private cause of action. *See* Letter from FTC, *supra* note 106, at 5–10.

[118] UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW § 8(b) (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 149 (1969)).

[119] UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW § 8(a) (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 148–49 (1969)) (allowing recovery of actual damages or $200, whichever is greater).

[120] FEDERAL TRADE COMMISSION, FACT SHEET: STATE LEGISLATION TO COMBAT UNFAIR TRADE PRACTICES (1973) [hereinafter FTC FACT SHEET].

[121] SHELDON & CARTER, *supra* note 113, at 132 & n.162. The jurisdictions adopting the first alternative were Connecticut, Florida, Hawaii, Illinois, Louisiana, Maine, Massachusetts, Montana, Nebraska, North Carolina, South Carolina, Vermont, Washington, and West Virginia. *Id.* As of November 1973, the FTC Fact Sheet listed twelve of these states under the first alternative: Connecticut, Florida, Hawaii, Illinois, Louisiana, Maine, Massachusetts, Montana, North Carolina, South Carolina, Vermont, and Washington. FTC FACT SHEET, *supra* note 120. The FTC Fact Sheet also listed Wisconsin. *Id.* Currently, however, Wisconsin only prohibits "unfair methods of competition," WIS. STAT. ANN. § 100.20(1) (West 2004 & Supp. 2004), and thus is not as broad as the first alternative model.

[122] UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW § 2 (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 146 (1969)).

[123] As of November 1973, these states were Arizona, Arkansas, California, Delaware, Iowa, Kansas, Kentucky, Maryland, Minnesota, Missouri, New Jersey, New York, and North Dakota. FTC FACT SHEET, *supra* note 120. Texas also was patterned after this second alternative, SHELDON & CARTER, *supra* note 113, at 132 n.163, as was Illinois, UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 142 (1969)). Notably, this second option was limited to fraud-based theories and did not cover "unfair practices." UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW § 2 (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 146 (1969)). No state currently uses this second alternative in its exact form. SHELDON & CARTER, *supra* note 113, at 132.

[124] UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW § 2 (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 146 (1969)).

[125] UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW § 2 (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 146–47 (1969)). The

passing "any act or practice which is unfair or deceptive to the consumer."[126] Sixteen states followed this model,[127] with some variation in the itemized list of prohibited practices.[128]

3. Uniform Consumer Sales Practices Act

Finally, in 1970, the NCCUSL proposed another consumer fraud statute: the Uniform Consumer Sales Practices Act.[129] Like the FTC's third alternative in the UTP/CPL,[130] the UCSPA provided an itemized list of prohibited practices, but it also barred "unconscionable act[s] or practice[s]."[131] Three states have used the UCSPA as a model.[132]

Thus, by the early 1970s, nearly every state had enacted a statute[133] designed to prevent consumer fraud.[134] For the most part, these statutes

---

enumerated practices were identical to the NCCUSL's 1964 UDTPA. *Compare* UNIFORM DECEPTIVE TRADE PRACTICES ACT § 2 (1964) (cited in HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS AND PROCEEDINGS OF THE ANNUAL CONFERENCE MEETING IN ITS SEVENTY-FIFTH YEAR 258–62 (1964)), *with* UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW § 2 (alternative form no. 3) (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 146–47 (1969)).

[126] UNFAIR TRADE PRACTICES AND CONSUMER PROTECTION LAW § 2 (alternative form no. 3) (1970) (cited in COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 147 (1969)).

[127] As of November 1973, these states were Alaska, Colorado, Idaho, Indiana, Michigan, Nevada, New Hampshire, New Mexico, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Dakota, Texas, Virginia, and Wyoming. FTC FACT SHEET, *supra* note 120. Currently, nine states appear to follow the third version: Alaska, Georgia, Idaho, Maryland, Mississippi, New Hampshire, Pennsylvania, Rhode Island, and Tennessee. SHELDON & CARTER, *supra* note 113, at 133 & n.165.

[128] SHELDON & CARTER, *supra* note 113, at 132–33 & n.165.

[129] UNIF. CONSUMER SALES PRACTICE ACT historical notes (1970) (The Act was amended in 1971.).

[130] *Id.* § 3(b).

[131] *Id.* § 4. *See generally* David A. Rice, *Critique: Uniform Consumer Sales Practices Act—Damages Remedies: The NCCUSL Giveth and Taketh Away*, 67 NW. U. L. REV. 369 (1972) (analyzing provisions of USCPA).

[132] Ohio, Utah, and Kansas followed this model. UNIF. CONSUMER SALES PRACTICE ACT tbl. of jurisdictions (1970) (noting adoption of UCSPA in Kansas in 1973, Ohio in 1972, and Utah in 1973); *accord* SHELDON & CARTER, *supra* note 113, at 133 & n.171; *see also* KAN. STAT. ANN. § 50-627 (1994 & Supp. 2005) (barring unconscionable acts); OHIO REV. CODE ANN. § 1345.03 (LexisNexis 2002 & Supp. 2005) (same); UTAH CODE ANN. § 13-11-5 (2001 & Supp. 2005) (same).

[133] This Article generally refers to these statutes as "consumer fraud statutes."

[134] Deceptive Trade Practices Act, ALA. CODE §§ 8-19-5 to -15 (LexisNexis 2002 & Supp. 2004); ALASKA STAT. §§ 45.50.471-.561 (2004); ARIZ. REV. STAT. ANN. §§ 44-1521 to -1534 (2003 & Supp. 2004); ARK. CODE ANN. §§ 4-88-101 to -115 (2001 & Supp. 2003); Consumer Legal Remedies Act, CAL. CIV. CODE §§ 1750-1785 (West 1998 & Supp. 2005); Unfair Competition Law, CAL. BUS. & PROF. CODE §§ 17200 to 17209 (West 1997 & Supp. 2005); Colorado Consumer Protection Act, COLO. REV. STAT. §§ 6-1-101 to -1001 (2004); CONN. GEN. STAT. ANN. §§ 42-110a to -110q (West 2000 & Supp. 2005); DEL. CODE ANN. tit. 6, §§ 2511-2527 (1999); DEL. CODE ANN. tit. 6, §§ 2531-2536 (1999) (deceptive trade practices); D.C. CODE ANN. §§ 28-3901 to -3911 (LexisNexis 2001 & Supp. 2005); Florida Deceptive and Unfair Trade Practices Act, FLA. STAT. ANN. §§ 501.201-.213 (West 2002 & Supp. 2005); Fair Business Practices Act of 1975, GA. CODE ANN.

sought to strengthen public enforcement of consumer protection. Nearly every state placed enforcement authority in the state's attorney general.[135]

---

§§ 10-1-390 to -407 (2000 & Supp. 2005); Uniform Deceptive Trade Practices Act, GA. CODE ANN. §§ 10-1-371 to -374 (2000 & Supp. 2005); HAW. REV. STAT. ANN. §§ 480-1 to -24 (LexisNexis 2002 & Supp. 2004); Uniform Deceptive Trade Practices Act, HAW. REV. STAT. ANN. § 481A-1 to -5 (LexisNexis 2002); Idaho Consumer Protection Act, IDAHO CODE ANN. §§ 48-601 to -619 (1997); Consumer Fraud and Deceptive Business Practices Act, 815 ILL. COMP. STAT. ANN. 505/1-12 (West 1999 & Supp. 2005); Uniform Deceptive Trade Practices Act, 815 ILL. COMP. STAT. ANN. 510/1-7 (West 1999 & Supp. 2005); IND. CODE ANN. §§ 24-5-0.5-1 to -12 (West 1995 & Supp. 2005); IOWA CODE ANN. § 714.16 (West 2003); Kansas Consumer Protection Act, KAN. STAT. ANN. §§ 50-623 (1994 & Supp. 2004); Consumer Protection Act, KY. REV. STAT. ANN. §§ 367.110-.370 (West 2002); Unfair Trade Practices and Consumer Protection Law, LA. REV. STAT. ANN. §§ 51:1401-1430 (2003 & Supp. 2005); Maine Unfair Trade Practices Act, ME. REV. STAT. ANN. tit. 5, §§ 205-A to 214 (2002 & Supp. 2004); Uniform Deceptive Trade Practices Act, ME. REV. STAT. ANN. tit. 10, §§ 1211-1216 (2002 & Supp. 2004); Maryland Consumer Protection Act, MD. CODE ANN., COM. LAW §§ 13-101 to -501 (LexisNexis 2000 & Supp. 2004); MASS. GEN. LAWS ANN. ch. 93A, §§ 1-11 (West 2005); Michigan Consumer Protection Act, MICH. COMP. LAWS ANN. §§ 445.901-922 (West 2002 & Supp. 2005); Unlawful Trade Practices Act, MINN. STAT. ANN. §§ 325D.09-.16 (West 2004 & Supp. 2005); Uniform Deceptive Trade Practices Act, MINN. STAT. ANN. §§ 325D.43-.48 (West 2004 & Supp. 2005); Prevention of Consumer Fraud Act, MINN. STAT. ANN. §§ 325F.68-.70 (West 2004 & Supp. 2005); MISS. CODE ANN. §§ 75-24-1 to -27 (West 1999 & Supp. 2004); MO. ANN. STAT. §§ 407.010-.307 (West 2001 & Supp. 2005); Montana Unfair Trade Practices and Consumer Protection Act of 1973, MONT. CODE ANN. §§ 30-14-101 to -143 (2003); Consumer Protection Act, NEB. REV. STAT. §§ 59-1601 to -1623 (2004); Uniform Deceptive Trade Practices Act, NEB. REV. STAT. §§ 87-301 to -306 (2004); NEV. REV. STAT. ANN. §§ 598.0903-0999 (LexisNexis 2004); N.H. REV. STAT. ANN. §§ 358-A:1-:13 (2004); N.J. STAT. ANN. §§ 56:8-1 to -106 (West 2001 & Supp. 2005); Unfair Practices Act, N.M. STAT. §§ 57-12-1 to -22 (2000 & Supp. 2003); N.Y. GEN. BUS. LAW §§ 349 to 349-c (McKinney 2004); N.C. GEN. STAT. §§ 75-1 to -49 (2003); N.D. CENT. CODE §§ 51-10-01 to -15 (1999); N.D. CENT. CODE §§ 51-15-01 to -11 (1999 & Supp. 2003); OHIO REV. CODE ANN. §§ 1345.01-.13 (LexisNexis 2002 & Supp. 2005); Oklahoma Consumer Protection Act, OKLA. STAT. ANN. tit. 15, §§ 751-789 (West 1993 & Supp. 2005); Oklahoma Deceptive Trade Practices Act, OKLA. STAT. ANN. tit. 78, §§ 51-55 (West 2002 & Supp. 2005); OR. REV. STAT. ANN. §§ 646.605-.656 (West 1983 & Supp. 1998); Unfair Trade Practices and Consumer Protection Law, 73 PA. STAT. ANN. §§ 201-1 to 201-9.2 (West 1993 & Supp. 2005); Unfair Trade Practice and Consumer Protection Act, R.I. GEN. LAWS §§ 6-13.1-1 to -27 (2001); South Carolina Unfair Trade Practices Act, S.C. CODE ANN. §§ 39-5-10 to -160 (1976 & Supp. 2004); S.D. CODIFIED LAWS §§ 37-24-1 to -35 (2000); Tennessee Consumer Protection Act of 1977, TENN. CODE ANN. §§ 47-18-101 to -125 (2001 & Supp. 2004); Deceptive Trade Practices-Consumer Protection Act, TEX. BUS. & COM. CODE ANN. §§ 17.41-.63 (Vernon 2002 & Supp. 2005); Utah Consumer Sales Practices Act, UTAH CODE ANN. §§ 13-11-1 to -23 (2001 & Supp. 2004); VT. STAT. ANN. tit. 9, §§ 2451-2480m (1993 & Supp. 2004); Virginia Consumer Protection Act of 1977, VA. CODE ANN. §§ 59.1-196 to -207 (2001 & Supp. 2005); Consumer Protection Act, WASH. REV. CODE ANN. §§ 19.86.010 to -.86.920 (West 1999 & Supp. 2005); West Virginia Consumer Credit and Protection Act, W. VA. CODE ANN. §§ 46A-6-101 to -110 (LexisNexis 1999); WIS. STAT. ANN. § 100.18 to 100.183, 100.20 (West 2004 & Supp. 2004); Wyoming Consumer Protection Act, WYO. STAT. ANN. §§ 40-12-101 to -114 (2005).

In addition to these general deceptive practices statutes, several states also have specific false advertising statutes. CAL. BUS. & PROF. CODE §§ 17500-17594 (West 1997 & Supp. 2005); GA. CODE ANN. §§ 10-1-420 to -427 (2000 & Supp. 2005); MINN. STAT. ANN. § 325F.67 (West 2004); N.M. STAT. §§ 57-15-1 to -10 (2000); N.Y. GEN. BUS. LAW §§ 350-350-f-1 (McKinney 2004); N.D. CENT. CODE §§ 51-12-01 to -14 (1999); VA. CODE ANN. § 18.2-216 (2004 & Supp. 2005).

[135] FTC FACT SHEET, *supra* note 120. In fifteen states, enforcement authority was

Indeed, several states established official "Consumer Protection Departments" or consumer counsel positions within the attorney general's office.[136] As of 1971, however, only eight states allowed private causes of action for damages.[137] Over the next decade, however, this landscape would change as states slowly amended their consumer fraud statutes to allow private damages actions.[138]

III. RELIANCE AND CAUSATION STANDARDS FOR PUBLIC AND PRIVATE ENFORCEMENT

The history of lax enforcement and the perceived need for new laws to protect consumers created an environment for broad judicial interpretations of consumer fraud statutes. To advance the government's public enforcement role in stopping fraud before the consumer is harmed, courts relaxed traditional fraud requirements for particular types of relief.[139] Thus, where the government sought injunctive relief to stop incipient fraud, courts held that government agencies do not need to prove consumer injury or consumer reliance.[140] The broad interpretation of public enforcement provisions strongly influenced how courts applied consumer fraud statutes to private causes of action. Thus, with little thought given to the different purposes of public and private actions, state courts incorporated the deterrence objective of government suits for injunctions[141] and loosened the traditional reliance-causation requirement in private claims for damages.[142] Thus, today's misrepresentation case—where no one may have been actually misled by the manufacturer's statement—was born.

---

shared between the attorney general and a local enforcement agency, such as the district, county or city attorney. *Id.*

[136] Lovett, *State Deceptive Trade Practice Legislation*, *supra* note 61, at 730; *accord* William A. Lovett, *Private Actions for Deceptive Trade Practices*, 23 ADMIN. L. REV. 271, 275 (1971) [hereinafter Lovett, *Private Actions*] (noting creation of state consumer protection agencies in several states). Some local jurisdictions, like New York City, even created their own consumer protection divisions. *E.g.*, NEW YORK CITY LAW NO. 83 (1969) (creating Department of Consumer Affairs), *available in* PRACTISING LAW INST., *supra* note 57, at 277–79.

[137] Lovett, *Private Actions*, *supra* note 136, at 275–76 & n.6 (stating that California, Colorado, Hawaii, Massachusetts, North Carolina, Rhode Island, Vermont, and Washington allowed private damages actions); *see also id.* at 282–83 (discussing requirements of these eight private action provisions).

[138] By November 1973, thirty-one states had adopted private cause of action provisions. FTC Fact Sheet, *supra* note 120. Washington, for example, added a private cause of action in 1970. Jennifer Rust Murray, *Proving Cause in Fact Under Washington's Consumer Protection Act: The Case for a Rebuttable Presumption of Reliance*, 80 WASH. L. REV. 245, 245 & n.3 (2005). Illinois added a private cause of action provision in 1973. *E.g.*, Oliveira v. Amoco Oil Co., 776 N.E.2d 151, 160 (Ill. 2002). Pennsylvania did not add a private cause of action provision until 1976. *E.g.*, Weinberg v. Sun Co., 777 A.2d 442, 446 (Pa. 2001).

[139] *See infra* Part III.A.

[140] *See infra* Part III.A.

[141] *See infra* Part III.B.2.

[142] *See infra* Part III.B.2.

A. Government Standards of Reliance and Causation

Government enforcement standards reflect the government's role in preventing incipient fraud. Section 5 of the FTC Act prohibits "unfair or deceptive acts or practices in or affecting commerce."[143] To establish a violation of Section 5, the FTC does not need to prove consumer reliance or an injury to the consumer.[144] Rather, courts have recognized a lower standard in order to effectuate the public purpose of the FTC Act and to allow the FTC to take preemptive action against deceptive practices.[145] Moreover, the FTC is authorized to seek injunctive relief whenever it "has reason to believe that any person, partnership or corporation is violating, or is about to violate, any provision of law enforced by the Federal Trade Commission," and an injunction "would be in the interest of the public."[146] No reliance-causation—or even injury for that matter—is required. Accordingly, to obtain injunctive relief[147] under the FTC Act, the FTC need only show that the misrepresentation was "likely to mislead" the consumer.[148]

Often overlooked, however, is the higher burden that the FTC faces when seeking consumer redress.[149] In a consumer redress claim, the FTC seeks a monetary award that it then uses to provide refunds to affected

---

[143] 15 U.S.C. § 45(a) (2000).

[144] *E.g.*, FTC v. Freecom Commc'ns, Inc., 401 F.3d 1192, 1203 (10th Cir. 2005); *see also* Robert Pitofsky, *Beyond Nader: Consumer Protection and the Regulation of Advertising*, 90 HARV. L. REV. 661, 676–77 (1977) (noting that "issues of . . . causality relating to whether consumers were influenced in purchasing decisions by the false claim are largely avoided by the Commission rules that it need show only capacity to deceive rather than actual deception, and capacity to affect purchasing decisions rather than actual effects").

[145] *See, e.g.*, *Freecom Commc'ns Inc.*, 401 F.3d at 1203; *accord* Regina Corp. v. FTC, 322 F.2d 765, 768 (3d Cir. 1963) ("The purpose of the Federal Trade Commission Act is to protect the public . . . and it is in the public interest to stop any deception at its incipiency.") (internal citations omitted).

[146] 15 U.S.C. § 53(b) (2000).

[147] Apart from enjoining the misrepresentations, the FTC also can order corrective advertising. *See, e.g.*, Warner-Lambert Co. v. FTC, 562 F.2d 749, 752 (D.C. Cir. 1977) (upholding injunction prohibiting company from representing that Listerine helps prevents colds and sore throats and requiring future advertising to state that "Listerine will not help prevent colds or sore throats or lessen their severity").

[148] *Freecom Commc'ns Inc.*, 401 F.3d at 1203.

[149] The FTC has the authority to seek injunctive relief under Section 13(b) of the FTC Act. 15 U.S.C. § 53(b). Courts have found that this provision authorizes the FTC to seek other equitable remedies such as disgorgement and consumer redress. *See, e.g.*, FTC v. Amy Travel Serv., Inc., 875 F.2d 564, 572 (7th Cir. 1989) ("[T]he statutory grant of authority to the district court to issue permanent injunctions includes the power to order any ancillary equitable relief necessary to effectuate the exercise of the granted powers."); *see also* FTC v. Sec. Rare Coin & Bullion Corp., 931 F.2d 1312, 1314 (8th Cir. 1991) (finding district court has authority to order consumer redress based on its "inherent equitable powers"); FTC v. H. N. Singer, Inc., 668 F.2d 1107, 1113 (9th Cir. 1982) (establishing principle that district court's inherent powers authorize ancillary relief under Section 13(b)). In addition, the FTC Act provides explicit authorization for consumer redress when a defendant violates a cease and desist order or FTC rule. 15 U.S.C. § 57b(b) (2000).

consumers.[150] While proof of consumer reliance and injury is not necessary to establish a violation of the FTC Act—and the government's accompanying right to injunctive relief—proof of reliance is required when the FTC seeks a monetary consumer redress award.[151]

Like the FTC, state attorneys general typically do not need to prove consumer reliance when establishing that a practice violates the state deceptive practices act.[152] Indeed, many state statutes provide that a practice may violate the consumer fraud statute, "whether or not any person has in fact been misled, deceived, or damaged thereby."[153] Thus, like the FTC, a state attorney general can establish that a misrepresentation violates a state consumer fraud statute simply by showing that the misrepresentation is likely to mislead consumers. Actual deception or reliance is not required.[154]

State courts, however, have taken different approaches to government claims for restitution. In a minority of jurisdictions, while the state does not need to show reliance to obtain a general restitution order, consumers do have to show reliance to obtain money from the restitution award.[155] Mostly, however, courts have applied the lower "no reliance-causation" standard to government restitution claims.[156] Two main reasons underlie this approach. First, as a matter of statutory construction, the statutory authorizations for equitable relief, including restitution, do not contain any

---

[150] *See generally* STEPHANIE W. KANWIT, FEDERAL TRADE COMMISSION § 21:6 (2003) (explaining FTC distribution of redress award).

[151] *Freecom Commc'ns Inc.*, 401 F.3d at 1205. To prove reliance, however, the FTC does not need to show that a particular consumer actually relied on and was injured by the misrepresentation. *Id.*

[152] *E.g.*, Consumer Prot. Div. Office of the Att'y Gen. v. Consumer Publ'g Co., 501 A.2d 48, 68–69 (Md. 1985). *See also* Michael S. Greve, *Consumer Law, Class Actions, and the Common Law*, 7 CHAP. L. REV. 155, 174 (2004) (discussing the "strengthen[ed] hand" of state attorneys general and finding that state enforcement typically does not require consumer reliance).

[153] *E.g.*, ARIZ. REV. STAT. ANN. § 44-1522(A) (2003 & Supp. 2004). Eleven states plus the District of Columbia include this or similar language in their consumer fraud statute. ARIZ. REV. STAT. ANN. § 44-1522(A) (2003 & Supp. 2004); DEL. CODE ANN. tit. 6, § 2513 (1999); D.C. CODE ANN. § 28-3904 (LexisNexis 2001 & Supp. 2005); 815 ILL. COMP. STAT. ANN. 505/2 (West 1999); IOWA CODE ANN. § 714.16(2) (West 2003); KAN. STAT. ANN. § 50-626(b) (1994 & Supp. 2004); MD. CODE ANN., COM. LAW § 13-302 (LexisNexis 2000); MINN. STAT. ANN. § 325F.69(subd. 1) (West 2004 & Supp. 2005); N.J. STAT. ANN. § 56:8-2 (West 2001 & Supp. 2005); N.D. CENT. CODE § 51-15-02 (1999 & Supp. 2005); S.D. CODIFIED LAWS § 37-24-6 (2003); W. VA. CODE ANN. § 46A-6-102(7)(M) (LexisNexis 1999 & Supp. 2005).

[154] *E.g.*, *Consumer Prot. Div. Office of the Att'y Gen.*, 501 A.2d at 68 ("In not requiring proof of actual deception or harm to consumers, the [Maryland] Consumer Protection Act follows the practice of the Federal Trade Commission.").

[155] *See, e.g.*, Consumer Prot. Div. v. Morgan, 874 A.2d 919, 943 (Md. 2005); State *ex rel.* Kidwell v. Master Distrib., Inc., 615 P.2d 116, 126 (Idaho 1980) (finding restitution order must include "procedure by which consumer claims may be efficiently and fairly processed").

[156] *See, e.g.*, Telcom Directories, Inc. v. Commonwealth, 833 S.W.2d 848, 850 (Ky. Ct. App. 1992) (noting that "proof of actual deception" not required for state restitution order); People v. Toomey, 203 Cal. Rptr. 642, 658–59 (Cal. Ct. App. 1984) ("[R]eliance and actual damages are not necessary elements to [a restitution] award.").

causal language or injury requirement.[157] More importantly, this approach focuses on the deterrent function of a restitution award: "Restitution is not intended to benefit the tendees by the return of money, but instead is designed to penalize a defendant for past unlawful conduct and thereby deter future violations."[158]

B. Private Enforcement Standards of Reliance and Causation

By embracing the deterrent function of government actions, some courts took the next step towards the modern misrepresentation class action—the wholesale application of government enforcement standards to private damages actions.[159] This approach, however, ignores the different purpose of a private damages action.

Unlike the FTC Act,[160] nearly every state[161] allows a private damages action for misrepresentation claims.[162] In addition to a deceptive practices

---

[157] *E.g.*, *Toomey*, 203 Cal. Rptr. at 658 (noting statute authorizes court to order defendant "to restore . . . any money or property . . . which may have been acquired by means of any [unlawful practice]") (alteration in original).

[158] *Id.* at 658–59.

[159] *See infra* notes 201–208 and accompanying text.

[160] *See, e.g.*, Moore v. N.Y. Cotton Exch., 270 U.S. 593, 603 (1926) (refusing to allow private claim under FTC Act); Holloway v. Bristol-Myers Corp., 485 F.2d 986, 988–89 (D.C. Cir. 1973) (holding that the FTC Act does not provide express private cause of action and refusing to imply private remedy). *But see* Guernsey v. Rich Plan of the Midwest, 408 F. Supp. 582, 586–88 (N.D. Ind. 1976) (recognizing private right of action where alleged conduct was subject to prior cease and desist order issued by FTC). In 1978, the FTC attempted to create a private cause of action under its rules governing franchise disclosure requirements. *See* 16 C.F.R. § 436.1–.3 (2005). In its "Statement of Basis and Purpose" accompanying this rule, the FTC stated its belief that a private cause of action should exist under the franchise disclosure rules. Final Trade Regulation Rule, 43 Fed. Reg. 59,614, 59,723 (Dec. 21, 1978). Still, courts refused to imply a private remedy, holding that the FTC's statement did not provide evidence of changed congressional intent. *See, e.g.*, Freedman v. Meldy's, Inc., 587 F. Supp. 658, 661–62 (E.D. Pa. 1984) (holding that FTC's intent was insufficient to imply a private cause of action).

[161] Iowa and North Dakota are the only two states that do not allow a private damages action. Only the Attorney General can bring suit under Iowa's consumer fraud statute. Molo Oil Co. v. River City Ford Truck Sales, Inc., 578 N.W.2d 222, 227–28 (Iowa 1998) (finding no implied private right of action under Iowa Code § 714.16). Iowa does provide a limited private right of action under its Consumer Credit Code. *See* IOWA CODE ANN. § 537.5201 (West 1997 & Supp. 2005) (listing specific violations, such as improper credit charges, that permit civil actions for damages). North Dakota allows a private civil suit, but only for injunctive relief. N.D. CENT. CODE § 51-10-06 (1999) (unfair practices); *id.* at § 51-12-14 (1999) (false advertising). The North Dakota Supreme Court has refused to imply a private right of action for damages. Trade 'N Post, L.L.C. v. World Duty Free Americas, Inc., 628 N.W.2d 707, 708, 710–12 (N.D. 2001).

[162] ALA. CODE § 8-19-10 (LexisNexis 2002 & Supp. 2004); ALASKA STAT. § 45.50.531(a) (2004); ARK. CODE ANN. § 4-88-113(f) (2001 & Supp. 2005); CAL. CIV. CODE § 1780(a) (West 1998 & Supp. 2005); COLO. REV. STAT. § 6-1-113(1) (2004); CONN. GEN. STAT. ANN. § 42-110g (West 2000 & Supp. 2005); DEL. CODE ANN. tit. 6, § 2525 (1999); D.C. CODE ANN. § 28-3905(k)(1) (LexisNexis 2001 & Supp. 2005); FLA. STAT. ANN. § 501.211(1) (West 2002 & Supp. 2005); GA. CODE ANN. § 10-1-399(a) (2000 & Supp. 2005); HAW. REV. STAT. ANN. § 480-13(b) (LexisNexis 2002 & Supp. 2004); IDAHO CODE ANN. § 48-608(1) (2000 & Supp. 2005); 815 ILL. COMP. STAT. ANN. 505/10a (West 1999 & Supp.

violation, a private plaintiff must satisfy the requirements of this private cause of action provision.[163] Adopting the deterrence purpose of government enforcement, state courts have loosened the traditional reliance-causation requirement,[164] even though the statutory text leaves no doubt that traditional tort causation principles apply to private damages actions.[165]

New York—the jurisdiction governing the McDonald's case[166]—illustrates this confusion. Under New York law, two statutes govern deceptive advertising: (1) the false advertising statute[167] and (2) the deceptive practices act.[168] To bring a damages claim under the false advertising statute, a private plaintiff must show reliance.[169] But when the plaintiff brings the exact same false advertising claim under the deceptive practices act,[170] courts no longer require reliance.[171] Both statutes contain identical causation requirements—only persons "injured by reason of" a deceptive state-

---

2005); 815 ILL. COMP. STAT. ANN. 510/3 (West 1999) (limited to injunctive relief); IND. CODE ANN. § 24-5-0.5-4(a) (West 1995 & Supp. 2005); IOWA CODE ANN. § 714.16(7) (West 2003 & Supp. 2005); KAN. STAT. ANN. § 50-634(a) (1994 & Supp. 2004); KY. REV. STAT. ANN. § 367.220(1) (West 2004); LA. REV. STAT. ANN. § 51:1409A (2003 & Supp. 2005); ME. REV. STAT. ANN. tit. 5, § 213 (2002 & Supp. 2005); MD. CODE ANN., COM. LAW § 13-408(a) (LexisNexis 2000 & Supp. 2003); MASS. GEN. LAWS ANN. ch. 93A, § 9(1) (West 2005); MICH. COMP. LAWS ANN. § 445.911(2) (West 2002); MINN. STAT. ANN. § 8.31(3a) (West 2005); MISS. CODE ANN. § 75-24-15(1) (West 1999 & Supp. 2004); MO. ANN. STAT. § 407.025 (West 2001 & Supp. 2005); MONT. CODE ANN. § 30-14-133(1) (2003); NEB. REV. STAT. § 59-1609 (2004); NEV. REV. STAT. ANN. § 41.600 (LexisNexis 2002); N.H. REV. STAT. ANN. § 358-A:10 (1995 & Supp. 2004); N.J. STAT. ANN. § 56:8-19 (West 2001 & Supp. 2005); N.M. STAT. ANN. § 57-12-10 (2000 & Supp. 2003); N.Y. GEN. BUS. LAW § 349(h) (McKinney 2004); N.C. GEN. STAT. § 75-16 (2003); OHIO REV. CODE ANN. § 1345.09 (LexisNexis 2002 & Supp. 2005); OKLA. STAT. ANN. tit. 15, § 761.1 (West 1993 & Supp. 2005); OR. REV. STAT. ANN. § 646.638(1) (West 1983 & Supp. 1998); 73 PA. STAT. ANN. § 201-9.2 (West 1993 & Supp. 2005); R.I. GEN. LAWS § 6-13.1-5.2(a) (2001 & Supp. 2004); S.C. CODE ANN. § 39-5-140(a) (1985 & Supp. 2004); S.D. CODIFIED LAWS § 37-24-31 (2000); TENN. CODE ANN. § 47-18-109(a)(1) (2001 & Supp. 2004); TEX. BUS. & COM. CODE ANN. § 17.50(a) (Vernon 2002 & Supp. 2004); UTAH CODE ANN. § 13-11-19 (2001); VT. STAT. ANN. tit. 9, § 2461(b) (1993 & Supp. 2003); VA. CODE ANN. § 59.1-204 (2001 & Supp. 2005); WASH. REV. CODE ANN. § 19.86.090 (West 1999 & Supp. 2005); W. VA. CODE ANN. § 46A-6-106 (LexisNexis 1999); WIS. STAT. ANN. § 100.20(5) (West 2004 & Supp. 2004); WYO. STAT. ANN. § 40-12-108(a) (2005). In addition, Arizona allows an implied private right of action under its consumer fraud statute. Sellinger v. Freeway Mobile Home Sales, Inc., 521 P.2d 1119, 1122 (Ariz. 1974) (en banc) (establishing that Arizona's consumer fraud statute implies a private right of action); *accord* Holeman v. Neils, 803 F. Supp. 237, 242 (D. Ariz. 1992) (noting that Arizona consumer fraud statute provides an implied right of action).

[163] *See supra* note 162.

[164] *See infra* notes 193–215 and accompanying text.

[165] *See infra* Part V.D.

[166] *See supra* notes 24–25 and accompanying text.

[167] N.Y. GEN. BUS. LAW § 350 (McKinney 2004).

[168] N.Y. GEN. BUS. LAW § 349 (McKinney 2004).

[169] *E.g.*, Small v. Lorillard Tobacco Co., 679 N.Y.S.2d 593, 599 (N.Y. App. Div. 1998) ("[I]ndividualized proof of reliance is essential to the causes of action for false advertising under [General Business Law] § 350."), *aff'd*, 720 N.E.2d 892 (N.Y. 1999).

[170] *E.g.*, *id.* (discussing false advertising claim brought under deceptive practices statute).

[171] Stutman v. Chem. Bank, 731 N.E.2d 608, 611–12 (N.Y. 2000).

ment can proceed.[172] Yet, one statute requires reliance-causation, while the other does not. This confusion in New York reflects the contradictory approaches taken throughout the nation.

1. Reliance-Causation Is Required

Remarkably, only a few state courts have recognized that, in a misrepresentation case, causation and reliance are essentially the same thing.[173] These courts recognize that, as a practical matter, damages cannot be "caused" by a defendant's misrepresentation without reliance on the statement.[174] In other words, if the defendant's statement did not have some influence on the plaintiff's decision to purchase the product, then it did not cause her any harm. The Minnesota Supreme Court, for example, took this approach in *Group Health Plan, Inc. v. Philip Morris Inc.*[175] While the analysis in that case essentially acknowledged that causality requires reliance, the court faltered in its application of this principle.

In *Group Health Plan, Inc.*, a group of health maintenance organizations brought suit against various tobacco companies under three of Minnesota's consumer fraud statutes.[176] On certification from the federal district court,[177] the Minnesota Supreme Court considered whether reliance was required under the state's consumer fraud statute.[178]

Unlike other courts,[179] the court correctly recognized the distinction between the elements necessary to establish a statutory violation and the additional elements necessary to satisfy the private cause of action provision.[180] Like many states, Minnesota's misrepresentation statute provided that any misrepresentation violated the act "whether or not any person has in fact been misled, deceived, or damaged thereby."[181] The court found that this language established the standard for a statutory violation: a mis-

---

[172] N.Y. GEN. BUS. LAW § 349-h (McKinney 2004); N.Y. GEN. BUS. LAW § 350-e (McKinney 2004).

[173] *See infra* notes 175–186, 190–192 and accompanying text.

[174] *See infra* notes 175–186, 190–192 and accompanying text.

[175] 621 N.W.2d 2 (Minn. 2001).

[176] *Id.* at 4. These statutes were: MINN. STAT. ANN. §§ 325D.13 (Unlawful Trade Practices Act), 325F.67 (false statement in advertising provision), 325F.69(1) (prevention of consumer fraud provision) (West 2004 & Supp. 2005). *Id.* at 3.

[177] The HMOs filed suit in the United States District Court for the District of Minnesota. Following motions to dismiss by the tobacco manufacturers, the federal district court certified two questions to the Minnesota Supreme Court concerning the Minnesota statutes: (1) whether a plaintiff must be a purchaser of the defendant's products; and (2) whether a plaintiff must plead and prove reliance on the defendant's statements or conduct. *Group Health Plan, Inc.*, 621 N.W.2d at 4; *see also* Group Health Plan, Inc. v. Philip Morris Inc., 188 F. Supp. 2d 1122, 1125 (D. Minn. 2002) (noting certification followed motions to dismiss).

[178] *Group Health Plan, Inc.*, 621 N.W.2d at 4–5.

[179] *See infra* notes 193–215 and accompanying text.

[180] *Group Health Plan, Inc.*, 621 N.W.2d at 12–13.

[181] MINN. STAT. ANN. § 325F.69(1) (West 2004); *see supra* note 153 (listing states with identical provision).

representation could violate the statute, and subject the manufacturer to government enforcement, without any showing of consumer reliance.[182]

The court then turned to the requirements of the private cause of action provision. The court focused on the provision's causation requirement, which allowed a damages action only by someone "injured by" a violation.[183] The court concluded that the required causal nexus between a defendant's misrepresentation and a plaintiff's injury could be shown only by reliance:

> [W]here, as here, the plaintiffs allege that their damages were caused by deceptive, misleading, or fraudulent statements or conduct . . . , as a practical matter it is not possible that the damages could be caused by a violation without reliance on the statements or conduct alleged to violate the statutes. Therefore, in a case such as this, it will be necessary to prove reliance on those statements or conduct to satisfy the causation requirement.[184]

Thus, although allegations of reliance were not necessary to violate the statute,[185] reliance was required to recover damages.[186]

---

[182] *See Group Health Plan, Inc.*, 621 N.W.2d at 12. Other states similarly have distinguished between public enforcement of a violation and a private claim for damages. *E.g.*, CitaraManis v. Hallowell, 613 A.2d 964, 969 (Md. 1992). In *CitaraManis*, the Maryland Court of Appeals explained:

> In a public enforcement proceeding "[a]ny practice prohibited by this title is a violation . . . whether or not any consumer in fact has been misled, deceived, or damaged as a result of that practice." In contrast, a private enforcement proceeding pursuant to § 13-408(a) expressly only permits a consumer "to recover for injury or loss sustained by him as the result of a practice prohibited by this title." Section 13-408(a), therefore, requires an aggrieved consumer to establish the nature of the actual injury or loss that he or she has allegedly sustained as a result of the prohibited practice. This statutory construction creates a bright line distinction between the public enforcement remedies available under the CPA, and the private remedy available under § 13-408(a).

Id. (citations omitted); *see also* Weinberg v. Sun Co., 777 A.2d 442, 445–46 (Pa. 2001) (noting that "tendency to deceive" language is a "consideration[ ] appropriate for a high public official responsible for protecting public interests").

[183] *Group Health Plan, Inc.*, 621 N.W.2d at 13; *see also* MINN. STAT. ANN. § 8.31(3a) (West 2005).

[184] *Group Health Plan, Inc.*, 621 N.W.2d at 13; *accord, e.g.*, Hageman v. Twin City Chrysler-Plymouth Inc., 681 F. Supp. 303, 308 (M.D.N.C. 1988) ("To prove actual causation, a plaintiff must prove that he or she detrimentally relied on the defendant's deceptive statement or misrepresentation.") (citing Pearce v. Am. Defender Life Ins. Co., 343 S.E.2d 174, 180 (N.C. 1986)); Feitler v. Animation Celection, Inc., 13 P.3d 1044, 1047 (Or. Ct. App. 2000) (holding causal element of misrepresentation claim requires reliance by the consumer); *cf.* Siemer v. Assocs. First Capital Corp., No. CV 97-281 TUC JMR (JCC), 2001 U.S. Dist. LEXIS 12810, at *12 (D. Ariz. Mar. 29, 2001) ("The injury element of the [state consumer protection statute] claim occurs when the consumer relies on the misrepresentations . . . .").

[185] *Group Health Plan, Inc.*, 621 N.W.2d at 12.

[186] *Id.* at 13.

Based on this analysis, however, the Minnesota Supreme Court mistakenly leapt to the conclusion that a private plaintiff was not required to allege reliance as an element.[187] The fact that a defendant can violate the statute (i.e., make a misrepresentation) without any reliance does not mean that a private plaintiff does not need to plead reliance to state a claim for damages. The Minnesota Supreme Court, however, skipped over this step and allowed a plaintiff to state a claim without allegations of reliance.[188] This error has led to easier certification of consumer class actions in Minnesota.[189]

Other states, however, correctly have recognized that reliance-causation is an essential element and do not allow a plaintiff to state a claim without allegations of reliance. In *Campbell v. Beak*,[190] for example, the Georgia Court of Appeals held that the state consumer fraud statute incorporates the "'reliance' element of the common law tort of misrepresentation into the causation element."[191] Likewise, the Pennsylvania Supreme Court has explained that the statute's causation language—"as a result of"—means that a plaintiff must allege reliance.[192]

---

[187] *Id.* at 12. The court explained its analysis as follows:

> The tobacco companies argue that, to the extent the legislature eliminated elements of common law fraud from a statutory misrepresentation action, it did so only for claims seeking injunctive relief, not damages. They point out that the express language eliminating the element of reliance is found only in one of the substantive statutes and that statute authorizes only injunctive relief. The tobacco companies argue that, because the statute that authorizes actions for damages, Minn. Stat. § 8.31, subd. 3a, contains no similar express exemption from the reliance requirement, there is therefore no such exemption intended for a damages action . . . . But subdivision 3a authorizes a damages action for a violation of the substantive statutes. As explained, those statutes define what constitutes a violation, and they do so in a manner that indicates that reliance is not a separate element of a violation. We will not read an element into a statutory claim that the legislature has not articulated and, to the contrary, has indicated should be eliminated.

Id.

[188] *Id.*

[189] *See, e.g.*, Curtis v. Philip Morris Cos., No. PI 01-018042, 2004 WL 2776228, at *3–*5 (D. Minn. Nov. 29, 2004). In *Curtis*, the federal district court certified a statewide class action of all Marlboro Lights purchasers. *Id.* at *5. Relying on *Group Health Plan, Inc.*, the court found that plaintiffs' allegation that "the lengthy course of misrepresentations concerning 'light' cigarettes, which affected a large number of Minnesota cigarette consumers, is sufficient evidence of reliance at this stage of the proceedings." *Id.* at *3.

[190] 568 S.E.2d 801 (Ga. Ct. App. 2002).

[191] *Id.* at 805.

[192] *E.g.*, Weinberg v. Sun Co., 777 A.2d 442, 446 (Pa. 2001) (finding that "[n]othing in the legislative history suggests that the legislature ever intended statutory language directed against consumer fraud to do away with the traditional common law elements of reliance and causation"); *see also* Toy v. Metro. Life Ins. Co., 863 A.2d 1, 11 (Pa. Super. Ct. 2004) (holding that *Weinberg* applies to all private actions under Pennsylvania's UTP/CPL).

2. Reliance-Causation Abandoned by the Courts

Taking an alternative view, courts in some other states have abandoned any causation requirement by eliminating reliance.[193] These decisions are premised on two underlying assumptions: (1) that the government's historical ineffectiveness justifies a broad private remedy,[194] and (2) that public enforcement standards also apply to a private damages claim.[195]

Underpinning these decisions is the historical context of the private cause of action provision—an era when government enforcement was lax or non-existent.[196] In *Slaney v. Westwood Auto, Inc.*,[197] for example, the Massachusetts Supreme Judicial Court based its decision to abandon the reliance element on the ineffectiveness of government enforcement in the 1960s.[198] The court repeated the same criticisms that had been leveled at the FTC of the 1960s: the state agencies were "understaffed and underfinanced, morassed in a sea of red tape, and unbearably slow acting."[199] Thus, the court justified an expansive private remedy based, in part, on the lack of government resources to obtain relief for the consumer.[200]

Given the perception of lax government enforcement, courts typically have ignored any distinction between the public enforcement provisions of the consumer fraud statute and the private cause of action provisions. Many courts thus blindly relied on the standards for violating the act instead of differentiating the separate requirements of the private cause of action provision.[201] A deceptive practice can violate a consumer fraud statute "whether or not any person has in fact been misled, deceived, or damaged thereby."[202] Courts recognized that this language signaled the legislature's intent to make it easier for the government to sue for statutory consumer fraud than it had been to sue for common law fraud.[203] Many states then applied this language and its lower threshold to the private damages

---

[193] *See infra* notes 197–215 and accompanying text.

[194] *See infra* notes 197–200 and accompanying text.

[195] *See infra* notes 201–208 and accompanying text.

[196] *See supra* Part II.A.

[197] 322 N.E.2d 768 (Mass. 1975).

[198] *Id.* at 775–77.

[199] *Id.* at 776 (citation omitted). According to the court, the Consumer Protection Division of the Massachusetts Attorney General spent most of its time responding to consumer complaints instead of pursuing violations of the statute. *Id.*

[200] *Id.*; *see also, e.g.*, Hinchliffe v. Am. Motors Corp., 440 A.2d 810, 815 n.5 (Conn. 1981) (relying on notion that government enforcement is "hampered").

[201] *See, e.g.*, Cole v. Hewlett Packard Co., No. 90,164, 2004 WL 376471, at *6 (Kan. Ct. App. Feb. 27, 2004) (relying on statutory language that act is violated "whether or not the consumer has been misled"); Odom v. Fairbanks Memorial Hosp., 999 P.2d 123, 132 (Alaska 2000) (expressly using state enforcement case law as basis of private action standard); Tallmadge v. Aurora Chrysler-Plymouth, Inc., 605 P.2d 1275, 1277 (Wash. Ct. App. 1979) (same).

[202] *See supra* note 153 and accompanying text.

[203] *E.g.*, State v. Alpine Air Prods., Inc., 500 N.W.2d 788, 790 (Minn. 1993).

claim.[204] Following this reasoning, courts have held that reliance is not required to state a private cause of action.[205]

Other courts have reached the same result by focusing on the FTC Act's standards for public injunctive relief, which do not require reliance.[206] As noted, the FTC need only show that a misrepresentation is "likely to mislead" a consumer and does not have to show causation or even injury.[207] Courts, however, have transposed these public injunctive relief standards to private *damages* actions and have failed to address the causal language present in the private damages provisions.[208]

Still, a few courts have attempted to justify their abandonment of reliance on the ground that reliance is not the same as causation. In *Collora v. R. J. Reynolds Tobacco Co.*,[209] for example, the court admitted that the "as a result of" language in the consumer fraud statute imposes a causation requirement.[210] Recognizing that normally causation would equal reliance, the court took pains to describe the causation requirement as "less strict" than a "proximate cause" requirement but failed to provide any authority for this principle.[211]

In *Varacallo v. Massachusetts Mutual Life Insurance Co.*,[212] the New Jersey Appellate Division similarly stated that there is a "distinction between proof of reliance and proof of causation."[213] At the same time, however, the court explained that plaintiffs could utilize a "presumption or inference of reliance *and* causation, where omissions of material fact are common to the class."[214] The court appeared to recognize that it was equating reliance with causation, stating that "if the plaintiffs in this case establish the core issue of liability, they will be entitled to a presumption of reliance and/or causation."[215]

---

[204] *E.g.*, Gennari v. Weichert Co. Realtors, 691 A.2d 350, 366 (N.J. 1997).

[205] *E.g.*, *id.*

[206] *E.g.*, Slaney v. Westwood Auto, Inc., 322 N.E.2d 768, 779 (Mass. 1975).

[207] *See* discussion *supra* Part III.A.

[208] *See, e.g.*, Oswego Laborers' Local 214 v. Marine Midland Bank, N.A., 647 N.E.2d 741, 745 (N.Y. 1995) (mentioning causal element without any discussion of applicable standard).

[209] No. 002-00732, 2003 WL 23139377 (Mo. Cir. Ct. Dec. 31, 2003).

[210] *Id.* at *2.

[211] *See id.* The court determined that causation is satisfied merely by showing purchase and receipt of a "product that would have been worth more if it in fact had truly been as represented." *Id.* The court found that whether or not the plaintiff purchased the product based on the defendant's alleged misrepresentation was "irrelevant." *Id.* The court, however, drew the wrong causal connection. In a misrepresentation case, a manufacturer's misrepresentation must induce purchase to cause an injury. *E.g.*, WILLIAM L. PROSSER, LAW OF TORTS § 108, at 714 (4th ed. 1971). That inducement to purchase provides the causal connection between the manufacturer's false statement and the consumer's resulting harm. *Id.*

[212] 752 A.2d 807 (N.J. Super. Ct. App. Div. 2000).

[213] *Id.* at 817.

[214] *Id.* (emphasis added).

[215] *Id.* at 818; *see also* Stutman v. Chem. Bank, 731 N.E.2d 608, 612 (N.Y. 2000) (stating that reliance and causation are "not identical").

Thus, some courts have relaxed the substantive requirement of reliance-causation and thereby allowed easy certification of misrepresentation class actions.

IV. THE PUBLIC LAW JUSTIFICATION FOR ABANDONING RELIANCE-CAUSATION

Implicit in the decision of these courts to abandon reliance is an acceptance of the "public-law" version of torts.[216] Historically, the tort system—including misrepresentation cases—was viewed as a means of determining whether a certain actor had wronged a certain victim and, if so, the nature of the appropriate remedy.[217] Up to the nineteenth century, "it is fair to characterize Anglo-American tort law as a peace system, a set of rules which provides a nonviolent way of resolving serious interpersonal disputes."[218] Under this "wrongs-based" conception of tort law, the court's function was not to issue public regulations but to resolve conflicts as presented by the parties.[219]

In recent years, however, commentators have championed a "public law" vision of torts.[220] Under this approach, tort law came to be viewed not as a method for resolving personal disputes, but as a social mechanism for maximizing collective welfare and deterring misconduct.[221] While tort law is nominally a system for providing redress to injured parties, public law advocates view it as a public policy tool that uses private disputes to make public rules.[222]

---

[216] *See* George L. Priest, *What We Know and What We Don't Know About Modern Class Actions: A Review of the Eisenberg–Miller* Study, 9 CIV. JUST. REP. 1, 7 (2005), *available at* http://www.manhattan-institute.org/pdf/cjr_09.pdf (noting that the class action problem has resulted from, among other reasons, "the expansion of liability standards since the mid-1970s based upon largely simple views about how liability judgments can improve societal welfare"); *see generally* David Rosenberg, *The Causal Connection in Mass Exposure Cases: A "Public Law" Vision of the Tort System*, 97 HARV. L. REV. 849 (1984) [hereinafter Rosenberg, *The Causal Connection*] (describing "public law" vision of torts).

[217] John C. P. Goldberg, *Tort Law for Federalists (And the Rest of Us): Private Law in Disguise*, 28 HARV. J.L. & PUB. POL'Y 3, 14 (2004) [hereinafter Goldberg, *Tort Law for Federalists*].

[218] John Hasnas, *What's Wrong with a Little Tort Reform?*, 32 IDAHO L. REV. 557, 562 (1996).

[219] Goldberg, *Tort Law for Federalists*, *supra* note 217, at 14.

[220] *E.g.*, Rosenberg, *The Causal Connection*, *supra* note 216, at 859.

[221] Hasnas, *supra* note 218, at 565. Scholars came to believe that "[t]he traditional account—under which tort law was understood as a set of rules and concepts, grounded in ordinary morality, for resolving disputes over alleged wrongs committed by *A* against *B*—was no longer obviously in tune with modern realities . . . ," such as the industrialized economy and the distance between manufacturer and consumer. John C. P. Goldberg, *Twentieth-Century Tort Theory*, 91 GEO. L. J. 513, 521 (2003) [hereinafter Goldberg, *Twentieth-Century Tort Theory*].

[222] *E.g.*, CHARLES FRIED & DAVID ROSENBERG, MAKING TORT LAW: WHAT SHOULD BE DONE AND WHO SHOULD DO IT 13–32 (2003) (arguing that tort law should be used to achieve "socially optimal management of accident risk"); *accord* Rosenberg, *The Causal Connection*, *supra* note 216, at 907 (arguing that consumer fraud law should "regulat[e]

This public law vision of torts upholds deterrence as the primary goal of the tort system.[223] Under a deterrence theory, the goal of tort law is to deter the risk of harm in an economically efficient manner.[224] From a deterrence perspective, the reliance-causation requirement of a misrepresentation claim is unnecessary: "[I]f one could show that, by some coincidence, television manufacturers were in the best position to deter future automobile accidents, then economic analysis would call for the imposition of liability on television manufacturers, notwithstanding the absence of any causal connection between their conduct and the accidents being deterred."[225] Because deterrence theorists impose liability on the best risk-avoider, there is no need to establish a single causal connection.[226] In a misrepresentation case, the manufacturer stands as the best entity to avoid making false statements about its product. Thus, there need not be a reliance-causation relationship with a particular plaintiff as long as the size of the damages award reflects the extent of the potential injuries that could have resulted from the misrepresentation. Deterrence analysis focuses on how

---

institutional policies or systematic practices that violate external substantive norms"). For a critical discussion of MAKING TORT LAW: WHAT SHOULD BE DONE AND WHO SHOULD DO IT, see MICHAEL S. GREVE, HARM-LESS LAWSUITS? WHAT'S WRONG WITH CONSUMER CLASS ACTIONS 10–13 (AEI Press 2005).

[223] Modern tort scholarship posits three primary functions of tort law: deterrence, corrective justice, and compensation. *See, e.g.*, Jeffery O'Connell & Christopher J. Robinette, *The Role of Compensation in Personal Injury Tort Law: A Response to the Opposite Concerns of Gary Schwartz and Patrick Atiyah*, 32 CONN. L. REV. 137, 138–39 (1999). For a succinct overview of the historical development of these competing theories, see Gary T. Schwartz, *Mixed Theories of Tort Law: Affirming Both Deterrence and Corrective Justice*, 75 TEX. L. REV. 1801, 1802–11 (1997).

[224] *See* George L. Priest, *Modern Tort Law and Its Reform*, 22 VAL. U. L. REV. 1, 20–21 (1987) ("[L]iability should be placed on the party that could have prevented the accident most effectively in order to create incentives to take such actions in the future."). *See generally* WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW (1987). Not all deterrence theorists embrace an economics approach to tort law. *See, e.g.*, Schwartz, *supra* note 223, at 1829. Rather, some scholars have taken a populist approach that views tort law as deterring big business from harming the "little guy" consumer. *Id.* (discussing work of Joan Claybrook). Others have approached deterrence from a socialist perspective that views tort defendants as capitalists likely to impose injury. *Id.* (discussing work of Richard Abel); *see also* Goldberg, *Twentieth-Century Tort Theory*, *supra* note 221, at 514 (dividing deterrence theory into two approaches: (1) compensation-deterrence theory, and (2) economic deterrence theory).

[225] Goldberg, *Twentieth-Century Tort Theory*, *supra* note 221, at 556 (paraphrasing GUIDO CALABRESI, THE COST OF ACCIDENTS 136 (1970)).

[226] Compensation theory likewise minimizes the necessity of cause:

> [S]uppose *D* drives carelessly down a city street without incident. Five blocks away, *P*, though no fault of her own, breaks her ankle stepping off a curb . . . . [I]f *P* can prove that she has in fact suffered an injury for which she needs compensation and further can prove that *D* has engaged in antisocial conduct, why should it matter that *D*'s conduct did not cause her injury?

Id. at 530. Thus, even when the primary emphasis is placed on the compensatory goal of the tort system, it does not matter whether a particular defendant caused this particular injury. Rather, the critical feature of tort law is that the plaintiff receives payment. *Id.*

to prevent the probable harm likely to result from the misrepresentation; in other words, deterrence theory embraces the FTC's "likely to deceive" standard.[227] The fact that more or less harm than predicted—more or less actual consumer reliance and injury—came to pass is irrelevant.

In short, deterrence theory finds little justification for any causation requirement.[228] At its core, the tension arises because the reliance-causation test calls for an ex post analysis, while public law or deterrence theory takes an ex ante view.[229] An ex ante view looks at an individual's preferences "under conditions of uncertainty, at a point in time before the person knows which of possible alternative fates will come to pass."[230] In the misrepresentation case, this approach analyzes the period prior to any consumer's purchase of a product and asks what harm is likely to flow from a manufacturer's misrepresentation.[231] Conversely, reliance-causation—and the ex post view—asks what harm actually came to pass because of the misrepresentation.[232]

In the consumer fraud context, the public law argument is that the "negative value" of misrepresentation claims—the low value of the claim combined with the high costs of litigation—precludes wronged consumers from vindicating their rights.[233] Public law advocates claim that this lack of enforcement leads to under-deterrence and inefficiency.[234] But eliminating reliance as an element of a statutory misrepresentation claim removes a stumbling block in the road to class certification.[235] Removing the reliance requirement allows misrepresentation claims to more easily satisfy class certification standards by eliminating a potential individual issue.[236]

Thus, from the deterrence theorist's perspective, one should weaken the traditional tort requirements and allow private parties to vindicate public rights in a misrepresentation class action. All that matters is that *someone* has come forward to deter misconduct; it does not matter that this plaintiff's injury has no connection to the defendant's misrepresentation. In-

---

[227] *See* discussion *supra* Part III.A.

[228] *See generally* Goldberg, *Twentieth-Century Tort Theory*, *supra* note 221, at 529–32, 556–57.

[229] Schwartz, *supra* note 223, at 1817 n.123 ("The core of the problem is that the actual causation test basically calls for an after-the-fact analysis, while the perspective of the economist is necessarily prospective."); *see also* FRIED & ROSENBERG, *supra* note 222, at 14 (finding that "ex ante and ex post preferences are mutually exclusive"); Christopher Schroeder, *Corrective Justice and Liability for Increasing Risks*, 37 UCLA L. REV. 439, 444 n.18 (1990) (explaining ex ante approach of economic-deterrence theory and ex post approach of causation).

[230] FRIED & ROSENBERG, *supra* note 222, at 14.

[231] GREVE, *supra* note 222, at 19.

[232] *See id.* at 18–19.

[233] *E.g.*, Samuel Issacharoff, *Group Litigation of Consumer Claims: Lessons from the U.S. Experience*, 34 TEX. INT'L L.J. 135, 144 (1999).

[234] *See id.*

[235] *See supra* notes 8–10 and accompanying text.

[236] *See supra* notes 8–10 and accompanying text.

tentionally or not, this framework underlies judicial decisions relaxing or eliminating the reliance requirement for consumer fraud class actions.

V. A Simple Fix To Rein In Consumer Class Actions: Require Reliance

As others have noted, the Class Action Fairness Act of 2005 (CAFA) is not going to solve the documented abuses of the consumer class action.[237] The real solution lies in a more reasoned approach to these lawsuits: judicial enforcement of the reliance-causation requirement in state consumer fraud statutes. The interpretative foundations of the misrepresentation class action—historical non-enforcement by government agencies and application of the "public law" theory—no longer hold true.[238] Absent that underpinning, the text of the state statutes provides a clear distinction between the standards for public enforcement and the standards that should govern a private damages claim.[239]

A. Today, Government Enforcement Is Strong

The abandonment of a reliance-causation requirement has been premised, in part, on lax public enforcement.[240] Unlike the days of "Nader's Raiders,"[241] today the FTC and its state counterparts rigorously enforce consumer fraud statutes.[242] These modern enforcement policies support a reliance requirement in private damages actions: allowing individual consumers to sue where there is causation and harm, leaving public agencies to act before the defendant's misrepresentation has caused harm or where the negative value of the claim effectively precludes private enforcement.

Unlike the FTC of the 1960s, today's FTC and state attorneys general take an active role in protecting the public from manufacturers' misrepresentations.[243] Recently, for example, the FTC settled a complaint against Tropicana Products, Inc.[244] The FTC alleged that Tropicana's advertisement

---

[237] For example, Professor Priest has characterized CAFA as a "modest reform," which "is not likely to solve the problems created by the modern class action." Priest, *supra* note 224, at 7. Professor Priest contends that the class action problem derives, at least in part, from "the expansion of liability standards since the mid-1970s based upon largely simple views about how liability judgments can improve societal welfare." *Id.* Other scholars contend that CAFA overestimates the superiority of federal adjudication of class actions and ignores the aggregation problems inherent in the class action device itself. *E.g.*, Mark Moller, *The Rule of Law Problem: Unconstitutional Class Actions and Options for Reform*, 28 Harv. J.L. & Pub. Pol'y 855, 891–93 (2005).

[238] *See infra* Parts V.A–C.

[239] *See infra* Part V.D.

[240] *See supra* notes 196–200 and accompanying text.

[241] *See supra* notes 76–93 and accompanying text.

[242] *See infra* notes 244–261 and accompanying text.

[243] *See infra* notes 244–261 and accompanying text.

[244] Agreement Containing Consent Order, *In re* Tropicana Prods. Inc., No. 042-3154 (FTC June 2, 2005), *available at* http://www.ftc.gov/os/caselist/0423154/050602agree0423154.

suggesting that its orange juice "improve[d] heart health" was misleading to consumers and lacked scientific support.[245] The FTC's settlement satisfied deterrence goals: Tropicana has agreed to stop using these statements and to file compliance reports with the FTC.[246]

Moreover, government relief is not limited to deterrence but also embraces the compensatory goals of the tort system.[247] In June 2005, for example, the FTC obtained a judgment of $4.9 million in consumer redress—a sum equal to the total amount of sales—from a company that deceptively marketed a weight-loss pill.[248] Indeed, from April 2004 to March 2005, the FTC has obtained the return of more than $480 million in consumer redress.[249]

Enforcement at the state level similarly has achieved both deterrence and compensatory goals. For example, the National Association of Attorneys General ("NAAG")[250] recently settled a false advertising claim against Blockbuster, Inc.[251] NAAG alleged that Blockbuster's "no late fee" advertisements were misleading because consumers who kept a rental more than seven days past its due date were charged the sales price of the item.[252] Un-

---

pdf. [hereinafter Consent Order]; *see also* Press Release, FTC, FTC Puts the Squeeze on Tropicana's Orange Juice Claims (June 2, 2005), *available at* http://www.ftc.gov/opa/2005/06/tropicana.htm.

[245] Complaint ¶ 5C & Ex. C, *In re* Tropicana Prods. Inc., No. 042-3154 (FTC June 2, 2005), *available at* http://www.ftc.gov/os/caselist/0423154/050602comp0423154.pdf. Other challenged statements included "A new clinical study shows enjoying two glasses of Tropicana Pure Premium every day can lower your blood pressure an average of ten points." *Id.* ¶ 5A & Ex. A.

[246] Consent Order, *supra* note 244, ¶¶ II–V; *see also* Tropicana Prods., Inc.; Analysis of Agreement Containing Consent Order to Aid Public Comment, 70 Fed. Reg. 35,434 (June 20, 2005) (summarizing provisions of consent order).

[247] *See supra* notes 149–151.

[248] Press Release, FTC, Defendants Who Deceptively Marketed the "Himalayan Diet Breakthrough" Settle FTC Charges: Agree to Pay $400,000 in Consumer Redress (June 20, 2005), *available at* http://www.ftc.gov/opa/2005/06/avsmarketing.htm. Upon determining that the defendant was unable to pay the full redress, the judgment was suspended upon payment of $400,000 to the FTC. *Id.* If, however, it is determined that the defendants misrepresented their financial condition, the full $4.9 million becomes due immediately. *Id.*

[249] Deborah Platt Majoras, Chairman, FTC, Keynote Address at the OECD Workshop on Dispute Resolution and Consumer Redress, at 2 (Apr. 19, 2004), *available at* http://www.ftc.gov/speeches/majoras/050419oecdworkshop.pdf; *see also* DEE PRIDGEN, CONSUMER PROTECTION AND THE LAW § 12:20, at 1045–46 (2004) (discussing representative sample of successful consumer redress claims in the 2002–2004 period, including nearly $13.5 million for several deceptive marketing cases).

[250] Founded in 1907, the National Association of Attorneys General facilitates cooperation among the state attorneys general. *See generally* NAAG, About NAAG, http://www.naag.org/naag/about_naag.php (last visited Nov. 4, 2005). NAAG's members are the attorneys general of all fifty states plus the District of Columbia, the Commonwealth of Puerto Rico, the Northern Mariana Islands and the territories of American Samoa, Guam, and the Virgin Islands. *Id.*

[251] Press Release, NAAG, Forty-Seven States and District of Columbia Settle With Blockbuster—Settlement Requires Refunds, Credits, Additional Disclosures (Mar. 30, 2005), *available at* http:www.naag.org/news/pr-20050330-blockbuster.php.

[252] *Id.*

der the settlement, customers will receive a one-time full refund or credit.[253] And, in contrast to private litigation,[254] the NAAG settlement achieved its deterrence goal: Blockbuster agreed to modify its advertisements and disclose the "no late fee" limitations.[255]

Similarly, in 2004, attorneys general from thirty-two states settled false advertising charges with three of the nation's largest wireless telephone carriers.[256] Under the settlement, Verizon, Sprint, and Cingular agreed to modify their advertising and provide comprehensive information about the costs and limits of their wireless services.[257] Here, too, state enforcement achieved its deterrence objective.

Individual states likewise are taking effective action. In 2003, New Jersey brought suit against Nutraquest, Inc., a manufacturer of a weight-loss product containing ephedra.[258] New Jersey alleged that the manufacturer falsely claimed that its product produced weight loss without dieting or exercise.[259] In July 2005, the company settled the suit for nearly $1 million.[260] Under the settlement, the company is prohibited from stating that its products can cause weight loss without diet or exercise.[261]

To be sure, public enforcement has its limitations.[262] Public agencies often lack sufficient financial resources and political will to monitor and detect all misrepresentations. But these perceived weaknesses may also be seen as strengths. Political[263] and budgetary restraints[264] limit the gov-

---

[253] *Id.*

[254] *See infra* notes 295–301 and accompanying text.

[255] Press Release, NAAG, *supra* note 251.

[256] Press Release, NAAG, Settlement: Thirty-two Attorneys General Settle with Verizon, Cingular, and Sprint PCS (July 22, 2004), *available at* http://www.naag.org/issues/20040722-settlement-wireless.php.

[257] *Id.*

[258] Charles Toutant, Suits & Deals, *$940,000 for Consumer Fraud*, 181 N.J. L.J. 173 (2005).

[259] *Id.*

[260] *Id.*

[261] NutraIngredients-USA, *NutraQuest Settles Over Exaggerated Adverts* (July 13, 2005), http://www.nutraingredients-usa.com/news/ng.asp?n=61258&m=1niu713&c=cgnwsiywf mwgwmf (last visited Nov. 4, 2005).

[262] Issacharoff, *supra* note 233, at 137–41. Professor Issacharoff notes five limitations on government enforcement: (1) lack of resources; (2) jurisdictional limits; (3) difficulty of acquiring information regarding consumer fraud; (4) limited consumer accessibility to government centers; and (5) dependence on political will. *Id.*

[263] Critics still charge that the FTC's consumer protection ability remains subject to politics. *See generally* Mark E. Budnitz, *The FTC's Consumer Protection Program During the Miller Years: Lessons for Administrative Agency Structure and Operation*, 46 CATH. U. L. REV. 371 (1997) (examining effect of Congress, industry, and consumer groups on FTC's enforcement activities during the Reagan administration); William E. Kovacic, *Congress and the Federal Trade Commission*, 57 ANTITRUST L.J. 869 (1989) (examining the FTC's relationship with Congress and its exposure to political attacks).

[264] In fiscal year 2006, the FTC received a $211 million budget for program outlays. OFFICE OF MGMT. & BUDGET, EXEC. OFFICE OF THE PRESIDENT, BUDGET OF THE UNITED STATES GOVERNMENT, FISCAL YEAR 2006, 1179 (2005), *available at* http://www.gpoaccess.gov/usbudget/fy06/pdf/appendix/oia.pdf. Of that amount, nearly half—$105 million—was earmarked for consumer protection activities. *Id.*; *see also* Consolidated Appropriations

ernment's ability to pursue potentially frivolous litigation, and thus harmonize public enforcement with public will. These limits encourage government agencies to focus on the most pressing wrongs: misrepresentations that have harmed the greatest number of consumers and misrepresentations that are truly deceptive.

Political accountability further provides a democratic check on misrepresentation suits.[265] Because most attorneys general are elected by the public,[266] a state's enforcement of consumer protection laws becomes accountable to the public.[267] If the electorate disagrees with an attorney general's enforcement decisions, the ballot box can register this disapproval, and the attorney general may be voted out of office. Indeed, consumer advocates often lose sight of the fact that public agencies act on behalf of *all* constituents, including manufacturers in their jurisdiction. Public accountability allows state attorneys general to channel resources to the benefit of the entire public, not just a small class of litigants.[268]

California, for example, provides an example of the public stepping in to limit the reach of the state's consumer fraud statute. Prior to November 2004, California's Unfair Competition Law ("UCL") was sweeping in its breadth.[269] The California UCL could be enforced by anyone. It was unnecessary to allege reliance by or even injury to anyone, never mind the plaintiff.[270] Rather, "'any person acting for the interests of . . . the general public,' [could] bring an action."[271] While a plaintiff could not re-

---

Act of 2005, Pub. L. No. 108-447, 118 Stat. 2809, 2908 (2004). By contrast, in 2003, the Senate approved double that amount—$210 million—for the State of North Dakota to use on highway and transportation projects. Press Release, Kent Conrad, U.S. Senator, North Dakota, Senate Approves More than $210 Million for North Dakota Highways (Oct. 24, 2003), *available at* http://conrad.senate.gov/~conrad/releases/03/10/2003A24B22.html.

[265] *See* Martin H. Redish, *Class Actions and the Democratic Difficulty: Rethinking the Intersection of Private Litigation and Public Goals*, 2003 U. CHI. LEGAL F. 71, 73 (arguing that "modern class action has undermined the foundational precepts of American democracy").

[266] Forty-three states elect their attorneys general by popular vote. NAAG, About NAAG, http://www.naag.org/naag/about_naag.php (last visited Nov. 4, 2005).

[267] *See* Redish, *supra* note 265, at 109–11 (discussing representation and accountability principles).

[268] In the consumer fraud class action, the argument that private parties would have greater incentive to bring suit than would a state attorney general is misplaced. The filing of a consumer fraud action is largely an attorney-driven creation. *See supra* note 4. Moreover, consumers typically receive no meaningful benefit from these suits. *See supra* notes 50–54. In any event, to the extent that a consumer does have individual incentive to sue, such as where damages are particularly large, the consumer remains free to bring her own suit, provided she can establish reliance on the manufacturer's statement.

[269] Unfair Competition Law, CAL. BUS. & PROF. CODE §§ 17200–17209 (West 1997 & Supp. 2005).

[270] *E.g.*, Blakemore v. Superior Court, 27 Cal. Rptr. 3d 877, 888 (Ct. App. 2005) ("Unlike common law fraud, a [UCL] violation can be shown even without allegations of actual deception, reasonable reliance and damage.").

[271] Kasky v. Nike, Inc., 45 P.3d 243, 249 (Cal. 2002) (quoting UCL § 17204).

cover damages under the UCL, monetary relief—restitution and disgorgement—was available under equitable principles.[272]

Many Californians, however, ultimately thought the statute went too far. In November 2004, voters enacted Proposition 64, which transformed California's UCL back into a private law statute.[273] Under Proposition 64, a UCL plaintiff now must show that he "has suffered injury in fact and has lost money or property as a result of such unfair competition."[274] Proposition 64 effectively forecloses suits by private plaintiffs who have not suffered a loss in reliance on a manufacturer's misrepresentation.[275]

Public law theorists would argue that limiting private consumer actions to individuals who can show reliance could under-deter misrepresentations that have not yet harmed individual consumers.[276] But these harms are *not* left undeterred. Public agencies remain able to address these quintessentially public harms. Moreover, government enforcement has the comparative advantage in articulating and applying a consumer protection policy that addresses these public harms. Because the government has substantial control over the selection of cases, it can direct a coherent body of law via both regulation and litigation.[277]

---

[272] CAL. BUS. & PROF. CODE § 17203 (West 1997); *see also* Bank of the West v. Superior Court, 833 P.2d 545, 557 (Cal. 1992) (holding that compensatory damages are not available under California UCL).

[273] *See* United Investors Life Ins. Co. v. Waddell & Reed, Inc., 23 Cal. Rptr. 3d 387, 388–89 (Ct. App. 2005). For the complete text of Proposition 64 and relevant portions of the Voter's Information Guide, including the statements by proponents and opponents, see Californians for Disability Rights v. Mervyn's, LLC, 24 Cal. Rptr. 3d 301 app. (Ct. App. 2005), *review granted*, 110 P.3d 1216 (Cal. 2005).

[274] CAL. BUS. & PROF. CODE § 17204 (West Supp. 2005); *see also United Investors Life Ins. Co.*, 23 Cal. Rptr. at 389 (noting text of Proposition 64).

[275] *See, e.g.*, *In re* Tobacco Cases II, No. JCCP 4042, 2005 WL 579720, at *6 (Cal. App. Dept. Super. Ct. Mar. 7, 2005) (decertifying class action because individual issues predominated the case, including whether each class member purchased cigarettes "as a result of" defendants' alleged false statements). The Texas legislature similarly amended its UDAP statute in 1995 to require reliance for misrepresentation claims. *See* Alford Chevrolet-Geo v. Murphy, No. 06-02-00059-CV, 2002 WL 31398487, at *1 n.2 (Tex. App. Oct. 25, 2002).

[276] *See supra* Part IV.

[277] For example, a state attorney general can take into account agency regulations and policy decisions when deciding whether to initiate litigation. Plaintiffs' lawyers, however, do not take these factors into account. In *Avery v. State Farm Mutual Automobile Ins. Co.*, for example, plaintiffs' lawyers filed a nationwide class action against State Farm, alleging that the use of non-original equipment manufacturer ("non-OEM") parts during repairs violated the Illinois Consumer Fraud and Deceptive Practices Act, even though at least forty states had enacted regulations or statutes expressly permitting the use of non-OEM parts as a means of reducing insurance premiums. 746 N.E.2d 1242, 1247, 1254 (Ill. App. Ct. 2001), *aff'd in part and rev'd in part*, 835 N.E.2d 801 (Ill. 2005). The plaintiffs' lawyers in *Avery* essentially usurped the roles of the state insurance commissioner and the state attorney general in deciding to sue State Farm for actions it had taken in compliance with state law. The $1.18 billion judgment against State Farm recently was reversed by the Illinois Supreme Court. 835 N.E.2d at 818.

B. Public Law Theory Is Misapplied to Misrepresentation Class Actions

Whatever the merits of the public law vision of tort law in other areas, the theory has taken courts down the wrong path in the consumer law context. In the misrepresentation class action, application of the public law theory has created both over-deterrence and under-deterrence. The absence of reliance-causation means that the damages award itself often is based on conduct that caused no harm, thereby over-deterring the defendant's conduct. Conversely, the ease of class certification can lead to settlements that are actually profitable for the defendant and thus provide insufficient deterrent against future misrepresentations. Moreover, the public law approach ignores the institutional reality of the tort system—someone gets a lot of money—and thus creates the potential for windfalls and incentives for frivolous litigation.[278]

1. Over-deterrence

The absence of reliance-causation eases the road to certification.[279] By eliminating reliance as an element, a misrepresentation class more easily satisfies the predominance inquiry for class certification.[280] This easier burden, however, creates an incentive to settle the case—not because the manufacturer has harmed the plaintiff, but because the case presents the risk of a bankrupting judgment.[281] Considering the recent Class Action Fairness Act of 2005, the Senate Judiciary Committee concluded that "[s]uch leverage can essentially force corporate defendants to pay ransom to class attorneys by settling—rather than litigating—frivolous lawsuits."[282] Thus, even though the defendant has done no wrong and there is nothing to deter, punishment is imposed through settlement.

Moreover, ignoring questions of reliance-causation creates an abstract inquiry into whether these particular class members are *entitled* to this money. To address the entitlement question, public law advocates suggest that litigation include subsequent compensation proceedings that resolve reliance questions and determine to what extent class members are entitled to share in the monetary award.[283] Under that system, however, the judge is forced to come up with a hypothetical total award that gives some class members the benefit of a bargain that they already have received. For example, the consumer class in the "light" cigarettes case, *Aspinall*,[284]

---

[278] *See infra* notes 283–287 and accompanying text.

[279] *See supra* notes 8–10 and accompanying text.

[280] *See supra* notes 8–10 and accompanying text.

[281] *See supra* notes 48–49 and accompanying text.

[282] S. REP. NO. 109-14, at 20 (2005).

[283] *See* David Rosenberg, *Decoupling Deterrence and Compensation Functions in Mass Tort Class Actions for Future Loss*, 88 VA. L. REV. 1871, 1885–88 (2002).

[284] 813 N.E.2d 476 (Mass. 2004). For a discussion of this decision, see *supra* note 39.

includes everyone from hypothetical health-conscious cigarette smokers who actually believed that they were purchasing a "safer" product to image-conscious smokers who are under no such illusions that any cigarette is safe as well as those who chose Marlboro Lights because they prefer the taste.[285] Including these knowledgeable class members in calculating the total damages award results in a sum greater than the harm caused by the misrepresentation.[286]

Thus, requiring manufacturers to pay compensation not simply to their actual victims but to all purchasers results in over-deterrence and subjects defendants to excessive liability. Indeed, even proponents of a public law vision of torts concede that "[p]unishing every error in judgment regardless of whether it has caused harm might result in excessive liability and could lead not only to overbearing and discriminatory enforcement, but also to a fearful and overcautious society."[287]

2. *Under-deterrence*

At the same time that it creates over-deterrence, the lack of a reliance-causation requirement creates under-deterrence. First, eliminating reliance-causation as an element of a misrepresentation case minimizes the showing needed to obtain certification[288] and thereby increases the likelihood of an early settlement.[289] This settlement process creates the potential for under-deterrence—true misrepresentations staying on the market with little to no penalty for the manufacturer and no real redress for the consumer. As the Institute for Civil Justice noted, early settlement can avoid full discovery, and the full extent of the defendant's wrongdoing is never exposed.[290]

Aside from avoiding discovery of wrongdoing, the manufacturer may be able to negotiate a settlement amount less than its total profit on the product. Thus, a misrepresentation class action settlement can undervalue the deterrent effect of the suit in order to achieve a quick resolution, tipping the scale on the side of private gain, not public good.[291] For example, a large number of consumer class actions are settled using a coupon method in which a defendant avoids liability by paying class members in

---

[285] *See supra* note 39 (discussing *Aspinall* decision).

[286] Thus, in *Aspinall*, the court noted that the potential damages would be measured based on the mere fact of purchase. *See* 813 N.E.2d at 490 n.23 (noting damages for class would be measured based on "difference between value paid and value received").

[287] Rosenberg, *The Causal Connection*, *supra* note 216, at 882.

[288] *See supra* notes 8–10 and accompanying text.

[289] *See supra* notes 48–49 and accompanying text.

[290] HENSLER ET AL., *supra* note 4, at 79–80; *see also id.* at 424 (noting that "class action attorneys were sometimes simply interested in finding a settlement price that defendants would agree to—rather than in finding out what class members had lost [and] what defendants had gained . . .").

[291] Furthermore, the costs of litigation are still passed on to all consumers in the form of higher prices, even though the consumer may not have received any commensurate benefit.

promotional coupons.[292] Even assuming that class members will redeem the coupons, a coupon settlement does not deter future misconduct by a defendant: "[A] class member's redemption of a settlement coupon merely decreases, but does not eliminate, a defendant's profit margin on a given sale. More importantly, from the defendant's perspective, the settlement coupons may encourage additional sales and thereby increase the defendant's net profits."[293] Thus, far from being punishment for a company's misrepresentation, a coupon settlement may benefit the defendant.[294]

Consider the private Blockbuster late fee litigation, a case noted by the Senate Judiciary Committee in its 2003 report on class actions.[295] In that case, Blockbuster settled a state court class action alleging that the company charged excessive late fees.[296] Under the terms of the settlement, class members received a coupon for one dollar off their next rental.[297] However, "[e]xperts . . . predicted that at most, only 20 percent of the class members will redeem the coupons."[298] The low redemption rate—characteristic of most coupon settlements[299]—effectively allowed Blockbuster to escape damages.[300] Moreover, the settlement had no deterrent effect and expressly allowed Blockbuster to continue its fee policy.[301]

---

[292] *See* Christopher R. Leslie, *The Need to Study Coupon Settlements in Class Action Litigation*, 18 GEO. J. LEGAL ETHICS 1395 (2005) [hereinafter Leslie, *The Need to Study Coupon Settlements*] (explaining that defendants avoid liability due to a low redemption rate for the settlement coupons); *see also* Leslie, *A Market-Based Approach to Coupon Settlements*, *supra* note 50, at 996 (noting that "in most cases, coupons are not punishment; they are promotional"). *See generally id.* (discussing coupon settlements and explaining how such settlements are worthless for most class members).

[293] *Id.* at 1014. Professor Leslie further explains that defendants can manipulate the value of a settlement coupon by increasing the price of the product or by lowering its quality. *Id.* at 1030–33.

[294] Professor Leslie also points out that, apart from the financial gain derived from a coupon settlement, coupon settlements reward defendants by "provid[ing] a competitive advantage." *Id.* at 1039. Because coupons generate sales, a settlement coupon can induce a class member to avoid a competing product and instead purchase the defendant's goods. *Id.*

[295] S. REP. NO. 108-123, at 16 (2003).

[296] *Id.*

[297] *Id.* Moreover, customers had to fill out claims forms to receive the one-dollar coupon and return the claim form before the deadline. *See* Cynthia Corzo, *Blockbuster Settles Suits Over Late Fees*, MIAMI HERALD, June 5, 2001, at A1. The plaintiffs' lawyers, however, received $9.25 million. S. REP. NO. 108-123, at 16.

[298] S. REP. NO. 108-123, at 16.

[299] In general, the redemption rate of class action coupons ranges from one to three percent. James Tharin & Brian Blockovich, *Coupons and the Class Action Fairness Act*, 18 GEO. J. LEGAL ETHICS 1443, 1445 (2005). *But see* Leslie, *A Market-Based Approach to Coupon Settlements*, *supra* note 50, at 1035 (noting that consumer class action redemption rates vary from as low as 3% to 13.1%).

[300] *See* Leslie, *A Market-Based Approach to Coupon Settlements*, *supra* note 50, at 1035–37. As Professor Leslie has explained, "[a] low redemption rate . . . is proof positive that the settlement failed. Low redemption rates mean no or low compensation. Similarly, if the redemption rate is low, then there was insufficient disgorgement." Leslie, *The Need to Study Coupon Settlements*, *supra* note 292, at 1402.

[301] S. REP. NO. 108-123, at 16. Indeed, in press reports, Blockbuster stated that it would not change its policy: "We're pleased we can end this litigation and that our rental policy will continue unchanged." Corzo, *supra* note 297 (quoting Ed Stead, executive vice

Although the Class Action Fairness Act of 2005 seeks to reform these coupon settlements, the effects will be limited. Under the new law, if attorneys' fees are based on a percentage of the coupon, fees must be calculated based on *redeemed* coupons, not merely issued coupons.[302] This provision falls short of solving the prroblem because, first, it applies only to actions in federal court[303] and only to settlement agreements that tie attorneys' fees to the coupon value.[304] Second, while tying attorneys' fees to coupon redemption may increase claims rates, nothing in CAFA requires a settling defendant to stop its deceptive practices. Thus, even under CAFA, Blockbuster could continue its deceptive fee practice unchanged.

Even where a consumer class action settles with a monetary award to the plaintiffs, these settlements often impose administrative burdens on class members that limit any actual dispersal of the defendant's money. Few individual plaintiffs will submit the necessary claim forms and ultimately receive compensation.[305] Thus, again, a low claims rate reduces the monetary penalty inflicted upon the defendant and thus lessens any deterrent function of the settlement.[306]

Moreover, allowing misrepresentation claims to proceed without any showing of reliance creates inefficient incentives by allowing the consumer to feign ignorance of information they actually have.[307] As a group, consumers would have less incentive to obtain information about the products they purchase,[308] and manufacturers as a class would have less incentives to inform customers about their products. For instance, does anyone really believe that a daily diet of fast food is healthy? Or do jam purchasers really believe they are buying 100% fruit? Reliance insures that the

---

president and general counsel for Blockbuster).

[302] 28 U.S.C.A. § 1712(a) (West 2005). CAFA also requires a settling defendant to notify "the appropriate State official" of a proposed class action settlement. *Id.* § 1715(b).

[303] *Id.* § 1711(2).

[304] *Id.* § 1712(b). *See also* Leslie, *The Need to Study Coupon Settlements*, *supra* note 292, at 1417 (discussing limits on CAFA's effect on coupon settlements).

[305] *See* Hillebrand & Torrence, *supra* note 54, at 747. A cumbersome disbursement process results in a low claims rate:

> Settlements and judgments in class action cases have often required class members to submit claims in order to share in the proceeds of the recovery. Recent cases suggest that claims procedures are ill-suited to consumer class actions in which the class size is very large and the amount of damages per class member is relatively small. These cases are characterized by very low claims rates.

Id.

[306] Similarly, these suits fail to serve any compensatory goal. Low claims rates mean that most consumers do not receive any compensation from these suits.

[307] *See* Goldberg, *Twentieth-Century Tort Theory*, *supra* note 221, at 550–51 (explaining how economic deterrence theory requires both parties to take steps to produce the "smallest sum of precaution costs and accident costs").

[308] As Judge Posner has explained, "the knowledge and the intelligence of the consumer" deter manufacturer misrepresentations. POSNER, REGULATION OF ADVERTISING, *supra* note 74, at 5. Eliminating reliance-causation as an element weakens the deterrent potential of the consumer.

consumer bear her share of responsibility in deciding to purchase a product based on a manufacturer's misrepresentation. By doing so, a reliance requirement screens out the insignificant misrepresentation that has no effect on the consumer's purchasing decision. The element of reliance-causation thus apportions deterrence between the manufacturer and the consumer.

Of course, it is impossible to eliminate all misrepresentation from the market. Rather, the optimal rule "will seek to preserve informational value while screening out the misrepresentations that induce social losses. That, precisely is the point of . . . detrimental reliance."[309]

C. Private Law Theory Provides a Better Controlling Framework for Misrepresentation Class Actions

Where dual public and private enforcement regimes exist side by side, the private law theory of torts should control the interpretation of private damages actions. Historically, tort law was "conceived as a law of personal redress rather than as a law of public regulation or punishment."[310] Thus, tort law allowed an injured person to bring suit against the wrongful party and recover money damages.[311]

The tort system thus serves as a means of achieving justice between the parties.[312] Under this view, a wrong has been done to a victim. Assuming the victim is innocent, the question becomes whether there is a person whose connection to the wrong borne by the victim warrants that the burden be shifted to that person.[313] In other words, is there a person whose connection to the wrong generates a responsibility on his part to fix it? Tort law then "corrects" the injustice by transferring the loss to the wrongdoer via a damages payment.[314]

In a misrepresentation class action, reliance-causation ties the plaintiff's loss[315] to an injustice by the defendant.[316] Reliance-causation thus

---

[309] GREVE, *supra* note 222, at 32.

[310] Goldberg, *Twentieth-Century Tort Theory*, *supra* note 221, at 517.

[311] *Id.* For an overview of the origins of the tort system as an alternative to "blood feuds" during the Middle Ages, see Hasnas, *supra* note 218, at 558–61.

[312] *See generally* Schwartz, *supra* note 223 (describing corrective justice theory of torts).

[313] *See* Goldberg, *Twentieth-Century Tort Theory*, *supra* note 221, at 570.

[314] *Id.*

[315] Corrective justice scholars dispute whether tort law responds to "wrongful conduct" or "wrongful loss." *Id.* at 571. Causation, however, is essential under both views. Under the "wrongful loss" view, the victim suffers a loss that she does not deserve to bear. *Id.* To transfer the loss, the victim must show that (1) another person acted wrongfully, and (2) that person caused her loss. *Id.* Critics have argued that this theory draws too sharp a distinction between wrongful conduct and causation. Thus, under the "wrongful conduct" theory, tort law corrects "wrongs" themselves, and looks for a causal connection between the wrong and the injury. *Id.* In either case, causation is essential.

[316] *See* Schwartz, *supra* note 223, at 1815–19 (finding that tort suits impose liability on "party whose tortuous conduct has 'caused' the plaintiff's injury"). *But see* Christopher J.

links "doer and sufferer," or institutionally speaking, the plaintiff and the defendant.[317] In the misrepresentation case, reliance answers the question "[w]hy can *this* plaintiff recover from *this* defendant?"[318] Reliance identifies *this specific* plaintiff as someone entitled to recover for her injury from all the persons who heard or saw the defendant's misrepresentation.[319] Because tort law functions within a litigation system, this understanding forms the basis of the entire structure of the tort system:

> [C]ausation particularizes by singling out this plaintiff from the class of persons whom the defendant has endangered. Through injury the general risk which the wrongdoing has unreasonably created lodges in a particular person. Similarly, wrongdoing serves to single out from among the numerous causal antecedents of the plaintiff's injury the particular cause that is juridically significant. Causation particularizes the plaintiff against the background of the defendant's wrongful risk creation, and wrongdoing particularizes the defendant against the background of the totality of the injury's causes.[320]

By identifying a victim, reliance defines whom the defendant must compensate: "[T]he fact that *A* causes *B* harm is normatively significant because it demonstrates that *B*, not someone else, was harmed by *A*. So if *A* must pay someone, it must be *B*, not *C*, *D*, or *E*, none of whom were harmed by *A*."[321] Indeed, even ardent public law advocates admit that the corrective justice approach "works best for intentional wrongs."[322] Though consumer fraud statutes have eliminated the intent requirement of common

---

Robinette, *Can There be a Unified Theory of Torts? A Pluralist Suggestion from History and Doctrine*, 43 BRANDEIS L.J. 369, 405–07 (2005) (arguing that, in some instances, the causation requirement can undermine corrective justice rationale because the causation element creates litigation hurdles that prevent injured parties from seeking justice); Schroeder, *supra* note 229, at 439 (arguing that liability for risk creation, regardless of whether actual harm is caused, is consistent with corrective justice theory); Kenneth W. Simons, *Corrective Justice and Liability for Risk-Creation: A Comment*, 38 UCLA L. Rev. 113, 137–38 (1990) (applying Professor Schroeder's corrective justice theory to intentional torts).

[317] *See generally* Ernest J. Weinrib, *Causation and Wrongdoing*, 63 CHI.-KENT L. REV. 407 (1987) (arguing that causation both identifies the victim and provides moral reason for requiring wrongdoer to compensate the victim).

[318] *See* Weinrib, *supra* note 317, at 414. For a critique of Professor Weinrib's article, see Jules L. Coleman, *Property, Wrongfulness and the Duty to Compensate*, 63 CHI.-KENT L. REV. 451 (1987).

[319] *See* Weinrib, *supra* note 317, at 416 ("Causation, then, has the function of particularizing the plaintiff in relation to the defendant."). Causation likewise identifies the defendant from all other actors who could cover the plaintiff's loss. *Id.* at 417–18. Actual causation ("but for"), however, can stretch back indefinitely. Thus, on the defendant's side, the causal inquiry stops with a wrongful act. *Id.*

[320] *Id.* at 429–30.

[321] Coleman, *supra* note 318, at 452.

[322] FRIED & ROSENBERG, *supra* note 222, at 31.

law misrepresentation claims,[323] these suits are fraud cases at their core and should be treated as such.

To be sure, tort law serves certain public functions. Through the imposition of damages via judgments and settlements, private misrepresentation cases deter manufacturer misrepresentations. But government agencies have been specifically charged with that very objective. Duplicating this public function through the misrepresentation class action is a waste of resources. Simply put, a "public law" vision of torts has no place where true public law—government enforcement—exists.

D. The Language of Consumer Fraud Statutes Establishes a Clear Distinction Between Public and Private Enforcement

Once the flawed theoretical underpinnings of the "no reliance-causation" approach are exposed, a clear distinction between public and private enforcement emerges in the text of these statutes. Nothing in these statutes suggests that the state legislatures intended to eliminate causation as an element of a private damages claim. Moreover, these statutes were enacted against a common law background that equated cause with reliance in misrepresentation cases.

Every jurisdiction, except the District of Columbia, imposes a causation requirement for a private cause of action under its consumer fraud statute.[324] Where a state has decided to abandon causation requirements, it has done so explicitly. For example, the consumer protection act for the District of Columbia contained a causation requirement prior to 2000: "Any consumer who *suffers damages as a result of* the use or employment of any practice by any person of a trade practice" could bring suit.[325] In 2000, however, the District amended the statute to eliminate the causation requirement.[326] The statute now reads "[a] person, whether acting for the interests of itself, its members, or the general public, may bring an action under this chapter."[327]

Thus, by including a causation requirement,[328] these consumer fraud statutes embrace a reliance requirement:

> The causal connection between the wrongful conduct and the resulting damage, essential throughout the law of torts, takes in cases

---

[323] *See infra* note 334 and accompanying text.

[324] *See supra* note 162.

[325] D.C. CODE ANN. § 28-3905 (LexisNexis 2001) (emphasis added).

[326] D.C. Law 13-172, § 1402(d), 47 D.C.R. 6308 (2000); *see also* Wells v. Allstate Ins. Co., 210 F.R.D. 1, 8 (D.D.C. 2002) (noting that the amendment eliminated the injury and causation requirements).

[327] D.C. CODE ANN. § 28-3905(k)(1) (LexisNexis 2001).

[328] Indeed, some states have acted to make the reliance-causation requirement explicit in the text of the statute. *See supra* notes 270–275 and accompanying text (discussing recent changes in California and Texas).

> of misrepresentation the form of inducement of the plaintiff to act, or to refrain from acting, to his detriment . . . . In order to be influenced by the representation, the plaintiff must of course have relied upon it, and believed it to be true. If it appears that he knew the facts, or believed the statement to be false, or that he was in fact so skeptical as to its truth that he reposed no confidence in it, it cannot be regarded as a substantial cause of his conduct.[329]

Thus, as a practical matter, damages cannot be "caused" by a defendant's misrepresentation without reliance on the statement.[330]

Courts that have rejected this principle have misunderstood the relationship between reliance and causation in a misrepresentation case. In *Smoot v. Physician's Life Insurance Co.*,[331] for example, the New Mexico Court of Appeals contended that "causation and reliance are distinct concepts," explaining, "causation requires a nexus between a defendant's conduct and a plaintiff's loss; reliance concerns the nexus between a defendant's conduct and a plaintiff's purchase or sale."[332] This argument, however, ignores the fact that in a misrepresentation case, the plaintiff's "loss" *is* the "purchase or sale." Consumer classes seek to recover the money spent on the product.[333]

To be sure, consumer fraud statutes have been viewed as lessening the requirements of common law claims. First, even without a reliance requirement, these statutes *do* lower the common law standards. Notably, consumer fraud statutes do not require proof of intent to deceive or scienter,[334] and many common law defenses are not allowed under the consumer fraud statutes.[335] Moreover, the "relaxed" non-common law standard originally was placed in the hands of government enforcement.[336] And government enforcement standards do not require reliance-causation or even injury for that matter—an understandably lower burden given the government's focus on the public interest and desire to deter fraud before any harm occurs.[337] But application of this lower standard to a private dam-

---

[329] WILLIAM L. PROSSER, LAW OF TORTS § 108, at 714 (4th ed. 1971).

[330] *Id.*; *see also supra* notes 183–184, 190–192 and accompanying text.

[331] 87 P.3d 545 (N.M. Ct. App. 2003).

[332] *Id.* at 550 (citation omitted).

[333] Typically, these misrepresentation class actions disclaim any personal injury losses and seek only to recover "benefit of the bargain" damages, which award the difference between the actual value of the product at the time of purchase and what its value would have been had the defendant's representations been true. *E.g.*, Price v. Philip Morris Inc., No. 00-L-112, 2003 WL 22597608, at *15 (Ill. Cir. Ct. Mar. 21, 2003); Aspinall v. Philip Morris Cos., 813 N.E.2d 476, 490 (Mass. 2004). Other states specify that the measure of damages is a refund of the purchase price. *E.g.*, N.J. STAT. ANN. § 56:8-2.11 (West 2001 & Supp. 2005).

[334] *See* SHELDON & CARTER, *supra* note 113, at 146–49.

[335] *See id.* at 179–89 (discussing defenses that do not apply to statutory misrepresentation claims).

[336] *See supra* notes 135–137 and accompanying text.

[337] *See supra* notes 143–148, 152–154 and accompanying text.

ages claim is unwarranted. By including the causation requirement in the private cause of action provisions, state legislatures signaled their intent that traditional reliance-causation limits apply to private damages actions.[338]

CONCLUSION

The problem of consumer class action abuse has been well documented. While CAFA has provided a modest step in the right direction, it fails to solve the underlying problem: relaxed substantive requirements that allow easy certification of misrepresentation class actions, regardless of the forum. The real remedy lies in enforcement of traditional reliance-causation requirements. Requiring reliance in private misrepresentation cases achieves the desired balance of public and private resources. Government agencies can seek restitution and injunctive relief before any harm occurs or where the negative value of the claim precludes private enforcement. The tort system, however, should be left to "those who have been wronged to seek redress from those who have wronged them."[339]

[338] *See, e.g.*, Weinberg v. Sun Co., 777 A.2d 442, 446 (Pa. 2001) ("Nothing in the legislative history suggests that the legislature ever intended statutory language directed against consumer fraud to do away with the traditional common law elements of reliance and causation.").

[339] Goldberg, *Tort Law for Federalists*, *supra* note 217, at 11.

ARTICLE

BENCHMARKING, CRITICAL INFRASTRUCTURE SECURITY, AND THE REGULATORY WAR ON TERROR

NICHOLAS BAGLEY*

The legal commentary surrounding terrorism, concerned with pressing constitutional questions, has to date largely ignored the role that regulatory agencies can play in reducing terrorism risks. This gap in the literature is peculiar because regulatory agencies have long assumed primary responsibility within the government for assessing and managing abstract risks and are therefore remarkably well-positioned to shore up our vulnerabilities against terrorist strikes. In an effort to turn the academic discussion to the regulatory aspects of the war on terror, this Article assesses the federal government's incipient regulatory effort to reduce terrorism risks to the nation's critical infrastructure. After concluding that the cornerstone of this administration's risk-reduction approach, the Critical Infrastructure Information Act of 2002, will fail to ensure that the government has adequate information to serve as a foundation for a coherent regulatory response, the Article proposes an alternative strategy based on a practice known as benchmarking. Under this approach, independent auditors would inspect high-risk firms within an industry and publicly rank those firms against each other according to their relative security vulnerabilities. Armed with this information, both the public and government regulators could bring continuous pressure to bear on private firms to reduce their vulnerabilities where most appropriate. Although a benchmarking approach would be plagued with some predictable weaknesses, the Article concludes that it nevertheless could prove to be an important feature of a comprehensive regulatory strategy to secure our critical infrastructure.

Homeland Security Secretary Michael Chertoff began his tenure by emphasizing that guarding against terrorist attacks requires a carefully calibrated *regulatory* response: "We cannot protect every person in every place at every moment. We cannot look in every container and every box. What we can do is use intelligent risk-based analysis, advanced technology and enhanced resources to manage that risk."[1] Yet the legal commentary sur-

* Law Clerk to Judge David S. Tatel, United States Court of Appeals for the District of Columbia Circuit, 2005–2006. J.D., New York University School of Law, 2005; B.A., Yale University, 2000. I would like to thank Richard Revesz, David Golove, Stephen Holmes, Rachel Barkow, Michael Burstein, and Kristina Daugirdas for their thoughtful comments.

[1] Michael Chertoff, U.S. Sec'y of Homeland Sec., Remarks at the George Washington Univ. Homeland Sec. Policy Inst. (Mar. 16, 2005) (remarks as prepared), *available at* http://www.dhs.gov/dhspublic/display?content=4391.

rounding terrorism, dominated by pressing constitutional questions,[2] has paid scant attention to the regulatory aspects of counterterrorism.[3]

Notwithstanding this lack of academic consideration, Chertoff is undoubtedly right that the security of the United States will fundamentally depend on whether the nation can capitalize on the potential of the regulatory state—an apparatus of tremendous sophistication that developed over the twentieth century in large part to mitigate the risks of an industrialized society—to maximize the possibility that rational decision-making can save lives and reduce the economic fallout from a devastating attack. Regulatory agencies have developed procedures for allocating scarce dollars to the most efficient risk-reduction programs; they have grappled with how to assess hard-to-quantify risks; they have focused attention on controversial valuations of human life and on the tradeoffs implicit in devoting scarce resources to projects that are not cost-effective; and they have developed and managed relationships with the private sector. They are our most potent counterterrorism weapons, and they have largely been ignored.

This Article intends to rectify this notable gap in the literature by refocusing the debate on the uneven American regulatory response to the war on terror.[4] It begins with the premise that terrorists with limited resources bent on inflicting maximum damage will concentrate on exploiting vulnerabilities in the nation's critical infrastructure. Whether by attacking chemical[5] or nuclear plants,[6] poisoning reservoirs,[7] releasing bio-

---

[2] *See, e.g.*, JEFFREY ROSEN, THE NAKED CROWD: RECLAIMING SECURITY AND FREEDOM IN AN ANXIOUS AGE (2004) (detailing risks of sacrificing constitutional freedoms because of terrorism fears); Bruce Ackerman, *The Emergency Constitution*, 113 YALE L.J. 1029 (2004) (arguing for an emergency constitution); David Cole, *The Priority of Morality: The Emergency Constitution's Blind Spot*, 113 YALE L.J. 1753 (2004) (arguing against an emergency constitution); Laurence H. Tribe & Patrick O. Gudridge, *The Anti-Emergency Constitution*, 113 YALE L.J. 1801 (2004) (same).

[3] *See* Eric A. Posner, *Fear and the Regulatory Model of Counterterrorism*, 25 HARV. J.L. & PUB. POL'Y 681, 681–83 (2002) (noting that a regulatory model of counterterrorism "deserves more public and academic attention than it has received"). Some academics have turned to questions relating to the quantification of terrorism risks. *See* RICHARD A. POSNER, CATASTROPHE: RISK AND RESPONSE (2004); *Special Issue on the Risks of Terrorism*, 26 J. RISK & UNCERTAINTY 99–249 (2003). None has carefully considered the adequacy of the country's overall regulatory response with respect to critical infrastructure security.

[4] For critiques of the Bush administration's invocation of the war paradigm to describe its efforts to weed out terrorism, see Bruce Ackerman, *This Is Not a War*, 113 YALE L.J. 1871 (2004), and David Golove & Stephen Holmes, *Terrorism and Accountability: Why Checks and Balances Apply Even in "The War on Terrorism,"* 2 N.Y.U. REV. L. & SECURITY 2 (2004).

[5] James V. Grimaldi & Guy Gugliotta, *Chemical Plants Are Feared as Targets: Views Differ on Ways to Avert Catastrophe*, WASH. POST, Dec. 16, 2001, at A01.

[6] *See* Address Before a Joint Session of the Congress on the State of the Union, 1 PUB. PAPERS 129, 130 (Jan. 29, 2002) [hereinafter State of the Union] ("[In Afghanistan we] have found diagrams of American nuclear powerplants and public water facilities, detailed instructions for making chemical weapons, surveillance maps of American cities, and thorough descriptions of landmarks in America and throughout the world.").

[7] *See id.*; *see also* Susan DeFord, *Water Safety a Priority for Utilities: Systems Exam-*

toxins into the food supply,[8] knocking out the national electrical grid,[9] blowing up natural gas lines,[10] disrupting telecommunications networks, or targeting financial centers, historical monuments, and transportation nodes,[11] terrorists can exploit the interdependence of critical systems and inflict serious harm at relatively low cost.[12]

These threats call for a regulatory regime that can both assess the risks posed by vulnerabilities in our critical infrastructure (risk assessment) and develop cost-effective methods to minimize those risks (risk management). The achievement of these twin goals is complicated by, among other things, the inconvenient fact that approximately eighty-five percent of the critical infrastructure in the United States is privately held.[13]

The cornerstone of the federal government's risk assessment strategy is the Critical Infrastructure Information Act of 2002 ("CIIA"),[14] enacted as part of the Homeland Security Act of 2002 ("HSA").[15] The CIIA attempts to promote information sharing by sweeping away certain disincentives that corporations previously faced when deciding whether to submit information about their facilities' security vulnerabilities to the government.[16] Most importantly, submissions to DHS that corporations designate as "critical infrastructure information" ("CII") are exempted from disclosure under the Freedom of Information Act ("FOIA").[17] The FOIA exemption, not incidentally, also prevents CII from falling into the hands of government-savvy terrorists who might otherwise look to DHS for ready-made assessments of how they can wreak the greatest havoc.

This Article begins by turning to administrative law principles to catalog the threefold trouble with this scheme. First, the CIIA will not promote any meaningful degree of public-private cooperation. Although

---

ined to Assess Vulnerability, WASH. POST, Oct. 24, 2002, at T3.

[8] Elizabeth Becker, *Shared Nightmare Over the Food Supply*, N.Y. TIMES, Dec. 11, 2004, at A14.

[9] Philip Shenon, *The Blackout of 2003: Domestic Security*, N.Y. TIMES, Aug. 15, 2003, at A24; Mark P. Mills & Peter W. Huber, *Can Terrorists Turn Out Gotham's Lights?*, CITY J., Autumn 2004, *available at* http://www.city-journal.org/html/14_4_gothams _lights.html (last visited Oct. 19, 2005).

[10] Ian Urbina, *Mapping Natural Gas Lines: Advise the Public, Tip Off the Terrorists*, N.Y. TIMES, Aug. 29, 2004, at A29.

[11] *See* MATTHEW BRZEZINSKI, FORTRESS AMERICA 7–10 (2004) (reciting a chilling litany of critical infrastructure vulnerabilities); Eric Lipton, *U.S. Report Lists Possible Terror Attacks and Likely Toll*, N.Y. TIMES, Mar. 16, 2005, at A1 (reporting on internal DHS assessments of the twelve most devastating terrorism scenarios, including biological attacks on subways or airports or the explosion of a truck bomb at a stadium).

[12] *The Lost Trail*, ECONOMIST, Oct. 22, 2005, at 15 ("[I]t does not cost much to commit an atrocity.").

[13] *See infra* note 25 and accompanying text.

[14] 6 U.S.C. §§ 131–134 (Supp. II 2002).

[15] Pub. L. No. 107-296, 116 Stat. 2135 (2002) (codified in scattered sections of 6 U.S.C. (Supp. II 2002)).

[16] *See infra* Part I.A.

[17] 6 U.S.C. §§ 133(a)(1)(A), (C) (Supp. II 2002) (referencing 5 U.S.C. § 552 (Supp. II 2002)).

it does remove one minor disincentive to sharing information, the statute simply does nothing to address industry's well-placed fear that the government will use CII as a basis for imposing costly security-related regulations on the private sector. Second, even assuming that DHS will receive a substantial amount of CII, what is it to *do* with it? There is an acute risk that the CII that DHS receives will be highly impressionistic, skewed by selection bias, and will under-represent the scope of the fallout from a potential attack. Furthermore, by refusing to impose any objective guidelines to standardize CII submissions, DHS has sharply limited its capacity to use CII to assess the relative risk profiles of different critical infrastructure facilities and prioritize risk-reduction efforts. Third, the CIIA's reliance on secrecy to ensure that CII does not inadvertently get into the hands of terrorists will produce only modest benefits even as it assures the private-sector domination of DHS's decision-making. Although nixing the CIIA's FOIA exemption would not by itself invigorate public oversight, throwing a veil over our critical infrastructure protection efforts will predictably lead to poor priority-setting and irrational decision-making.

The Article then turns to risk management and argues that stubborn gaps in the set of risk-related information available to a regulator—gaps that cannot be filled by the CIIA—will limit the efficacy and efficiency of conventional risk management approaches. Reliance on conventional strategies therefore could potentially divert scarce resources away from the cost-effective protection of critical infrastructure and toward expensive compliance with minimally effective regulations.

In light of these shortcomings, this Article proposes a different kind of regulatory response, one that could further both risk assessment and risk management goals. A new wave of administrative law literature has embraced informational strategies, and particularly a practice known as benchmarking, to deal with difficult regulatory problems that are not amenable to solution by traditional regulatory approaches. This Article suggests that an effective regulatory answer to the terrorist threat might implement a benchmarking regime incorporating three major elements. First, firms that own and operate critical infrastructure could be required to undertake vulnerability assessments and provide them to DHS. Second, the CII could be compiled by independent inspectors who could standardize security information across a limited number of variables so as to facilitate comparisons of the relative security of facilities, firms, and industries. Third, this benchmarked information could be publicized and made easily accessible. Public and market pressure could thereby be focused on those firms that pose the greatest threats to homeland security, prompting meaningful and continuous reductions in terrorism vulnerabilities while reserving to firms the choice as to the most cost-effective means of making security improvements.

To be sure, a benchmarking regime would have some predictable weaknesses—most notably that some firms will predictably lag behind their competition; that assessments of facility-wide security gaps will be bad proxies for understanding the scope of our security risks; and that improvements in our risk profile will not be tethered to cost-benefit analysis and will consequently be quite vulnerable to public overreaction. Benchmarking nevertheless holds substantial promise as a mechanism to assess and manage the risks that the terrorist threat poses to our critical infrastructure.

I. BACKGROUND

A. *The Threat to Our Critical Infrastructure*

To begin with, definitions. What is "critical infrastructure"? The CIIA defines it as those "systems and assets, whether physical or virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters."[18] Importantly, this defines critical infrastructure *not* with reference to the identity of the target, but by the consequences of an attack on it. DHS, citing a desire for "flexibility,"[19] has declined to elaborate on the meaning of critical infrastructure.[20] The term is better fleshed out in the Bush administration's *National Strategy for the Physical Protection of Critical Infrastructures and Key Assets* ("*National Strategy*"), which breaks down the country's critical infrastructure into thirteen separate "sectors" (agriculture, food, water, public health, emergency services, government, defense industrial base, information and telecommunications, energy, transportation, banking and finance, chemical industry and hazardous materials, and postal and shipping) plus two categories of "key assets" (national monuments and particularly high-risk targets like nuclear power plants, dams, and hazardous materials storage facilities).[21] A rough assessment that they would make "lucrative" or "high-payoff" targets links these "highly complex, heterogeneous, and interdependent . . . facilities, systems, and functions."[22]

Whether something qualifies as "critical infrastructure," then, turns not on its importance to our national economy per se, but rather on an as-

---

[18] 6 U.S.C. § 101(4) (Supp. II 2002) (referencing Critical Infrastructures Protection Act of 2001, 42 U.S.C. § 5195c(e) (Supp. I 2001)).

[19] 6 C.F.R. § 29.2 (2005).

[20] *See* Procedures for Handling Critical Infrastructure Information, 69 Fed. Reg. 8073, 8076 (Feb. 20, 2004) (interim rule) [hereinafter CII Procedures].

[21] WHITE HOUSE, NATIONAL STRATEGY FOR THE PHYSICAL PROTECTION OF CRITICAL INFRASTRUCTURES AND KEY ASSETS 6 (2003), *available at* http://www.whitehouse.gov/pcipb/physical_strategy.pdf [hereinafter NATIONAL STRATEGY].

[22] *Id.* at 2.

sessment of the total costs, including cascade effects, associated with a successful terrorist attack on that "facility, system, or function." The CIIA thus leaves it to DHS to judge whether a target is enough of a "high-payoff" to make it "critical." The necessary implication is that the relative critical-ness of particular infrastructures will "evolve"[23] as risk assessments improve.

According to the *National Strategy*, the critical infrastructure of the United States includes—among many other facilities—over 66,000 chemical processing plants, 26,600 financial institutions, 5000 public airports, 1800 federal reservoirs, 1600 municipal waste water facilities, 2800 electrical plants, and 80,000 dams.[24] Because approximately eighty-five percent of the facilities likely to be deemed "critical" are in private hands,[25] the key to improving risk assessments and thereby shoring up the nation's vulnerable infrastructures will be enlisting private industry to contribute to homeland security efforts.

B. "Cooperation and Partnership" Through Secrecy

Although Congress and the White House had a number of options, several of which are discussed below, the government's chosen strategy for working with the private sector is to rely on a "new paradigm" of "cooperation and partnership."[26] The CIIA ostensibly facilitates partnership by providing for the secrecy of CII that private industry voluntarily submits to the Critical Infrastructure Protection Program ("CIPP") at DHS.[27] Because DHS has no authority to compel private industry to disclose its security gaps,[28] in practice all such information will be supplied "voluntarily."[29]

CII—defined in vague terms by the CIIA and DHS regulations as information "not customarily in the public domain and related to the secu-

---

[23] *Id.* at 35.

[24] *Id.* at 9.

[25] The source of the eighty-five percent figure is obscure, but journalists and government officials throw it around with abandon. *See, e.g.*, *id.* at 8; John Mintz, *U.S. to Keep Key Data On Infrastructure Secret: Firms Encouraged to Report Security Gaps*, WASH. POST, Feb. 19, 2004, at A21. Given the fluidity of the definition of "critical infrastructure," this number cannot be anything more than a very rough estimate.

[26] NATIONAL STRATEGY, *supra* note 21, at 8, 20 ("Safeguarding our critical infrastructures . . . from terrorism in today's fluid marketplace and threat environment requires a new, more cooperative set of institutional relationships and attitudes. The need for partnering is clear.").

[27] 6 U.S.C. § 133 (Supp. II 2002).

[28] *See* Mintz, *supra* note 25, at A21 ("U.S. officials have no power under the Homeland Security Act to compel industries to provide data about their security gaps.").

[29] This includes information that is required to be submitted to other agencies, so long as the information is not being provided to DHS in lieu of the firm's other regulatory obligations and DHS has not exercised its (nonexistent) legal authority to compel disclosure. DHS Protected Critical Infrastructure Information, 6 C.F.R. § 29.2 (2005); *see also* CII Procedures, *supra* note 20, at 8076.

rity of critical infrastructure"[30]—cannot be disclosed in response to FOIA requests, is not subject to any restrictions on ex parte communications, and cannot be used by anyone in any civil action.[31] Ideally, by alleviating industry's concerns that damaging CII might become public or might be used in enforcement actions against them, firms will voluntarily provide needed CII to the federal government. The FOIA protection also serves the secondary purpose of preventing terrorists from exploiting open-government laws to learn about the country's vulnerabilities.[32]

Once it has compiled enough information, DHS will engage in a risk assessment of our critical infrastructure.[33] Risk management, naturally, will come next; in coordination with federal, state, and local officials, the CIIA requires DHS to "recommend measures necessary to protect the . . . critical infrastructure of the United States . . . ."[34]

Although the details are hazy, the guiding principle of the Bush administration's strategy for managing critical infrastructure risks is to look to the market.[35] The premise is that "[c]ustomarily, private sector firms prudently engage in risk management planning and invest in security as a necessary function of business operations and customer confidence."[36] The federal government's quite limited role in risk management will therefore be to coordinate public and private security responses, provide a clearinghouse for information on security threats, and facilitate intra-industry cooperation by identifying, promoting, and sharing "industry-specific best practices."[37] This healthy cooperative enterprise will (in theory) "encourage[ ] the private sector to make prudent investments earlier and at all levels of the risk management spectrum."[38]

The *National Strategy* concedes that, in some circumstances, the threat will "excee[d] an enterprise's capability to protect itself beyond a reason-

---

[30] 6 U.S.C. § 131(3) (Supp. II 2002). DHS Protected Critical Infrastructure Information, 6 C.F.R. § 29.2 (2005).

[31] 6 U.S.C. § 133(a)(1) (Supp. II 2002).

[32] *See* Comment, Kristen Elizabeth Uhl, 53 AM. U. L. REV. 261, 264–65 (2003) (reporting that FOIA restrictions were a "knee-jerk reaction" to investigators' discovery of "'detailed maps and drawings of sensitive infrastructure locations' in caves in Afghanistan and Al Qaeda training camps").

[33] 6 U.S.C. § 121(d)(2) (Supp. II 2002):

> [DHS shall] carry out comprehensive assessments of the vulnerabilities of the . . . critical infrastructure of the United States, including the performance of risk assessments to determine the risks posed by particular types of terrorist attacks within the United States (including an assessment of the probability of success of such attacks and the feasibility and potential efficacy of various countermeasures to such attacks).

[34] 6 U.S.C. § 121(d)(6) (Supp. II 2002).

[35] *See* NATIONAL STRATEGY, *supra* note 21, at 12 ("[T]he federal government strives to encourage proactive, market-based protective solutions.").

[36] *Id.* at x.

[37] *Id.* at 20.

[38] *Id.*

able level of security investment" and therefore "[t]he private sector may . . . require incentives to stimulate investment."[39] While the administration has embraced the possibility of subsidies, however, it eschews top-down regulation for at least two reasons: first, because private industries know best how to protect their facilities against terrorist threats; and second, because regulation is anathema to building a culture of trust that is necessary for the success of a voluntary program.[40]

II. SHORTCOMINGS

A. Risk Assessment: The Three-Part Failure of the CIIA

The CIIA will fail to achieve its ambitious risk assessment goals for three distinct reasons. First, it will not provide the necessary incentives for companies to voluntarily provide CII to the federal government. Second, even if it were to promote public-private cooperation, the CIIA's scattershot approach to CII would result in the production of skewed and impressionistic information that DHS would be unable to compile and analyze in a meaningful way. Third, the FOIA exemption simultaneously hampers public oversight of DHS's efforts to promote critical infrastructure security and ensures that industry groups will be well-positioned to dominate agency decision-making.

1. Failing To Facilitate Public-Private Cooperation

Most of the considerable controversy surrounding the CIIA has come from detractors who argue that the CIIA's secrecy protections are too solicitous of private industry.[41] This Article takes the opposite view: the real problem with the CIIA is that it does not go far enough in accommodating the private sector. Firms will predictably underinvest in terrorism-related security, which will in turn make them understandably wary about providing the government with information that could be used to craft costly security-related regulations. Because the government cannot credibly bind itself to the promise that the provision of CII will not lead to the imposition of regulations, private firms will lack sufficient incentives to voluntarily participate in the CIPP. The CIIA is consequently likely to prove grossly inadequate as an information-gathering tool.

---

[39] *Id.* at x, 23–24.

[40] *Id.* at 8.

[41] *See, e.g.*, Uhl, *supra* note 32, at 290–304; Editorial, *Fix This Loophole*, WASH. POST, Feb. 10, 2003, at A20.

a. Homeland Security as a Public Good

A rational owner of a piece of critical infrastructure will mitigate its vulnerabilities to a terrorist attack only so long as her marginal cost of paying for additional security is less than the marginal benefit she gets from improved security. However, a rational owner will weigh only her *private* benefit in calculating how much security to provide, and will ignore benefits that inure to the public at large.[42]

In every case involving critical infrastructure, the social benefit of secure critical assets (i.e., the diminution of the risk of a "debilitating impact on security, national economic security, national public health or safety"[43]) will dwarf an owner's private benefit in securing her own piece of critical infrastructure (i.e., the diminution of the risk that the value of her investment will be destroyed). To provide just one example, EPA has estimated that 123 separate chemical plants in the United States store toxic chemicals that, if released from any single plant into the atmosphere as the result of a terrorist attack, could endanger well over a million people.[44] The vast public benefits associated with reducing the risk of a catastrophe of this magnitude will certainly far outstrip a plant-owner's private benefit in shoring up her facility's security.

This divergence between private and social benefits means that rational owners and operators of critical infrastructure will systematically provide a suboptimally low level of security.[45] Critical infrastructure security is a classic example of a "nonexcludable public good,"[46] similar to a national military or clean air, that the private market simply will not provide at efficient levels.[47] Competitive pressures cement this tendency by allowing firms that underinvest in security to obtain a competitive advantage relative to firms that do expend resources on terrorism security.[48]

---

[42] *See* Garrett Hardin, *The Tragedy of the Commons*, 162 SCIENCE 1243 (1968).

[43] *See* 42 U.S.C. § 5195c(e) (Supp. I 2001) (*referenced in* 6 U.S.C. § 101(4) (Supp. II 2002)).

[44] Grimaldi & Gugliotta, *supra* note 5, at A30.

[45] For at least some facilities, it is possible that the risk of a terrorist attack is so low that, however grave the potential fallout from an attack, any additional investment in security would prove more costly than beneficial. *Cf.* W. Kip Viscusi & Richard J. Zeckhauser, *Sacrificing Civil Liberties to Reduce Terrorism Risks*, 26 J. RISK & UNCERTAINTY 99 (2003) (investigating whether tradeoffs in civil liberties pass cost-benefit muster). Even if this were true for some facilities, however, it is unlikely to be the case across the board. *See* Lipton, *supra* note 11, at A16 (reporting on DHS's estimates of enormous costs associated with several types of devastating attacks).

[46] A nonexcludable public good is a commodity that (a) cannot be restricted only to those individuals that pay for it and (b) for which the marginal cost of an additional person's consumption is zero. MICHAEL L. KATZ & HARVEY S. ROSEN, MICROECONOMICS 613–17 (3d ed. 1998).

[47] *Id.*; *see also* Eric Pianin & Bill Miller, *Businesses Draw Line on Security: Firms Resist New Rules for Warding off Terror*, WASH. POST, Sept. 5, 2002, at A1.

[48] *See* Howard Kunreuther & Geoffrey Heal, *Interdependent Security*, 26 J. RISK & UNCERTAINTY 231, 231 (2003) (reporting that, in a game-theoretic model of terrorism reduction, the incentives for parties to reduce their security vulnerabilities approaches zero as

Because of the structure of this problem, the federal government's apparent belief that market forces will normally prompt the private sector to invest in "prudent" levels of security is fanciful.

b. The Cost and Benefits of Information Sharing

Private industries normally understand better than the government how vulnerable they are to a terrorist attack and how little they have chosen to invest in security. The private sector's biggest disincentive to sharing CII will consequently be its fear that, when the true scope of its vulnerabilities become known, legislators will force it to adopt more stringent security standards.[49] The political appointees at DHS are keen to this; Robert Liscouski, who until recently led CIPP as the assistant secretary for infrastructure protection, has said "that while some infrastructure industries are eager to discuss security issues with the government, others are dragging their feet for fear of prompting later government requirements that they spend money to protect their networks."[50]

Because the CIIA does nothing to assuage companies' well-placed fear of regulation, it is not surprising that, in the first ten months of CIPP's operation, only thirty-one firms out of the hundreds of thousands that own pieces of critical infrastructure in the United States have submitted *any* CII to DHS.[51] (CIPP declined this author's query as to whether these disclosures came predominantly from any particular sector on the grounds that the CIIA prohibited disclosure because "identifying the sectors could risk disclosure of the submitters' identity or contents of the submissions."[52] The response is baffling: knowing from which *sectors* CII contributions have come could not, without more, possibly reveal the identity of individual firms that have submitted CII—much less the *content* of that information.)

The CIIA's FOIA exemption does not by itself, however, provide any benefits to industry. The exemption serves only to minimize the risk of

---

the number of parties increase).

[49] *See* Cary Coglianese et al., *Seeking Truth for Power: Informational Strategy and Regulatory Policymaking*, 89 MINN. L. REV. 277, 290 (2004) ("Firms usually have an interest in maintaining silence, in withholding or not even generating information that would help government regulate.").

[50] Mintz, *supra* note 25, at A21.

[51] E-mail from Emily Rochelle Hickey, Senior Communications Officer, Protected Critical Infrastructure Info. ("PCII") Office, Dep't of Homeland Sec., to Nicholas Bagley, author (Dec. 27, 2004, 12:51:18 EST) (on file with author); *see also Hearing before the H. Select Subcomm. on Infrastructure and Border Sec.*, Apr. 21, 2004 (statement of Robert Liscouski, Assistant Secretary for Infrastructure Protection), *available at* http://www.iwar.org.uk/cip/resources/energy/liscouski091703.htm (reporting that only two firms and two associations participated between February 2004 and April 2004).

[52] E-mail from Emily Rochelle Hickey, Senior Commc'ns Officer, PCII Office, Dep't of Homeland Sec., to Nicholas Bagley, author (Mar. 5, 2005, 15:36:24 EST) (on file with author).

public disclosure of sensitive CII voluntarily submitted to DHS. But a firm is under no obligation to participate in CIPP, and can achieve the same level of secrecy simply by choosing not to submit any CII information in the first place. Moreover, firms cannot use DHS as a "black hole" in which to hide information that would otherwise have come to light:[53] the FOIA exemption does not extend to the information submitted to another agency, "even where it is identical to information voluntarily submitted to DHS pursuant to the [CIIA]."[54] The same goes for the grant of civil immunity for CII disclosures and the exemption of CII from *ex parte* communication rules:[55] the CIIA only protects information that the private sector does not otherwise disclose to the government.

What, then, *are* the benefits of participation in CIPP? When asked at a Senate hearing, Liscouski replied that "[t]he motivation of the companies giving the information . . . is 'doing public good in protecting the country.'"[56] In interviews, he has repeated a similar belief that "the private sector will act responsibly . . . [and] do the right thing on their own."[57] As an initial matter, premising a critical initiative on the patriotism of profit-seeking corporations seems quite unwise. Even conceding that patriotism is a meaningful motivator for a substantial fraction of firms, however, abstract nationalism will not frequently outweigh concrete fears that the end result of information sharing will be costly regulations.

Beyond patriotism, however, there may be some occasional benefits associated with corporate participation in a voluntary government program. Most importantly, some industries may have cause to believe that the government will subsidize their security investments. Firms within those industries may therefore submit damning CII in an effort to extract greater rents from the government.[58]

This brinksmanship strategy is quite risky, however. As explored below, the government is likely in many cases to force the private sector to bear the costs of additional security.[59] At least in most cases, industries are

---

[53] *See Bush Seeks FOIA Exemption in Homeland Security Bill*, OMB WATCH, June 16, 2002, http://www.ombwatch.org/article/articleview/844/1/246?TopicID=1 ("[T]his exemption could create a black hole into which companies dump information, and avoid public scrutiny."); *see also* 149 CONG. REC. S3632 (2003) (statement of Sen. Leahy) (arguing that the CIIA "effectively allows companies to hide information about public health and safety from American citizens simply by submitting it to DHS").

[54] 6 C.F.R. § 29.3(a) (2005).

[55] *See* 6 U.S.C. §§ 133(a)(1)(B), (C) (Supp. II 2002).

[56] Mintz, *supra* note 25, at A21.

[57] BRZEZINSKI, *supra* note 11, at 189.

[58] As OMB Watch, a Washington public-interest group, has explained: "Conceivably companies reluctant to improve old failing infrastructure could submit information to the CII program and seek financial assistance from the government. The company could attempt to leverage grants, low interest loans and other government resources, or threaten to leave the problem unresolved." Letter from Sean Moulton, Senior Policy Analyst, OMB Watch, to Janice Penya, Office of the General Counsel, Dep't of Homeland Sec. 4 (May 20, 2004), *available at* http://www.ombwatch.org/ homeland/InterimFinalCII-OMBW.pdf.

[59] *See infra* Part II.A.1.c.

more likely to shy away from information sharing than gamble on the chance that the government will subsidize their security efforts.

Alternatively, some firms may submit information in the hopes that it will provoke the imposition of regulations because their critical infrastructures are relatively cheaper to secure than those of their competitors.[60] And some firms may take a long view, and believe that the value of establishing relationships with the nascent DHS outweighs the risk that regulations will be slapped on them. Particularly if they think that security regulations will inevitably be enacted, those relationships might be leveraged to influence the content of those regulations.

In the final estimation, however, these potential benefits pale in comparison to the costs of regulations that require further investments in security measures. In general, firms will therefore avoid cooperating with the government.[61]

c. Perverse Institutional Incentives

This lack of voluntary private-sector cooperation puts the directors of DHS in a bit of a bind. To make CIPP successful, DHS will have to credibly assure firms that no regulation based on the CII will be forthcoming. Liscouski, CIPP's director, is pursuing this strategy assiduously, and has gone on the record to say in no uncertain terms that "[o]ur job is not to regulate. . . . Regulating is not our role."[62]

DHS will face two difficulties in making these assurances credible. First, DHS has no capacity to bind the government into the future. Even if it recklessly promises not to enact regulations today, it can not promise that it will not enact regulations tomorrow. Any such promise, if made, would last only as long as the sitting administration (if even that long), after which point DHS may prove all too willing to use the information garnered under the CIIA to recommend costly regulatory options.

Second, no one seriously expects that the government will have the resources or the will to foot the entire bill for securing our critical infrastructure. As an initial matter, subsidies are costly to the government; regulations are not. Regulations are therefore quite appealing to budget-conscious legislators. Whatever its protestations to the contrary, the government inevitably will flex the state's coercive powers to mandate the provision of a more efficient level of protection against potential terrorist attacks. Indeed, this is already happening: despite its demonstrated willingness to spend governmental funds to improve airline security,[63] the

---

[60] Coglianese et al., *supra* note 49, at 291.

[61] *Id.* at 288 ("Government cannot count on self-interested holders of information to reveal it fully and without bias.").

[62] BRZEZINSKI, *supra* note 11, at 189.

[63] Eric Lichtblau, *Security Report on U.S. Aviation Warns of Holes*, N.Y. TIMES, Mar. 14, 2005, at A1.

federal government is now insisting that airports raise fares and borrow money to pay the additional $5 billion required for new baggage screening equipment.[64] We should expect this pattern—governmental regulation followed by squabbling over who should pay for it—to repeat itself many times over.

Even if subsidies succeeded in boosting overall levels of expenditures in security improvements, they are nevertheless an imprudent way to go about protecting our critical infrastructure because they are likely to prove terrifically inefficient. Security costs will vary among different firms in the same industry, depending on a number of different factors (e.g., their level of inter-connectedness with other parts of our critical infrastructure, their location, their pre-existing level of security). If it enacted regulations requiring that firms meet a baseline security standard, the government would force industries to treat threat-abatement costs as simply another production input. This would, in turn, give a competitive advantage to those firms that could more cheaply abate security risks.

Subsidies, however, allow those firms that are peculiarly vulnerable to terrorist threats to continue in operation even if shutting those firms down would be a preferable alternative. For one example, shoring up the security at Indian Point, a nuclear power plant located just thirty-five miles from Manhattan, may not make sense when nuclear reactors in sparsely settled areas of the country can be adequately protected for far less.[65] Thus one disturbing irony of the White House's *National Strategy* is that its ostensible "market solution" is affirmatively designed to retard the effective functioning of the market.

In sum, the CIIA has created a system in which the government must abandon regulation as a risk management technique *before* it can successfully gather the information necessary to assess our risk profile. Just as they have proven necessary to combat a wide variety of other social harms, however, regulations are likely to be an invaluable part of our response to the terrorist threat. No one seriously argues that "voluntary programs" and "subsidies" would be adequate to prevent private firms from spilling a suboptimally high level of pollution into our air or water.[66] And no one invokes a chemical manufacturer's "patriotism" in explaining why it will take sufficient precautions when disposing of hazardous materials. Why we should expect owners and operators of elements of our critical infrastructure to be more solicitous of the public good is mystifying. The efficacy of the CIIA as a risk assessment tool should not depend on the

---

[64] Thomas Frank, *Airports Push for U.S. Aid*, USA TODAY, Mar. 14, 2005, at A1.

[65] *Cf.* Rory Kennedy, Op-Ed., *A Target on the Hudson*, N.Y. TIMES, Sept. 5, 2004, at 14WC (describing danger of Indian Point's location in a "more densely populated area than any other nuclear plant in the country").

[66] *See, e.g.*, Guy Gugliotta & Eric Pianin, *Bush Plans on Global Warming Alter Little; Voluntary Programs Attract Few Firms*, WASH. POST, Jan. 21, 2004, at A01 (documenting dismal failure of voluntary plans to curb global warming).

government's willingness to fight the regulatory war on terror with one hand tied behind its back.

2. *The CIIA's Obstruction of Careful Analysis*

But assume for a moment that a substantial fraction of firms did participate in CIPP. At that point, DHS's job would shift from *collecting* information to *analyzing* it. Oddly, however, the CIIA works *against* rigorous analysis in two important ways. First, the information that DHS does receive will be tainted by selection bias, contributing to a misleading and highly impressionistic picture of our vulnerabilities and of the fallout associated with a terrorist attack that exploits those vulnerabilities. Second, the CII that firms do submit will be qualitative and will not be keyed to any standardized assessment of their facilities' security profiles. DHS will have only a highly circumscribed capacity to use that information to tell whether a particular firm should invest more in security measures, much less make comparisons of relative security risks within an industry or across industries. The lack of quantified standards means that neither DHS nor the public will have any way to hold firms accountable (either through subsidies or regulations) for achieving a greater measure of security.

a. *Selection Bias*

The more effectively that firms can downplay the risk that their facilities pose to homeland security, the less likely that regulation will be forthcoming—or, if regulation does come, that it will be excessively burdensome.[67] Firms that fear the prospect of costly regulation will consequently have enormous incentives to provide information that minimizes the extent of their true security gaps and the damage that a terrorist attack might cause.

Since firms have absolute discretion in choosing what CII goes (and, more importantly, does not go) to DHS, they can manipulate the CII that they submit to DHS so as to provide the agency with an incomplete and misleading picture of the risk profile of our nation's critical infrastructure. The set of CII with which DHS will work will therefore be highly prone

---

[67] As explained earlier, see *supra* note 61 and accompanying text, under normal circumstances only a small fraction of firms will voluntarily submit CII to DHS, perhaps because they are likely to receive subsidies or they believe regulations will confer on them a competitive advantage. That CII will itself be skewed, albeit toward *emphasizing* the risks posed by gaps in critical infrastructure security. In an effort to demonstrate difficulties in analyzing large amounts of CII, however, the following discussion assumes that a large majority of firms will provide information to CIPP—but that in providing CII, their incentives to downplay their vulnerabilities remain. Under either scenario, however, the central point stands that voluntarily submitted CII will reflect *some kind* of selection bias and will be an imprecise and skewed guide to decreasing critical infrastructure risks.

to multiple levels of selection bias and may affirmatively hinder the agency's capacity to make accurate risk assessments and set thoughtful risk-reduction priorities.[68]

Consider first an assessment of a facility's relative vulnerability to terrorist attacks. A comprehensive assessment of any facility's security profile will necessarily take into account a wide array of variables that will be relevant to assessing how vulnerable a facility is to terrorist attacks. Those variables will include the facility's layout, the architectural plans of its buildings, the security measures already in place (both live security and automated security), an analysis of the efficacy of those security measures, the facility's proximity to police responders, the set-up of a facility's electrical system and its computer networks, the composition and background of the facility's work force, and so on.

Any given facility will score better on some dimensions of this risk profile and worse on others. These differences across hundreds of different variables give firms substantial latitude to present their facilities in the most flattering light possible to DHS, resulting in profound selection bias. One firm with a sprawling and hard-to-secure facility may, for example, decline to provide information about the facility's layout even as it submits every shred of evidence relating to investments in securing its primary building. In contrast, a firm with a compact facility will be more inclined to share information on its layout while neglecting to tell DHS that it has done little to secure the facility.[69] Moreover, many firms own multiple facilities—allowing them to put their best foot forward by selectively submitting CII on only those facilities that are most secure.

Compounding these two levels of selection bias, firms that have lagged behind their competition will be less inclined to participate in CIPP than their relatively more secure counterparts. Even worse, whole industries that pose relatively greater threats to homeland security will opt not to participate in CIPP at higher rates than their less vulnerable counterparts.

These four levels of selection bias—within a facility, across facilities owned by a single firm, across firms, and across industries—will systematically operate to skew any CII that firms do happen to submit to the

---

[68]

> Selection bias is commonly understood as occurring when the *non*random selection of cases results in inferences, based on the resulting sample, that are not statistically representative of the population . . . . A common problem arising from such selection is that it may over represent cases at one or the other end of the distribution on a key variable.

David Collier, *Translating Quantitative Methods for Qualitative Researchers: The Case of Selection Bias*, 89 AM. POL. SCI. REV. 461, 462 (1995).

[69] *Cf.* Susan DeFord, *supra* note 7, at T26 ("Whereas a nuclear plant is concentrated [in one location], it's more of a daunting task to protect every [water] storage tank, every intake site.") (quoting Martin J. Allen, Director of Technology Transfer for the Research Foundation of the American Water Works Association in Denver)).

agency. The information that DHS will have at its disposal will consequently paint a rosier picture of the security profiles of our country's facilities than is warranted.[70]

The relative vulnerability of various facilities, firms, or industries is only half the picture, however. To understand what additional security measures are or are not appropriate, DHS must also gauge the potential fallout from a successful terrorist attack at a particular facility. Because firms will have strong incentives to downplay the kinds of risks that their facilities pose, the four types of selection bias detailed above will operate with equal force to damage assessments. Take the risks posed by facilities that process dangerous chemicals as an example. A facility that divulges information about the minor risks posed by the release of a relatively innocuous chemical may balk at sharing information about more dangerous chemicals, even if both chemicals are present at one facility in large amounts. Similar pressures would operate to skew information relating to damage assessments across facilities, firms, and industries.

Reduced to the simplest terms, firms have both the motive and the opportunity to use their CII submissions to downplay the risk of a terrorist threat to their facilities.[71] DHS's blithe assumption that, in making its risk assessments, it can rely on what we can safely predict will be an impressionistic and highly skewed set of information is stunning.

b. Lack of Standards

Even assuming that DHS did receive a substantial amount of information from private firms that fairly represented the vulnerabilities of our critical infrastructure, the next step would involve compiling and analyzing that information so as to fashion an assessment of the risks posed by those vulnerabilities. That assessment would ideally include determinations of the likelihood that terrorists could *successfully* mount a particular attack on a particular facility, taking into account that facility's security profile (the nature of risk) and the damage that such an attack would likely cause (the extent of risk).[72] Alternatively, acknowledging that predictions about the likelihood of terrorist attacks are highly contingent, a risk assessment would *at least* allow for comparisons of the relative risks between facilities so as to better understand how we could most efficiently allocate limited resources to critical infrastructure protection.

---

[70] It is true that if DHS were aware that the CII it received predictably understated true risks, it could attempt to compensate for that by skewing its final risk assessments in the opposite direction. It would be an exceedingly peculiar regulatory approach, however, that stressed the *need* for information but was premised on the explicit *distrust* of that same information.

[71] *See* Coglianese et al., *supra* note 49, at 288.

[72] Risk assessments gauge the nature and scope of risk. *See* RICHARD L. REVESZ, FOUNDATIONS OF ENVIRONMENTAL LAW AND POLICY 50 (1997). *See generally* WILLIAM D. RUCKELSHAUS, RISK, SCIENCE, AND DEMOCRACY (1985).

In practice, however, the CII that DHS receives will be difficult to use in assessing either the nature or the extent of risks. CIPP's first problem is one of scale. The sheer size of our sprawling infrastructure means that, if the CIIA were successful, CIPP would be inundated with a veritable blizzard of information from hundreds of thousands of different sources. CIPP would be in a similar situation as the FBI found itself after the enactment of the USA PATRIOT Act. Although the PATRIOT Act made it easier for the FBI to acquire terrorism-related information through secret wiretaps and electronic surveillance,[73] it was so overwhelmed and understaffed that by September 2004 it had amassed a backlog of over 120,000 hours of recordings that it had not yet translated or analyzed.[74]

A more profound objection to CIPP—and one not amenable to resolution by an increase in spending—is that the lack of *any* standards to guide CII submissions will necessarily frustrate thoughtful analysis. To see why, consider first that the CIIA's definition of CII is so vague that "it will cover, and thus keep secret, virtually anything a company decides to fork over."[75] DHS has refused to provide any further guidance as to what might constitute CII, or as to how a firm might best organize that information to facilitate analysis. DHS is instead

> [M]indful that private sector submitters, as the owners and operators of most of the nation's critical infrastructures, are the most well versed as to what information in their particular sector or industry might qualify as CII; therefore, [DHS] does not wish to unduly restrict the scope of what may be submitted as CII under the Act.[76]

The result of this deferential approach is that firms can designate anything from architectural plans to maps of computer networks to security contracts to inventory rosters to evacuation plans to employee background checks as CII. Moreover, because of the utter lack of standards as to what does (or does not) constitute CII, submissions from different firms will vary dramatically.

DHS's job will be to take these unstandardized paper submissions and somehow transform them into risk assessments.[77] DHS has said very little about how it proposes to go about this, but using standardless CII to make any meaningful assessment of risk is certain to be a challenge. Consider first the problem at the facility level. An employee at CIPP will

---

[73] STEPHEN L. SCHULHOFER, THE ENEMY WITHIN 29–48 (2002).

[74] Eric Lichtblau, *F.B.I. Said to Lag on Translations of Terror Tapes*, N.Y. TIMES, Sept. 28, 2004, at A1.

[75] *See* Editorial, *Fix This Loophole*, WASH. POST, Feb. 10, 2003, at A20.

[76] Procedures for Handling Critical Infrastructure Information, 69 Fed. Reg. 8076 (Feb. 20, 2004) (interim rule).

[77] NATIONAL STRATEGY, *supra* note 21, at 35.

have at her disposal an enormous *amount* of information, but very little sense of how that information should be aggregated and analyzed so as to make nuanced judgments about how vulnerable that facility is to a terrorist attack. She might know that a facility has invested in a particular kind of security system, for example, but have little sense of the efficacy of that system when applied to that particular facility. She might know that the facility has invested resources in protecting itself against one kind of attack, but not know whether it would be appropriate for the firm to protect against a different type of threat. And she might know that the facility is located close to a major metropolitan center, but not understand to what degree that should color her analysis. Even if she could assume that the information at her disposal was complete and wholly reliable—which she probably could not—her judgments about the relative safety of the facility she is evaluating will almost certainly be highly general and fundamentally qualitative.

This problem will be more acute when DHS attempts to compare the qualitative assessments of different facilities with each other. Recourse to CII—which will vary across facilities, firms, and industries and which provides nothing more than a partial paper image of a firm's security profile—will provide, at best, a partial answer.

Even worse, DHS will have a difficult time comparing security across industries. Different kinds of threats pose different degrees of risk, and every industry will need to shore up its security in different ways to protect against different kinds of terrorist attacks. Appropriate security precautions at a nuclear plant will not be appropriate at a water-treatment plant, and vice-versa. But CII that is not standardized across industries will provide only the roughest guide to the relative threats that gaps in the security of our critical infrastructure pose.

I do not want to overstate these difficulties. Risk assessments based on qualitative and standardless CII would be an improvement on risk assessments undertaken without reference to *any* private information. DHS could, for example, use CII to garner a deeper understanding of industry best practices and how best to protect facilities against the most likely terrorist threats. The agency could undoubtedly develop some measure of expertise with these issues over time, identifying leaders and laggards and focusing legislative attention on what it deemed to be the most critical problems.

But the core problem with CIPP analysis as it now stands is summed up nicely by that old standby of sound management: what doesn't get measured won't get done.[78] Even if DHS received abundant information about what firms are doing, it would have little capacity to assess whether those efforts were improving homeland security. As in any organization in

[78] *Cf.* Louis Lowenstein, *Financial Transparency and Corporate Governance: You Manage What You Measure*, 96 COLUM. L. REV. 1335 (1996).

which outcomes are hard to measure, DHS will struggle to find ways to ensure that firms are in fact making needed investments in security.[79] Without recourse to the opinions of trained experts on the ground who can inspect facilities and give a holistic, contextual, and, above all, *quantitative* sense of the degree of a facility's security against particular kinds of terrorist attacks, CIPP will be hard pressed to make the kind of rigorous risk assessments needed to address the vulnerabilities of our critical infrastructure.

3. *The Trouble with the FOIA Exemption*

The CIIA's FOIA exemption ostensibly serves a secondary salutary purpose of protecting CII from falling into terrorists' hands. Seen in this light, however, the CIIA's FOIA exemption is quite peculiar: the statute requires DHS to withhold CII that, as I explain below,[80] it *already* has the legal authority under FOIA to withhold.

a. *The Peculiar FOIA Exemption*

Even without the CIIA, DHS already has the authority under a long-standing FOIA exemption to rebuff requests for *all* voluntarily submitted CII. On its face, Exemption 4 of FOIA (the trade secrets exemption) allows DHS to withhold information that is "privileged or confidential."[81] In its en banc decision in 1992 in *Critical Mass Energy Project v. Nuclear Regulatory Commission*, the D.C. Circuit explained that Exemption 4 "protects *any* financial or commercial information provided to the Government on a voluntary basis if it is of a kind that the provider would not customarily release to the public."[82]

---

[79] *See* JAMES Q. WILSON, BUREAUCRACY: WHAT GOVERNMENT AGENCIES DO AND WHY THEY DO IT 159–60 (1989) (describing difficulties in managing agencies in which outputs are hard to measure).

[80] *See infra* Part II.A.3.a.

[81] 5 U.S.C. § 552(b)(4) (2000).

[82] 975 F.2d 871, 880 (D.C. Cir. 1992) (en banc) (emphasis added). DHS should have wide latitude to reject requests for CII under Exemption 1 as well, which protects from disclosure information that has been classified "in the interest of national defense." 5 U.S.C. § 552(b)(1) (2000). Per the language of the standing executive order on secrecy, at least some CII will fit cleanly into the category of "defense information . . . which requires the highest degree of protection" and the disclosure of which "could result in exceptionally grave damage to the Nation." Because the President has delegated authority to DHS to classify such information as "Top Secret," Exec. Order No. 13,228, 66 Fed. Reg. 51,812 (Oct. 8, 2001), DHS will have substantial latitude to classify CII and refuse to divulge it. Despite the availability of judicial review for denials of FOIA requests, courts are likely to prove "reluctant to inspect classified documents in camera or to second-guess executive decisions in national security classifications." STEPHEN G. BREYER ET AL., ADMINISTRATIVE LAW AND REGULATORY POLICY: PROBLEMS, TEXT, AND CASES 875 (5th ed. 2002). Given DHS's substantial discretion to classify sensitive CII, it is hard to see exactly why it needs *another* FOIA exemption to keep *less* sensitive CII secret.

Because the CIIA's FOIA exemption covers only "voluntarily submitted" CII,[83] the CIIA protects *only* that information that could *already* have been withheld under the D.C. Circuit's interpretation of Exemption 4.[84] The CIIA therefore *tracks* a pre-existing FOIA exemption rather than *alters* it, and its effect on the free flow of information should therefore be negligible.

The CIIA's sole additional bite is that it imposes criminal penalties, including a possible one-year jail term, on any federal employee who "knowingly publishes, divulges, discloses, or makes known in any manner or to any extent not authorized by law, any critical infrastructure information protected from disclosure . . . ."[85] This transforms DHS's *discretionary* decision to withhold information under Exemption 4 into a legal mandate. It is hard to believe, however, that DHS would ever have exercised its discretion to disclose CII in light of its professed commitment to ensuring that firms are not deterred from disclosing CII because they are afraid the information will be vulnerable to FOIA requests. There is consequently substantial reason to believe that the CIIA's additional FOIA exemption will only have a limited effect on the public's access to CII.

b. The Costs and Benefits of Secrecy

Whether DHS relies on the CIIA or on Exemption 4 to veil our CII, however, a first-order question remains: what are the costs and benefits of this extensive secrecy in the terrorism context? Although the appropriateness of keeping CII secret has been framed as a tradeoff between security and openness, it is better understood as a tradeoff between security (i.e., hiding CII from terrorists) and security (i.e., rational priority setting guided by public accountability). The CIIA reflects a calculation that shrouding our risk-reduction decision-making in secrecy will prevent terrorists from exploiting our vulnerabilities while still allowing for rational decision-making. This Section argues that the government is likely to have calculated incorrectly. The level of secrecy contemplated by the current regime is likely to confer only slim benefits in deterring the terrorist exploitation of our critical infrastructures, while at the same time hampering the kind of public oversight that promotes reasoned decision-making. Although repealing the CIIA would not, without more, be sufficient to trigger mean-

---

[83] 6 U.S.C. § 133(a)(1) (Supp. II 2002).

[84] *See* Brett Stohs, *Protecting the Homeland by Exemption: Why the Critical Infrastructure Information Act of 2002 Will Degrade the Freedom of Information Act*, 2002 DUKE L. & TECH. REV. 0018, ¶ 13, *available at* http://www.law.duke.edu/journals/dltr/articles/2002dltr0018.html ("[Exemption 4] has precisely the same goal as does the CIIA."); Uhl, *supra* note 32, at 293 ("Cases suggest that Exemption 4 . . . already protects critical infrastructure information.").

[85] 6 U.S.C. § 133(f) (Supp. II 2002).

ingful public oversight, its emphasis on *additional* secrecy is nevertheless troubling.

i. Realism About Terrorism

There are two major reasons to be skeptical that keeping CII secret will have much of an effect on terrorist behavior. First, terrorists have repeatedly demonstrated that they do not need CII to carry out their attacks. The State Department's most recent update on the *Patterns of Global Terrorism* catalogs the terrorist attacks that took place in 2003. Of these, the vast majority were bombings in public, urban spaces, and required no information beyond a bombing manual and matériel.[86] Even better-financed and larger-scale attacks normally rely only on information that is widely available. The most critical piece of information used to coordinate the attacks of September 11 was an airline timetable, and the terrorists who planned the recent plot against the Citigroup skyscraper took the unsophisticated approach of "casing the joint" in person.[87] CII would be unnecessary to poison our food supply[88] or to crash a plane into a chemical plant.[89] The same goes for using shipping containers to transport weapons and hazardous materials.[90] This is not to say that CII will never play a role in terrorist planning—it probably will—but is rather to suggest that terrorists do not need CII to plan and successfully execute horrific attacks. We are so vulnerable in obvious ways that terrorists do not need CII to learn about our less-obvious weaknesses.

Second, even if terrorists did want to use CII to carry out their next attack, they could probably get it whether or not DHS kept that CII secret. To begin with, a startling amount of information is publicly available, particularly (although not exclusively) over the internet.[91] Fine-grained satellite images are available for free, making the general layout of any critical infrastructure facility public knowledge.[92] Detailed maps

---

[86] U.S. DEP'T OF STATE, PATTERNS OF GLOBAL TERRORISM 2003 5 (2004), *available at* http://www.state.gov/documents/organization/31912.pdf (reporting that bombings constituted 137 out of 190 terrorist attacks in 2003).

[87] Douglas Jehl & David Johnston, *Reports that Led to Terror Alert Were Years Old, Officials Say*, N.Y. TIMES, Aug. 3, 2004, at A1.

[88] Becker, *supra* note 8 ("'For the life of me, I cannot understand why the terrorists have not attacked our food supply because it is so easy to do.'" (quoting Tommy Thompson, Sec'y of Health and Human Services)).

[89] *See* Grimaldi & Gugliotta, *supra* note 5, at A1 (reporting allegations that Mohammed Atta, the leader of the September 11 plot, considered ramming a plane into chemical plants).

[90] Eric Lipton, *Audit Faults U.S. for its Spending on Port Defense*, N.Y. TIMES, Feb. 20, 2005, at 1.

[91] For one extensive website that is dedicated to posting sensitive information, see GlobalSecurity.org—Reliable Security Information, http://www.globalsecurity.org/ (last visited Oct. 21, 2005).

[92] *See, e.g.*, GoogleLocal, http://maps.google.com (providing satellite imagery of the majority of the country, including major metropolitan areas).

of New York City's electrical grid were posted on the internet for some time,[93] and primers on different ways of taking down electrical grids—some written by officers in the United States Armed Services—are readily available.[94] Site-owners have published "detailed maps of nuclear storage facilities in New Mexico" and diagrams of natural gas pipelines.[95] The Nuclear Regulatory Commission recently shut its online reading room because a national watchdog group notified the Commission that the reading room contained links to floor plans of nuclear research laboratories.[96] Doubtless other government agencies have posted information that terrorists could use to equal advantage.

The government's response to information on the internet has been almost laughable. Immediately after September 11, federal[97] and state[98] governments began removing sensitive information from their websites in the hopes that doing so would eliminate it from the public domain. Enterprising NGOs and website owners immediately re-posted cached versions of much of that information, and have continued to post information that the government would prefer was out of the public view.[99] What the government has been slow to grasp is that once information gets onto the internet, it is *out*.[100] There is no way to put that genie back into the bottle. Given the ease with which information can be posted to the internet and the rapidity with which it can then be disseminated worldwide, stemming the flow of all but the best-controlled information will prove an unworkable strategy to keep critical information out of terrorist hands.

---

[93] James C. McKinley, Jr., *State Restricts Data on Internet in Attempt To Thwart Terrorists*, N.Y. TIMES, Feb. 26, 2002, at B1.

[94] *See, e.g.*, MAJOR THOMAS E. GRIFFITH, JR., STRATEGIC ATTACK OF NATIONAL ELECTRICAL SYSTEMS (1994), *available at* http://www.comw.org/pda/fulltext/griffith.pdf.

[95] Urbina, *supra* note 10, at A29.

[96] *NRC Removes All Nuclear Information from Its Public Website*, OMB WATCH, Nov. 1, 2004, http://www.ombwatch.org/article/articleview/2498/1/297/.

[97] *Access to Government Information Post September 11th*, OMB WATCH, Feb. 1, 2002, http://www.ombwatch.org/article/articleview/213/1/1/ (last visited Oct. 21, 2005) (containing updated list of information removals from government websites).

[98] James C. McKinley, Jr., *State Restricts Data on Internet in Attempt to Thwart Terrorists*, N.Y. TIMES, Feb. 26, 2002, at B1.

[99] *See* Tom McNichol, *Peeking Behind the Curtain of Secrecy*, N.Y. TIMES, Nov. 13, 2003, at G7; McKinley, Jr., *supra* note 98, at B5 ("Some search engines save information from old Web sites, for instance, so a terrorist might still be able to find a map of New York's power grid.").

[100] In one particularly rich example, the Nuclear Regulatory Commission (NRC) threatened to prosecute staff members at a government-watchdog group, the Project on Government Oversight (POGO), for publishing a study on security gaps at the Indian Point nuclear facility. After a public scuff-up in which the NRC asked POGO to remove discussion of "some of those issues [that] would not be in the best interests of the United States," but declined to specify *which* issues were highly sensitive because that would itself supposedly imperil security, POGO revised some of the sections of its study. The original study was immediately posted on memoryhole.com. *See* R. Jeffrey Smith, *Nuclear Security Decisions Are Shrouded in Secrecy; Agency Withholds Unclassified Information*, WASH. POST, Mar. 29, 2004, at A1.

Even without the internet, terrorists could still quite easily procure needed CII. Most obviously, the owners and operators of our critical infrastructure employ hundreds of thousands of American workers. Surely a few of those employees share sympathies with transnational terrorists. One of the 1993 World Trade Center bombers, for example, worked as a chemical engineer at Allied Signal,[101] and an Al Qaeda training manual advocates that its operatives seek out "[i]ndividuals who are recruited either as volunteers or because of other motives" who can provide the group with needed information.[102] (Although it canvasses a number of ways of obtaining information about potential targets, the training manual nowhere mentions FOIA.) It would be far easier—and far more effective—for terrorists to turn to these human sources to learn about a facility's vulnerabilities rather than pore over information contained in reams of arid technical documents acquired from DHS.[103]

In short, classifying information to make it more difficult for terrorists to learn about our vulnerabilities does little to diminish their capacity to do so through other readily available means. This is not to say that we are helpless to mitigate our critical infrastructure vulnerabilities; by minimizing terrorists' capacity to exploit our highest-profile targets, we can certainly reduce our exposure to terrorist risks. And this is also not to say that there is not some CII which we would do well to restrict—it is hard to see what benefits would be had by making public highly technical information like the layouts of our nuclear facilities.[104] It *is* to say, however, that while restricting CII will confer some security benefits, those benefits are in general likely to be modest.

ii. The Costs of Secrecy

In contrast, the costs of draping regulatory decision-making in secrecy are acute. Both Exemption 4 and the CIIA grant an *informational* and an *access* monopoly to the private sector, making it vastly more difficult for the public to assess DHS decisions critically and to hold the agency accountable for the mistakes it is bound to make. Hornbook principles of

---

[101] LINDA-JO SCHIEROW, CRS REPORT FOR CONGRESS: CHEMICAL PLANT SECURITY 4 (2004).

[102] AL QAEDA TRAINING MANUAL, CHAP. 11, BM-82, *available at* http://www.usdoj.gov/ag/manualpart1_3.pdf (last visited on Oct. 20, 2005).

[103] For this reason, part of the critical infrastructure protection plan includes a certification program for background-screening companies so as to ensure that companies can get reliable information on their employees. NATIONAL STRATEGY, *supra* note 21, at 29. Nevertheless, *Time Magazine* recently reported that our nuclear power plants do not have sufficient mechanisms in place to "prevent a saboteur from engineering a catastrophe." Mark Thompson & Bruce Crumley, *Are These Towers Safe?; Why America's Nuclear Power Plants are Still so Vulnerable to Terrorist Attack—and How To Make Them Safer*, TIME, June 20, 2005, at 37.

[104] *NRC Removes All Nuclear Information from Its Public Website*, OMB WATCH, Nov. 1, 2004, http://www.ombwatch.org/article/articleview/2498/1/297/.

administrative law, however, hold that agencies make better decisions when they have as much information at their disposal as possible and when they have thoughtfully weighed the interests of their various constituencies.[105] Consequently, the APA and FOIA are normally seen as imposing a presumption in favor of disclosure.[106] Private citizens and watchdog groups then have an opportunity to make their voices heard, either by participating in notice and comment[107] or by demanding political action from their representatives.

That participatory role is hamstrung if DHS does not release the information upon which it bases its decisions. As Judge Leventhal, a founder of modern administrative law, wrote in an influential opinion, "It is not consonant with the purpose of a rulemaking proceeding to promulgate rules on the basis of inadequate data, or on data that [in] critical degree, is known only to the agency."[108] The CIIA, however, *ensures* that DHS will make its regulatory decisions based on information that it cannot disclose widely. This is troubling. As no less eminent an authority than Judge Patricia Wald has written concerning FOIA, "too much secrecy breeds irresponsibility"[109]—irresponsibility that, in the risk assessment context, will markedly diminish our security.

Not long before he died, Senator Daniel Patrick Moynihan made the same point in *Secrecy*, a book in which he provided an account of the United States government's growing obsession with its secrets over the course of the twentieth century. The Senator laid out a compelling case for seeing our modern penchant for secrecy—born in the crucible of the Cold War—as a bureaucratic pathology motivated by a bureaucratic instinct for self-preservation.[110] Secrecy thus became a governmental reflex action and an end in itself. In Moynihan's view, this reflex is pernicious not simply because it is undemocratic and authoritarian, but for the more straightforward reason that secrecy clouds judgment, as exhibited most strikingly in our intelligence services' gross mischaracterization of the stability of the Soviet Union in the seventies and eighties.[111] Moynihan's (rather radical) conclusion is that the costs of secrecy *in virtually all circumstances* are far greater than its benefits.[112] Although it runs counter to

---

[105] *See, e.g.*, Michael C. Dorf & Charles F. Sabel, *A Constitution of Democratic Experimentalism*, 98 COLUM. L. REV. 267, 373 (1998) ("[I]t is a staple of democracy that in most contexts publicity encourages change for the better.").

[106] PETER L. STRAUSS ET AL., GELLHORN & BYSE'S ADMINISTRATIVE LAW: CASES & COMMENTARIES 734 (10th ed. 2003).

[107] Government Organization and Employees Act, 5 U.S.C. § 553(c) (2000).

[108] Portland Cement Ass'n v. Ruckelshaus, 486 F.2d 375, 393 (D.C. Cir. 1973).

[109] Patricia M. Wald, *The Freedom of Information Act: A Short Case Study in the Perils and Paybacks of Legislating Democratic Values*, 33 EMORY L.J. 649, 654 (1984).

[110] DANIEL PATRICK MOYNIHAN, SECRECY 142–43 (1998).

[111] *See id.* at 194–99.

[112] In this, Moynihan agrees with the essential conclusion of a task force established in 1970 by the Defense Department: that "more might be gained than lost" if the United States were to adopt "a policy of *complete openness in all areas of information*." *Id.* at 175

the instincts of a country that has yet to emerge fully from the mind-set of the Cold War, his takeaway conclusion resonates with even greater force with the security threat of transnational terrorism: "*Analysis*, far more than secrecy, is the key to security."[113]

The informational and access monopolies also ensure that industries will dominate DHS decision-making—or, to put it another way, that industries will capture the agency.[114] As James Q. Wilson notes in his classic text on bureaucracy, capture is virtually inevitable when an agency finds itself "confronting an environment where much of the information it needs and many of the political resources to which it must respond will be in the hands of an interest fundamentally hostile to its purposes."[115] Behind-the-scenes decision-making will thus amplify "the danger that the redistributive authority of agencies will be exercised in favor of a limited group of organized interests with a special stake in an agency's policies."[116]

This is not to say that restricting the flow of CII would be without its benefits: "[I]n truth, the FOIA, like all basic freedoms, sometimes hurts the worthy and sometimes helps the unworthy."[117] But the looming terrorist threat should not distract us from the principle that open government is not simply a good in itself (although it is that too), but is rather a basic component of thoughtful, reasoned decision-making. But it is curious that the CIIA can explicitly recognize the need for thoughtful analysis of the threat to our critical infrastructures and yet palpably fail to provide a space in which the contours of that threat can be fully assessed.

B. Risk Management: The Limits of Command and Control

Even if the CIIA were an effective risk assessment tool, DHS would be left with the unenviable task of managing those risks. As explained earlier, a blind reliance on the market will be inadequate because firms will predictably under-invest in security.[118] DHS's job will therefore be to goad firms into taking adequate precautions to ensure that terrorist attacks will not expose the population to inordinate risks. In principle, this could be done through command-and-control regulations. DHS could assess the risks posed by certain socially desirable behavior, weigh the costs of risk reduction against the benefits of such reduction, and require firms to meet

---

(citing Defense Science Board, *Final Report of the Defense Science Board Task Force on Secrecy* (July 1, 1970)) (emphasis added).

[113] MOYNIHAN, *supra* note 110, at 222.

[114] BREYER ET AL., *supra* note 82, at 879 ("FOIA can be viewed as a mechanism for correcting agency 'failure' by providing broader public knowledge and scrutiny of administrative practices and providing another court-enforced procedural mechanism for citizen involvement in government.").

[115] WILSON, *supra* note 79, at 78.

[116] MOYNIHAN, *supra* note 110, at 147.

[117] Wald, *supra* note 109, at 683.

[118] *See supra* Part II.A.1.a.

regulatory requirements that reduce risks in such a way as to optimally balance costs against benefits.[119]

In practice, however, command-and-control regulations—the bread and butter of regulatory policymaking—are unlikely to be well-suited to reducing terrorism risks.[120] As Justice Breyer has noted, this kind of regulatory standard setting relies for its success on a large amount of "accurate, relevant" information.[121] At a minimum, setting non-arbitrary regulatory standards in this context would require information on different *types* of terrorism risks to which particular facilities are vulnerable (i.e., airplane attack, sabotage), the *likelihood* that terrorists will mount a particular attack (taking into account that terrorists will adapt their strategies to changing risk profiles), the *scope* of the damage associated with a successful type of attack (including cascade effects), the identification of different security *strategies*, the relative risk *reductions* associated with the implementation of security strategies, and the relative *costs* of those security strategies.[122]

Simply reciting this list gives some idea of the magnitude of the informational problems associated with terrorism. Even the Office of Information and Regulatory Affairs (OIRA), an agency with the primary task of assessing the costs and benefits of agency regulations, has acknowledged that it is struggling to adapt conventional regulatory approaches to reductions in terrorism risks. John Graham, OIRA's current administrator and an expert in risk assessment techniques, explained in a 2003 speech:

---

[119] *See* Revesz, *supra* note 72, at 7–23 (discussing cost-benefit analysis in the environmental arena).

[120] This Article leaves to one side serious discussions of alternative approaches—particularly liability schemes and the creation of risk-reduction markets—that are not plausible strategies to mitigate terrorism threats.

Ex post approaches, such as the imposition of liability for damages resulting from a successful attack, are likely to fail because rational firms will invest in additional security only up to the level of its solvency, and not beyond. *See* Revesz, *supra* note 72, at 8–17. That solvency will in every case be far exceeded by the "debilitating costs" associated with an attack on our critical infrastructure. Consequently, ex post liability will not transmit adequate incentives to shore up security to a level that will adequately protect the public welfare. Even if firms had the proper incentives, the enormous uncertainty surrounding terrorism risks means that firms would rarely (if ever) have sufficient information to make rational choices about how to reduce their vulnerabilities to such attacks.

Tradeable permit schemes are often advanced as a way to tap into market forces to help alleviate information-gathering burdens on regulators. *See* Bruce A. Ackerman & Richard B. Stewart, *Reforming Environmental Law*, 37 STAN. L. REV. 1333 (1985). Although one could theoretically devise a market for terrorism risk in which firms swapped allowances for risk, the practical and ethical difficulties with such an approach would be insurmountable. *Cf.* Jeffrey Rosen, *This Year in Ideas: Total Information Awareness*, N.Y. TIMES MAG. § 6, Dec. 15, 2002 (detailing one such abortive effort to create a market in terrorism risks).

[121] STEPHEN G. BREYER, REGULATION AND ITS REFORM 103 (1982).

[122] For a thoughtful discussion of the multiplicity of factors that go into risk assessments, see Alon Rosenthal et al., *Legislating Acceptable Cancer Risk from Exposure to Toxic Chemicals*, 19 ECOLOGY L.Q. 269 (1992).

> [H]omeland security has emerged as a new growth area for federal regulation. . . . At [OIRA] we have been humbled by the challenge of analyzing these ideas. How should agencies quantify the benefits of rules aimed at reducing the probability of (or damages from) future terrorist acts? How should agencies quantify the costs of homeland security rules, whether they come in the form of time losses at airports or intrusions into privacy or freedoms of foreign students and visitors to our country[?] Quite frankly, the agencies and [OIRA] need help on how homeland security ideas should be evaluated.[123]

As Graham's comments highlight, any assessment of terrorism risks will be dogged by unknowns on a scale that outstrips other areas of regulation.[124] And although the government has taken fumbling steps to inquire into some of these areas,[125] a dearth of information about the country's vulnerabilities remains. Particularly because terrorists are human agents who can adapt to changes in our risk profile, the information gap promises to remain intractable.

Environmental regulation, another arena in which regulators lack accurate and relevant information, provides an illuminating analogy. As Richard Stewart explains, the most important "inheren[t] limitation" of command-and-control regulation is "the inability of central planners to gather and process the information needed to write directives appropriately responsive to the diverse and changing conditions of different economic actors."[126] The result is that "regulation writers face grave difficulties in gathering information about the diverse circumstances of different facilities and devising requirements that are responsive to these different circumstances. In order to economize on decision-making costs, regulators adopt uniform measures of procrustean character that are often inappropriate for particular facilities," thereby creating enormous waste and impeding sensible priority-setting.[127] Bradley Karkkainen, another astute observer of the informational deficits that plague environmental regulation, agrees:

---

[123] John D. Graham, Administrator, Office of Information and Regulatory Affairs of the Office of Management and Budget, CATO Institute Hill Briefing: Reigning in the Regulatory State: The Smart-Regulation Agenda (Oct. 3, 2003), http://www.whitehouse.gov/omb/inforeg/speeches/031003graham.html.

[124] *See generally* JOHN D. GRAHAM ET AL., IN SEARCH OF SAFETY: CHEMICALS AND CANCER RISK (1998) (discussing difficulties of quantifying the degree of increased risk of cancer from exposure to carcinogenic compounds).

[125] To identify possible terrorist threats, DHS has pulled together focus groups from the public and private sectors to "think like terrorists" and brainstorm about potential vulnerabilities. *See, e.g.*, John Donnelly, *Fighting Terror the Scientific Approach: Panel of Scientists to Tackle Terrorism*, BOSTON GLOBE, Oct. 13, 2001, at A1.

[126] Richard B. Stewart, *United States Environmental Regulation: A Failing Paradigm*, 15 J.L. & COM. 585, 587 (1996).

[127] *Id.* at 588.

> Conventional approaches to environmental regulation are nearing a dead end, limited by the capacity of regulators to acquire the information necessary to set regulatory standards and keep pace with rapid changes in knowledge, technology, and environmental conditions. A pervasive *information bottleneck* constrains the extent, effectiveness, efficiency, and responsiveness of the regulatory system.[128]

Given the dynamic nature of terrorist threats and the vast difficulties associated with predicting the consequences of security gaps, this information bottleneck will be at least as profound in the terrorism context—and the constraints on the "extent, effectiveness, efficiency, and responsiveness" of command-and-control regulations will be similarly severe.[129]

The daunting information problems associated with a command-and-control strategy call into serious question such a strategy's utility in combating critical infrastructure vulnerabilities. Similar to command-and-control regulations in the environmental context, precious security dollars are likely to be spent adhering to misguided, misapplied, or minimally effective regulations rather than on appropriately securing the country's critical infrastructure against terrorist threats.[130] The costs associated with developing the information needed to impose even those inefficient regulations will deter DHS and other agencies from rulemaking and will likely provide large disincentives to revisiting the substantive standards that do make it through the regulatory gauntlet. Regulatory flexibility risks being replaced by regulatory ossification, which will in turn substantially impede efforts to respond to terrorists that can adapt nimbly to changes in our risk profile.[131]

In sum, while command-and-control regulations might make us marginally safer, the costs of adhering to "procrustean" regulations would likely outweigh the benefits. As a consequence, alternative approaches merit careful consideration.

---

[128] Bradley C. Karkkainen, *Information as Environmental Regulation: TRI and Performance Benchmarking, Precursor to a New Paradigm?*, 89 GEO. L.J. 257, 263 (2001).

[129] *See, e.g.*, Viscusi & Zeckhauser, *supra* note 45, at 101 ("Even the insurance industry, which is thoroughly acquainted with estimating unusual risks, has a hard time gauging the risk of terrorism losses.").

[130] *See* Karkkainen, *supra* note 128, at 269 (noting that agencies that suffer from this information bottleneck must inevitably turn to "crude categorical" requirements that "introduce familiar rigidities and inefficiencies into the regulatory scheme."); *see also* Stewart, *supra* note 126, at 588.

[131] *See* Thomas O. McGarity, *Some Thoughts on "Deossifying" the Rulemaking Process*, 41 DUKE L.J. 1385 (1992).

III. BENCHMARKING

Because the CIIA fails to provide the government with adequate information upon which to premise an effective regulatory response, and because command-and-control regulations are unlikely to be effective tools for managing critical infrastructure risks, the government should consider the implementation of a benchmarking strategy. Benchmarking is an information-based alternative that shows promise as a cost-effective and democratically accountable mechanism to assess and manage our critical infrastructure risks.

This Part argues that, if firms were required to provide to DHS standardized information on their security profiles, gleaned from on-the-ground holistic inspections, and if that information were published in an easily comparable and easily accessible format, informed public oversight could create substantial incentives for firms to improve their security profiles. Moreover, because the information would be formatted to facilitate oversight rather than facilitate terrorism, benchmarked data would avoid the pernicious consequences of the FOIA exemption while at the same time minimizing the risk that terrorists could exploit our sensitive information.

This strategy, by spurring cost-effective security reductions even in the face of uncertainty, may prove superior to clumsy command-and-control regulations that rely on large amounts of information for their success. Benchmarking will have some predictable weaknesses: some firms may not have adequate incentives to invest in security measures and will lag behind others; benchmarking focuses on a narrow portion of the problem, limiting its utility as a risk-reduction device; and relying on public oversight mechanisms gives rise to fears of public irrationality. But benchmarking—whether used alone or in combination with command-and-control approaches—nevertheless holds great promise as a strategy to more adequately protect the country's critical infrastructure.

A. Implementing a Benchmarking Approach

Recognizing the serious difficulties associated with conventional regulation in the face of profound uncertainty, commentators in recent years have begun to pay more attention to informational approaches in the regulatory arena.[132] Of particular interest has been a practice, drawn from the

---

[132] *See, e.g.*, MARY GRAHAM, DEMOCRACY BY DISCLOSURE (2002); Coglianese et al., *supra* note 49; Dorf & Sabel, *supra* note 105, at 345–88 (describing success of TRI, Forest Service initiatives, and informational approaches to the regulation of nuclear power plants); Karkkainen, *supra* note 128, at 286–94 (detailing success of EPA's TRI); Cass R. Sunstein, *Informational Regulation and Informational Standing:* Akins *and Beyond*, 147 U. PA. L. REV. 613 (1998) [hereinafter Sunstein, *Informational Regulation*]; *see also* WESLEY A. MAGAT & W. KIP VISCUSI, INFORMATIONAL APPROACHES TO REGULATION (1992).

business world's emphasis on "total quality management,"[133] known as benchmarking. Put simply, benchmarking is a process whereby firms monitor their performance across a particular variable (e.g., profitability, efficiency, employee relations) and compare their performance with that of other firms in order to figure out the most effective techniques.[134] The advantages of a benchmarking approach are manifold:

> This benchmarking comparison of actual with potential performance disrupts established expectations of what is feasible. By casting pragmatic doubt on the advisability of current methods, benchmarking spurs exploration of the possibilities immediately disclosed and may lead to discovery of entirely new solutions through investigation of the surprising similarities and differences among various approaches.[135]

Drawing on this emerging literature, this Article argues that a benchmarking approach holds substantial promise as a regulatory strategy to mitigate critical-infrastructure risks. The argument in favor of benchmarking advances in three parts, each of which details an important component of a successful benchmarking approach. First, because firms have limited incentives to develop and share CII, the government should *mandate* that firms participate in CIPP. Second, in order to facilitate making comparisons across firms, the government should *standardize* the information it receives by generating simple metrics along which vulnerabilities can be measured and requiring on-site risk inspections that rate firms along those metrics. Third, in order to facilitate political and market oversight, the government should *publicize* the information in a readily comprehensible format on a publicly accessible database.

To help demonstrate the desirability and feasibility of these various features, this Section draws analogies to EPA's Toxics Release Inventory (TRI), an "innovative, if barebones, national system of benchmarking [the]

---

[133] For another application of total quality management (TQM) in the regulatory arena, see E. Donald Elliott, *TQM-ing OMB: or Why Regulatory Review Under Executive Order 12,291 Works Poorly and What President Clinton Should Do About It*, 57 LAW & CONTEMP. PROBS. 167 (1994).

[134] Karkkainen cites to a definition of benchmarking as "the continuous process of measuring products, services and processes against those of industry leaders or the toughest competitors, resulting in a search for best practices, those that will lead to superior performance, through measuring performance, continuously implementing change, and emulating the best." Karkkainen, *supra* note 128, at 260 n.8 (citing JOHN S. OAKLAND, TOTAL QUALITY MANAGEMENT: THE ROUTE TO IMPROVING PERFORMANCE 181 (2d ed. 1993)). Michael Dorf and Charles Sabel describe benchmarking as the practice of undertaking "an exacting survey of current or promising products and processes [to] identif[y] those products and processes superior to those the company presently uses, yet are within its capacity to emulate and eventually surpass." Dorf & Sabel, *supra* note 105, at 287.

[135] Dorf & Sabel, *supra* note 105, at 286–87.

self-reported releases of toxic substances,"[136] and argues that TRI can be seen as a rough template for a novel approach to terrorism risks. The structure of the TRI program is simplicity itself. EPA lists 581 chemicals and thirty chemical "categories" as hazardous substances,[137] and requires that firms disclose, in standardized units,[138] any discharges or transfers of those chemicals from any facility.[139] EPA then collects that standardized data and publishes it on several readily available databases to facilitate public oversight.[140] Although TRI does not mandate that firms reduce their discharges, the program has nevertheless spurred dramatic decreases in the discharge of hazardous substances.[141] This Section suggests that a benchmarking program structured along the lines of TRI but designed particularly to address critical infrastructure risks could spur similarly dramatic improvements in the country's domestic security.

1. Mandating Disclosure

Firms cannot benchmark in a vacuum. They must have information about themselves and about their competition in order to make relevant comparisons. Although the government has blithely assumed that firms will invest substantial resources in investigating their vulnerabilities to terrorist attacks and then sharing that information with the government, firms normally lack adequate incentives to generate or to share information about their vulnerabilities.[142]

So why not mandate that owners and operators of our critical infrastructure provide CII to DHS? As Cass Sunstein argues, "compulsory disclosure of information can provide the simplest response" to the failure of the market to provide adequate information.[143] And there is certainly a pressing public need to secure CII from owners of our critical infrastructure, many of whom make a profit in part by externalizing some of their security risks onto the public. Requiring them to run their own risk assessments and then forcing them to provide that CII to the government may very well be the only way for the federal government to get an accurate sense of the country's true threat profile.

---

[136] *Id.* at 375.

[137] EPA website, TRI Program, http://www.epa.gov/tri/chemical/index.htm (last visited Oct. 12, 2005).

[138] 42 U.S.C. § 11023(g)(2) (2000).

[139] *Id.* § 11023(a).

[140] *Id.* § 11044; *see also* EPA website, What Is TRI?, http://www.epa.gov/tri/whatis.htm (last visited Oct. 12, 2005).

[141] Karkkainen, *supra* note 128, at 287–88.

[142] *See supra* notes 47–57 and accompanying text; *cf.* Karkkainen, *supra* note 128, at 283 ("Private parties will generally lack adequate incentives to produce and disclose much of the information relevant to solving environmental problems.").

[143] Sunstein, *Informational Regulation*, *supra* note 132, at 624.

Proponents of the CIIA counter with three separate arguments. First, they claim that government coercion will not produce high-quality CII.[144] Firms have incentives to minimize their apparent vulnerabilities in order to protect themselves from costly regulatory responses; any coerced risk assessments will thus understate the extent of the "real" threat.[145] In contrast, firms will be fully forthcoming about their security risks (or so the argument runs) only when they have reason to believe that the information will not be used against them. According to CIIA proponents, the CII garnered under a voluntary approach will therefore be more accurate than CII secured through a coercive approach.[146]

The implicit assumptions in this account are unsupportable. As explained above, most firms will be deterred from providing CII to DHS voluntarily.[147] The proper comparison is thus not between the quality of the CII provided under mandatory and voluntary programs, but between coerced CII and little or no CII at all. Moreover, voluntarily provided CII is unlikely to be of high quality. The few firms that *do* provide CII voluntarily will still be sensibly conscious of the risk of provoking a regulatory response and will have continuing incentives to obscure the full extent of their vulnerabilities. Voluntary CII, because it can be provided selectively and can exclude information about a firm's most acute risks, is quite likely to be even *more* misleading than a coerced, and therefore comprehensive, assessment of a firm's risk profile.[148]

This is not to deny that receiving low-quality CII is a problem, but rather to point out that it is a problem that a voluntary program does not solve. Furthermore, securing high-quality CII is not an insurmountable obstacle. Many of our regulatory strategies rely on creating incentives for the private sector to engage in meaningful self-audits, and several obvious strategies gleaned from other regulatory programs could be successfully employed. The Resource Conservation and Recovery Act (RCRA), for example, requires generators of hazardous wastes to utilize certified transport, storage and disposal (TSD) facilities that the Act subjects to extensive regulation.[149] Similarly, the government could require owners and operators of our critical infrastructure to hire security consultants that have been certified as meeting certain professional standards to undertake their risk assessments. Or, like the IRS, DHS could audit a certain percentage of firms' risk assessments to ensure their accuracy and impose stiff civil or criminal penalties if the firms have understated their vulnerabilities. Better still, the government could lift a page from the Sarbanes-Oxley Act, which requires executives to certify on pain of criminal prosecu-

---

[144] NATIONAL STRATEGY, *supra* note 21, at 8.
[145] *See supra* Part I.
[146] NATIONAL STRATEGY, *supra* note 21, at 12–13.
[147] *See supra* Part II.A.1.
[148] *See supra* Part II.A.2.a.
[149] Resource Conservation and Recovery Act, 42 U.S.C. §§ 6922, 6924 (2000).

tion that the information they are submitting is true, fair, and accurate.[150] Jail terms have a remarkable capacity to focus the executive mind on providing the government with what it wants. The point is not that any of these alternatives are good ideas, but rather that the problem of low-quality self-assessments is amenable to solution.

The second objection is that a coercive approach will deter firms from forming cooperative relationships with the federal government. As the *National Strategy* avers, "stimulating voluntary, rapidly adaptive protection activities requires a culture of trust and ongoing collaboration among relevant public- and private-sector stakeholders." Consequently, a voluntary program is preferable to "more traditional systems of command and control."[151] The owners of our critical infrastructure (the argument goes) might otherwise not share vital life-saving information in the event of an emergency, and might not be willing to work with DHS to develop best practices on how to address the terrorist threat.

This objection is similarly wrongheaded. A mandatory CII program that creates poisonous interactions between private industry and the federal government would *still* be vastly preferable to a voluntary program that creates no relationships at all. Furthermore, there is little reason to believe that a requirement to undertake risk assessments will generate friction between industry and DHS. Quite the opposite, in fact. Private industries, now forced to work with DHS, would prefer to form productive relationships so that they are well-positioned to influence its decision-making. This explains why private industries often work amicably with EPA, even as EPA implements costly environmental statutes.[152] This is not to say that owners of critical infrastructure will "like" DHS any more than industrial plants "like" EPA. But if they are forced to work together, they will find ways to do so productively.

The third objection to a more coercive regulatory regime is cost. The government may be loath to force private industries to undertake expensive risk assessments, especially because "[m]any of the critical infrastructure sectors are currently highly regulated."[153] This objection is quite curious. In quantifying the costs that the voluntary CIPP would impose on the private sector, DHS has explained patiently that most private firms have already committed the resources to assessing their security vulnerabilities and that turning CII over to the government would therefore impose only negligible costs.[154] Indeed, in quantifying the additional "costs"

---

[150] 18 U.S.C. § 1350 (2000); *see also* WILLIAM T. ALLEN & REINIER KRAAKMAN, COMMENTARIES AND CASES ON THE LAW OF BUSINESS ORGANIZATION 275–76 (2003) (detailing the criminal penalties associated with violations of a host of federal statutes).

[151] NATIONAL STRATEGY, *supra* note 21, at 8.

[152] *See* Cary Coglianese, *Litigating Within Relationships: Disputes and Disturbance in the Regulatory Process*, 30 LAW & SOC'Y. REV. 735, 750–51 (1996) (describing positive working relationship between regulated industries and EPA).

[153] NATIONAL STRATEGY, *supra* note 21, at 12.

[154] Initial Regulatory Flexibility Determination, 69 Fed. Reg. 8081 (Feb. 20, 2004)

of participating in CIPP, DHS listed only the price of labels and stamps used to designate the CII it submits to DHS ($9.90–$10.25 for stamps, $7.87–$34.92 for labels).[155] If firms have already undertaken risk assessments, then mandating their disclosure will also not impose any additional costs on the private sector. If, on the other hand, firms have not engaged in risk assessments, then a voluntary CIPP could not possibly succeed. Why on earth would a firm incur significant costs to participate in a voluntary program? Either the private sector already has sufficient incentives to undertake its own risk assessments or it does not. The government cannot have it both ways.

The bottom line is that, because there is a pressing public need for this information, a regulation requiring owners and operators of our critical infrastructure to provide CII is sensible and practical. The invocation of appealing phrases like "cooperation and partnership" and "market solutions" should not obscure that private industry does not have the proper incentives to provide this information voluntarily. Moreover, throwing minor sops to the private sector—like the FOIA exemption—will do nothing to promote information exchange and partnerships.

2. *Standardizing Critical Infrastructure Information*

Benchmarking requires more than raw information, however. That information must also be standardized and keyed to quantifiable metrics so as to allow for comparisons across facilities. Without such a metric, firms will have no way to assess their facilities' performances against their competition, reducing the incentives that a benchmarking program might otherwise create. Nor will firms have a meaningful way to assess their own performance, set goals for improvements, or monitor progress toward those goals.

Much of the success of EPA's TRI program flows from its reliance on a simple, standardized metric—namely, the number of pounds of release of a particular chemical by a particular plant.[156] By drawing attention to this standardized number, "firms and facilities are compelled to self-monitor and, therefore, to confront disagreeable realities concerning their environmental performance in detail and early on."[157] Simply by virtue of having something to measure, TRI has thus drawn firms' attention to a problem that they had previously ignored,[158] "jarring firms into action" and allowing "managers [to] set firm-wide improvement targets

("DHS believes that affected entities will incur minimal *costs* from complying.").

[155] *Id.* at 8081, 8082.

[156] Karkkainen, *supra* note 128, at 289–90.

[157] *Id.* at 295 (internal quotes removed).

[158] *See id.* at 297 ("Many top corporate managers, previously unaware of the volumes of toxic pollutants their firms were generating, were indeed surprised by the information produced in the first rounds of TRI.").

and gauge progress toward their achievement."[159] Although TRI only offers a very rough proxy of environmental performance, EPA's impulse to standardize one measure of industry's polluting activities reflects a regulatory effort to reap the benefits of benchmarking.

Because "security" cannot be readily measured against a commonly recognized yardstick, however, providing a similar standardized metric in the terrorism context is deeply problematic. Rigid checklists of required security practices will not take into account a high degree of variation among firms, nor are they likely to provide a meaningful measure of facilities' actual security vulnerabilities. Similar to the way in which OSHA's "going by the book" strategy of regulatory enforcement has been the subject of longstanding criticism for the inefficiencies that result from its focus on technical violations rather than on the substantive goal of worker safety, a checklist approach threatens to submerge our substantive objective of reducing critical infrastructure vulnerabilities under technical requirements that may not substantially advance that goal.[160]

On the other hand, undertaking a more holistic assessment of a firm's security practices would be difficult, costly, and potentially subjective. Distilling facility-wide assessments into a set of standardized quantifiable metrics could strip the information of qualitative nuance. Regulatory agencies that "grade" firms on their security profiles may also provoke claims of politically driven unfairness, which will in turn impose an enormous burden on regulatory agencies to demonstrate even-handedness. Many agencies will attempt to prove even-handedness by retreating to rigid objective standards,[161] an outcome that would impede efforts to reduce critical infrastructure vulnerabilities.

These substantial objections would carry more weight, however, were it not for the fact that at least one regulatory program designed *precisely* to measure safety and allow for intra-industry benchmarking is already in place and functioning quite effectively. In the wake of the near-meltdown at a nuclear plant on Three Mile Island and in an effort to "restore its image,"[162] the nuclear power industry established in 1979 "a new system of benchmarking regulation of the nuclear power industry housed in the Institute of Nuclear Power Operations (INPO) . . . to reduce the potential

---

[159] *Id.* at 297.

[160] *See generally* EUGENE BARDACH & ROBERT A. KAGAN, GOING BY THE BOOK (1982) (describing inefficacy of OSHA's rigid enforcement strategy); STEVEN KELMAN, REGULATING AMERICA, REGULATING SWEDEN (1982) (comparing OSHA's highly detailed procedures and punitive approach to compliance to its loosely structured, collaborative Swedish counterpart).

[161] WILSON, *supra* note 79, at 113–36 (describing how institutional constraints will "induce rational managers to base their decisions on the most defensible criteria," which will normally be "more objective, quantifiable, and visible," to the detriment of the substantive goal).

[162] Thomas W. Lippman, *Is Industry Usurping NRC Functions?: Seabrook Controversy Sheds Light on Secretive Nuclear Institute*, WASH. POST, Mar. 26, 1990, at A9.

for catastrophic incidents in the industry."[163] INPO's core function is information-oriented: it collects information on effective safety strategies, designs benchmarks that it can use to assess plant-wide safety, and then evaluates and ranks plants according to those benchmarks.[164] Despite the fact that measuring nuclear "safety" is as elusive as measuring facility "security," INPO has been hailed as "one of the most successful schemes of [industry] self-regulation ever documented."[165]

In order to evaluate nuclear power plant safety, INPO trains small teams of experts at INPO headquarters in a "results-oriented" inspection methodology. That methodology does not prescribe methods for achieving safety objectives, but rather "anticipat[es] that [alternative] methods and procedures in use may be as effective as those described in good practices."[166] This approach consequently "emphasizes achievement of the performance objectives" while "strongly discourag[ing] a rule-bound and compliance-oriented approach."[167]

INPO then dispatches teams to "spen[d] two weeks of twelve-hour days doing nothing but watching what is going on at [nuclear] plant[s]."[168] Notably, "the inspection process involves two key tasks: *observing* operational activities at the plant and *interpreting* their significance."[169] The INPO inspectors put facts they uncover into context, and their objective evaluations of nuclear plant safety are based on a holistic picture of actual safety on the ground. INPO then compiles and analyzes the results of the inspectors' reports and ranks nuclear power plants against each other with reference to a set of "industry-wide performance indicators."[170] Those rankings are disclosed to CEOs and boards of directors of nuclear energy utilities, as well as to the Nuclear Regulatory Commission (NRC).[171]

Leaving aside the factors that motivate nuclear power plants to improve their safety rankings, the key to INPO's rather startling success has been its capacity to give firms a limited set of quantitative benchmarks that tell an uncompromising and yet informed story about a firm's safety profile. Firms, importantly, can rely on those benchmarks to set internal

---

[163] JOSEPH V. REES, HOSTAGES OF EACH OTHER: THE TRANSFORMATION OF NUCLEAR SAFETY SINCE THREE MILE ISLAND 1 (1994).

[164] Dorf & Sabel, *supra* note 105, at 371 ("In practice, INPO's chief activities consist of pooling the industry's operating experience, establishing benchmarks that distill the lessons it contains, and then evaluating individual power plants according to their ability to meet the relevant benchmarks.").

[165] Neil Gunningham & Darren Sinclair, *Integrative Regulation: A Principle-Based Approach to Environmental Policy*, 24 LAW & SOC. INQUIRY 853, 869 (1999); *see also* REES, *supra* note 163, at 4 ("Safety-related performance indicators . . . all show clear signs of improvement.").

[166] REES, *supra* note 163, at 76.

[167] *Id.* at 76. For a list of assessment guidelines, see *id.* at 70–87, tbls. 5.1–5.7.

[168] Dorf & Sabel, *supra* note 105, at 372 (internal quotes omitted).

[169] REES, *supra* note 163, at 141–42 (emphasis in original).

[170] *Id.* at 98.

[171] Dorf & Sabel, *supra* note 100, at 372.

improvement goals and measure progress toward meeting those goals without meeting rigid requirements that may be ill-suited for their particular facility. Moreover, INPO's ranking system facilitates regulatory oversight, as demonstrated by the NRC's reliance on the rankings to prod lower-ranked firms to implement more robust safety measures.[172]

In principle, a similar kind of intensive on-the-ground inspection strategy could be employed to assess facility-wide security against potential terrorist threats. In fact, it is already happening. OMB reports that "[c]ompanies representing more than 90 percent" of the chemical industry "have adopted a comprehensive security code *that includes mandatory inspections*" by independent third parties.[173] Implemented under the auspices of the chemical industries' pre-existing Responsible Care program,[174] the security code requires independent inspectors to make "security vulnerability assessments" (SVAs) of every participating facility.[175] While the results of those inspections remain secret, leaving many open questions as to the quality or efficacy of the inspection regime, the chemical industry's embrace of SVAs suggests that regulated industries recognize that independent inspections play a central role in minimizing critical infrastructure risks. Similarly, Congress has required drinking water systems that serve more than 3300 people to conduct vulnerability assessments and submit them to EPA,[176] and EPA has provided extensive and thoughtful documentation to help guide these self-inspections.[177] Although self-inspections will be kept secret[178] and may reflect self-interested bias, the requirement that facilities assess their security readiness reflects a legislative recognition that inspections are critical to shoring up the country's critical infrastructure vulnerabilities.

In order to spur the same kind of inspection schemes in other industries, DHS could form industry-specific centralized offices to develop expertise in the kinds of terrorist threats most likely to plague that industry. These already exist in nascent form within DHS as industry working groups called Information Sharing and Analysis Centers (ISACs), which are intended to "allow critical sectors to share information and work together

---

[172] *See id.* at 345–88; Gunningham & Sinclair, *supra* note 157, at 869.

[173] OFFICE OF MGMT. & BUDGET, EXEC. OFFICE OF THE PRESIDENT, 2003 REPORT TO CONGRESS ON COMBATING TERRORISM 35 (2003) (emphasis added).

[174] For a discussion of Responsible Care, see generally Neil Gunningham, *Environmental Management Systems and Community Participation: Rethinking Chemical Industry Regulation*, 16 UCLA J. ENVTL. L. & POL'Y 319 (1997).

[175] American Chemistry Council, Fact Sheet: Responsible Care Security Code (Apr. 2005), http://www.americanchemistry.com/s_acc/bin.asp?SID=1&DID=1232&CID=258&VID=109&DOC=File.PDF.

[176] Public Health Security and Bioterrorism and Response Act of 2002, Pub. L. No. 107-188, §§ 401–403, 116 Stat. 682 (2002) (codified at 42 U.S.C.A. § 300i-2 (West 2003)).

[177] *See* EPA website, Vulnerability Assessments, http://cfpub.epa.gov/safewater/watersecurity/home.cfm?program_id=11 (last visited Sept. 29, 2005).

[178] 42 U.S.C.A. § 300i-2(a)(5) (2005).

to help better protect the economy."[179] The ISACs could be given a measure of independence and the resources, first, to develop standardized industry-specific benchmarks to gauge the security readiness of facilities within that industry, and second, to train and deploy teams of security experts to evaluate facilities according to those benchmarks. DHS could even tailor the intrusiveness (and therefore the cost) of these inspections across different industries, depending on the relative difficulty of assessing different industries' security vulnerabilities and a rough assessment of the relative risks that particular industries pose to homeland security. As in the INPO setting, facilities could then be "graded" along these benchmarks according to their relative security.

Much turns on the quality of these inspections. But because there is little reason to believe security readiness is more difficult to assess than nuclear safety—both involve holistic assessments of facility-wide readiness that are contingent on a number of difficult-to-quantify and hard-to-observe variables—INPO's "unqualified success"[180] stands as a powerful counterexample to those who argue that rigorous security evaluations are simply too contingent and subjective to have any value.

To be sure, such an inspection regime (whether financed by taxpayers or by firms within an industry) would be costly.[181] But it would be far less costly than a regulatory alternative that relied on spotty and inadequate information to prescribe one-size-fits-all solutions that failed to protect the nation's most vulnerable critical infrastructure. It is also likely to be less costly in the long run than doing nothing to protect ourselves because we lack enough information to form an appropriate response. Insisting that the country assess its vulnerabilities on the cheap—as the CIIA would apparently have it—is simply not a strategy for effective infrastructure protection.

An inspection regime faces one profound obstacle that INPO does not. INPO (like Responsible Care) is an industry-sponsored organization that exists to assuage public fears of nuclear power and therefore serve the economic interests of its members. DHS's efforts to uncover facilities' security gaps are unlikely to have a similarly broad base of industry support. Although some firms within an industry may be pleased that the inspections show them to be more secure than their competition, the industry as a whole would probably prefer to keep quiet the degree to which it has externalized terrorism risks onto the American public.[182] As public choice pathologies emerge, industry hostility to DHS's evaluation efforts

---

[179] DHS, Sharing Information to Protect the Economy, http://www.dhs.gov/dhspublic/display?theme=73&content=1375 (last visited Nov. 20, 2005).

[180] Dorf & Sabel, *supra* note 105, at 372.

[181] INPO's annual budget in 1989 was reported to be $51.8 million before 1990, and was financed by contributions from the fifty-five electric utilities that operate nuclear power plants. Lippman, *supra* note 162, at A9.

[182] *Cf.* Coglianese et al., *supra* note 49, at 291.

will assuredly translate into significant political pressure.[183] Claims of favoritism and political cronyism will proliferate (justifiably or not) against the agency, and any inspection regime will operate under a constant political shadow. Faced with substantial opposition, DHS will be sorely tempted to retreat to a less subjective, but less effective, inspection regime that focused on technical requirements rather than meaningful evaluations of facility-wide security.[184]

Again, however, this is not an insuperable obstacle. Many regulatory agencies operate quite effectively under similarly long political shadows, which is hardly surprising given that agencies developed in part as a mechanism to resist day-to-day political pressures.[185] In any event, terrorism risks are salient enough that public attention can provide at least a partial counterweight to public choice pressures that might otherwise threaten to cripple an inspection regime. And on a more chilling note, even if public choice pathologies were to sap much of the efficacy of an inspection regime, a single devastating terrorist attack could immediately realign political forces to support extraordinary regulatory efforts to uncover security gaps before they could be exploited.

3. *Publishing Benchmarks*

Even with a fistful of benchmarked CII in hand, however, firms must have the proper incentives to use benchmarking as an affirmative tool for risk management. With respect to changes that will inure to the firm's private benefit, such as improvements in manufacturing processes or management practices, market pressures will predictably push firms to engage in an appropriate level of benchmarking activity. With respect to changes that will inure to the public good, however, a puzzle remains: "Although firms have flexibility to choose their own improvement targets, why should they bother to do so at all?"[186]

Benchmarking's solution is a paradoxical one: Perhaps all DHS needs to do is make its benchmarked information public. In so doing, it will provide a ready-made source to promote meaningful public oversight—which may *by itself* be enough to promote some measure of risk management.[187] Although the view that a toothless regulatory regime could be

---

[183] *See generally* George J. Stigler, *The Theory of Economic Regulation*, 3 BELL J. ECON. & MGMT. SCI. 3, 11 (1971).

[184] *See* WILSON, *supra* note 79, at 122, 127 (reporting on agencies' need to demonstrate fairness in the face of political pressure).

[185] Woodrow Wilson, *The Study of Administration* (1887), *reprinted in* CLASSICS OF PUBLIC ADMINISTRATION, at 10 (Jay M. Shafritz & Albert C. Hyde eds., 1978) ("[A]dministration lies outside the proper sphere of politics."); Richard B. Stewart, *The Reformation of American Administrative Law*, 88 HARV. L. REV. 1669, 1671–80 (1975) (detailing the traditional model of administrative law).

[186] Karkkainen, *supra* note 128, at 295.

[187] *See* Sunstein, *Informational Regulation*, *supra* note 132, at 626 ("A primary virtue

effective may sound like hopeless naïveté, several programs have demonstrated that the approach can be remarkably successful in practice.[188] EPA's TRI program—which has been lauded as "one of the nation's most effective environmental laws" by at least two EPA administrators,[189] despite mandating no reductions in hazardous waste disposal[190]—is the most important and best studied of these.

a. TRI

In his careful study of TRI's success as a benchmarking program, Karkkainen discusses a number of factors that may have contributed to firms' reductions of their pollution discharge in spite of the lack of enforceable regulatory targets.[191] He contends that the mandatory production and disclosure of TRI information forces firms to develop information about their own firms' pollution practices that, prior to TRI, many firms did not compile.[192] Managers then can use that information "to evaluate [their] own performance and production processes" against those of other firms and work toward improvement goals.[193] TRI also provides a tool for industry-organized groups to bring pressure to bear on laggard firms whose substandard environmental performance threatens to undermine the environmental credentials of the rest of the industry.[194]

More importantly, TRI sends a powerful message that the government is watching. "Adverse facility-, firm-, or industry-level TRI data . . . carry the implicit threat that regulatory action may follow, whether at the initiative of regulators themselves or in response to rising political demand for regulatory action."[195] That threat of regulatory action means that industries will have strong incentives to self-police in order to preempt governmental regulation.

This threat of regulation is all the more acute because TRI, by providing the necessary informational predicate about one aspect of environmental performance, also fosters community oversight. Community groups can use TRI information to reward or punish firms based on their discharge records. Punishment may involve pickets, boycotts, lawsuits, or

---

of informational regulation is that it triggers political safeguards and allows citizens a continuing oversight role—one that is, in the best cases, largely self-enforcing.").

[188] For a lengthy discussion of informational strategies, see *id.* at 618–24.

[189] Karkkainen, *supra* note 128, at 287 (discussing views of Carol Browner and William Reilly, former EPA administrators).

[190] Dorf & Sabel, *supra* note 105, at 375–76 (TRI "neither fixes targets for the reduction of aggregate levels of pollution, nor requires specific pollution-abatement efforts by particular classes of polluters.").

[191] Karkkainen, *supra* note 128, at 294–331.

[192] *Id.* at 297–98.

[193] *Id.* at 295–96.

[194] *See id.* at 309.

[195] *Id.* at 311.

(worst of all) efforts to force national and local governments to regulate.[196] By driving up the costs of business to offending firms, community organizations can turn even a toothless TRI program into a highly effective tool to bring political pressure to bear to achieve social goals.

Market forces should also serve to make firms see positive TRI ranking as part and parcel of a profitable business strategy, not as an onerous government regulation with which to grudgingly comply. Several studies have concluded that reports of low TRI scores can bring about substantial decreases in firms' share prices and raise firms' cost of capital[197]—which then, in turn, may cause firms to reduce their emissions.[198] The translation of adverse TRI rankings into financial pain should focus managerial attention on guaranteeing that their hazardous waste discharges do not provoke a negative response in the capital markets. Adverse TRI ranking may also make it more difficult to recruit and maintain a high-quality labor force, either because workers fear hazardous waste or because they prefer to work in firms with better environmental records.[199] Some environmentally conscious consumers may shy away from goods from firms with poor TRI records; similarly, poor TRI rankings may impose severe reputational costs.[200]

In the final estimation, Karkkainen concludes that TRI's success is "overdetermined":

> The underlying genius of TRI, then, is that by measuring and continuously tracking facility- and firm-level environmental performance by using objective and comparable metrics, it creates a transparent and information-rich environment, enabling monitoring and benchmarking by multiple actors—by managers and directors, as well as by markets, communities, and the regulatory apparatus at all levels of government.[201]

---

[196] *Id.* at 316 ("By driving up the cost of doing business, these measures may force the polluting firm to negotiate over de facto (even if not legally mandatory) environmental standards.").

[197] *See* James T. Hamilton, *Pollution as News: Media and Stock Market Reactions to the Toxics Release Inventory Data*, 28 J. ENVTL. ECON. & MGMT. 98, at 108–11 (1995); Karkkainen, *supra* note 128, at 260 n.7 (listing studies). Changes in share value, of course, only reflect investors' beliefs about the likelihood that TRI reports will influence a firm's expected value. Stock fluctuations therefore depend on community and political responses to TRI information for their effect—price signals "are messengers, not the message." Karkkainen, *supra* note 128, at 324.

[198] Shameek Konar & Mark A. Cohen, *Information as Regulation: The Effect of Community Right To Know Laws on Toxic Emissions*, 32 J. ENVTL. ECON. & MGMT. 109, 120 (1997) (concluding that firms that suffer the largest reductions in share price then undertake the largest reductions in emissions).

[199] Karkkainen, *supra* note 128, at 325–26.

[200] *Id.* at 327–28.

[201] *Id.* at 329.

Although TRI imposes no enforceable regulations, these multiple sources of public oversight give the program substantial bite.

b. Prompting Vulnerability Reductions

TRI's success at inducing private actors to make voluntary emissions reductions should be replicable in the context of reducing terrorist risks. Consider that the core reason TRI's oversight mechanisms have proven so effective is that communities (local, statewide, and national) now have the information they need to identify and punish those firms that thwart the public's preference for a cleaner environment.[202] Informational regulation thus acts as a kind of democratic enabler by helping to transform undirected public will into directed regulatory strategy.

It is no stretch to say a post–September 11 public has at least as strong a desire to minimize vulnerabilities to our critical infrastructure as it does to reduce toxic effluent discharges. A well-crafted benchmarking regime that allowed for vulnerability comparisons would consequently focus considerable public scrutiny on low-performing firms with demonstrated security gaps. The managers of our critical infrastructure would have an assessment of their vulnerabilities on hand and a benchmark against which to measure progress; firms would have incentives to reduce their vulnerabilities in an effort to preempt governmental regulation; community groups and NGOs could engage in grassroots organizing to promote further vulnerability reductions; and market discipline could impress upon firms the importance of taking the public's preferences seriously. Market forces might even operate with particular force in the terrorism context because the liabilities associated with a successful attack are so much larger than the liabilities associated with hazardous waste discharge. Labor costs at lower-ranked firms could increase sharply as managers and workers decide that going to work at certain facilities was simply too risky. And the reputational costs associated with being seen as a facilitator of terrorism would be profound. This confluence of oversight mechanisms would be sure to prompt at least some measure of vulnerability reduction across a wide range of sectors.

Better still, firms would have *continuous* incentives to improve facility safety so as to keep pace with their competition and not run afoul of public antipathy.[203] As Michael Dorf and Charles Sabel argue, a nimble benchmarking regime promotes the adoption of "rolling best-practices," by which they mean best practices that shift based on what industry leaders have demonstrated is possible. The implementation of rolling best practices is particularly appealing in the terrorism context. As they explain:

---

[202] *Id.*

[203] *Id.* at 276.

> [R]olling best-practice can be used potentially, to reduce sources of risk in novel or experimental products, even before the precise nature of those sources can be identified. Potential rolling best-practice rules are useful where product life cycles are short with respect to the time needed fully to test and improve the safety of a product under real-world conditions (computers, much software, and complex financial products) or where initial real-world failures would be catastrophic (pharmaceuticals, foodstuffs, and products bound for space or the battlefield). The way to reduce risks under these circumstances is to characterize more and more precisely the sources from which hazards may derive and to reduce and monitor each precisely characterized source more and more effectively.[204]

Although the analogy is not precise, critical infrastructure risks correspond to those product risks in which "real-world failures would be catastrophic," and in which the implementation of rolling best practices over time may be the only effective way to minimize the possibility of a disastrous results. And, like with products for which safety assessments are difficult to make quickly, there will be no point at which we have reached "the time needed *fully* to test and improve the safety" of our critical infrastructure facilities. Just as terrorists will adapt to changes in our security profile, we in turn must adapt our security profile in the face of the changing contours of the terrorist threat. Consequently, benchmarking's capacity to promote innovation and punish stultification will prove highly useful in the terrorism context.

Informational deficiencies have stymied public oversight of our critical infrastructure vulnerabilities and have allowed firms to continue externalizing their security risks onto a largely unsuspecting population. A benchmarking regime could provide the public with the information it needs to hold these firms accountable and ensure that its security preferences were not subsumed by industry's desire for secrecy and deregulation.

c. Formatting for Oversight vs. Formatting for Terrorism

Although a blanket FOIA exemption for CII is troubling in the absence of an alternative oversight mechanism,[205] it is not the case that honoring FOIA requests for CII would be the *ideal* way to allow the public to scrutinize DHS's efforts. In the terrorism context, in fact, there is reason to think that it may not prove particularly efficacious. In the same way that DHS would face overwhelming difficulties in compiling and analyzing CII, members of the public (with presumably fewer resources than

---

[204] Dorf & Sabel, *supra* note 105, at 353.

[205] *See supra* Part II.A.3.

the government) would be flummoxed when trying to use a limited subset of that raw CII to draw conclusions about our homeland-security gaps.[206] A mass of raw CII produced in response to a FOIA request—while better than nothing—would consequently be of only limited utility as an oversight mechanism.

If benchmarks were published and easily accessible, however, then FOIA would no longer be the only mechanism through which the public could get information about security vulnerabilities. Published benchmarks could instead provide the information necessary to affirm the public's central role in regulatory decision-making. Because these benchmarks would be formatted specifically to facilitate public oversight, moreover, they would be superior monitoring devices to cumbersome FOIA requests. Benchmarks are, after all, the end products of a costly analytical process in which a motley assortment of raw security information is converted into a quantitative snapshot of a facility's security profile. Effective oversight requires only that these quantitative assessments are made available; public overseers can then use those assessments and rankings to shine a spotlight on those firms that have provided a sub-optimally low degree of security.

If benchmarking can meet public oversight needs, then FOIA's costs weigh more heavily.[207] Exposing sensitive CII may make sense if that exposure is necessary to effectuate public oversight; if benchmarking can better involve the public in regulatory decision-making, however, then that same exposure may be unnecessary and counterproductive. Although the risk that terrorists will exploit CII in planning and executing their attacks is modest, even that modest risk counsels against disclosure if the public has alternative ways to hold DHS and critical infrastructure owners accountable.

The publication of benchmarks, in contrast to FOIA requests, gives rise to fewer security concerns. Benchmarks would, to be sure, identify those facilities that are relatively more vulnerable to particular kinds of attacks than other comparable facilities, and could help direct terrorists' attention to our most vulnerable installations. The degree to which benchmarks would operate as aids to terrorists could be easily overstated, however. Benchmarks offer only limited information, after all; they are merely rankings that can be formatted so as to allow for the optimal targeting of public opinion while withholding a complete set of CII. At best, then, they would help terrorists answer the question as to *which* facilities to target, but it would not provide sufficient information to show them *how* to go about attacking them. Terrorists would still need to turn to alternative sources of information to find a "roadmap" to crippling our critical infrastructure. Moreover, private firms that perceive that their low rankings will put their facilities at greater risk of exploitation will have strong in-

---

[206] *See supra* Part II.A.2.

[207] For a discussion of those costs, see *supra* Part II.A.3.b.

centives to shore up their security gaps immediately. Benchmarking may therefore prompt a kind of dynamism in the security profiles of firms within an industry, making it easy for the public to know which firms under-invested in security *yesterday* (and therefore are likely to under-invest in the future) while making it risky for terrorists to rely on benchmarks as reliable indicators of actual security gaps *today*.

What is more, those who argue that benchmarking will provide terrorists with a roadmap to exploiting our critical infrastructure must first explain how published security rankings raise more profound risks than readily available accounts in newspapers, magazines, and books.[208] Most startlingly, the *New York Times* recently published a detailed account of an internal report that DHS had compiled on the "dozen possible strikes it views as most plausible or devastating," complete with estimates of casualties and economic consequences. While the report remains confidential, the *Times* listed the potential casualty counts for several types of attacks as well as detailing the three "most devastating of the possible attacks."[209] Other examples are legion. In November 2003, *60 Minutes* dispatched camera crews to *identified* chemical plants across the United States and "found gates unlocked or wide open, dilapidated fences, and unprotected tanks filled with deadly chemicals that are used to manufacture everything from plastics to fertilizer."[210] *New York Times* reporter Matthew Brzezinski recently published *Fortress America*, a book that reads like a litany of worst-case disaster scenarios, in which he identifies (among other risks) highly vulnerable targets in the Baltimore port.[211] And many articles have detailed the security deficiencies of nuclear power plants generally[212] and Indian Point in particular.[213]

The key insight here is that information on our critical infrastructure vulnerabilities can and should be formatted to promote public oversight while minimizing its utility for the purposes of terrorism. Although any approach in which the public maintains some oversight role over our critical infrastructure vulnerabilities raises some risk that terrorists will use that information against us, a benchmarking approach does far better than FOIA at striking a sound balance between these competing objectives.

---

[208] *Cf.* New York Times Co. v. United States, 403 U.S. 713 (1971) (allowing the publication of the Pentagon Papers over government's objections).

[209] Lipton, *supra* note 11, at A1.

[210] CBS website, U.S. Plants: Open to Terrorists, June 13, 2004, http://www.cbsnews.com/stories/2003/11/13/60minutes/main583528.shtml.

[211] BRZEZINSKI, *supra* note 11, at 3–7.

[212] *See, e.g.*, *Nuclear Security Training Lacking; Plants Eliminated or Reduced Drills Designed To Repel Attacks, U.S. Says*, WASH. POST, Mar. 18, 2004, at A20; Editorial, *Our Unnecessary Insecurity*, N.Y. TIMES, Feb. 20, 2005, at 8, A29.

[213] *See, e.g.*, Kennedy, *supra* note 65, at 14WC.

B. Some Practical Objections

As with any regulatory regime, benchmarking will have some predictable shortcomings. The following Section discusses three such deficiencies, arguing that they can be mitigated and, in any event, do not justify the abandonment of benchmarking as a regulatory strategy.

1. Laggards

One difficulty with a benchmarking program is that different firms will respond differently to public oversight pressures.[214] Predictably, then, a benchmarking regime—like any voluntary regulatory program—will over time produce leaders (firms with relatively more secure facilities) and laggards (firms with relatively less secure facilities).[215] The consequences in the terrorism context may be particularly extreme. As leader firms continually improve their security profiles, laggard firms will become the low-hanging fruit of our critical infrastructure, ripe for terrorists to pluck. Benchmarking will thus only divert terrorists from relatively "hard targets" (leader firms) to more vulnerable "soft targets" (laggard firms).

Without minimizing the force of this argument, however, a regulatory regime that produced *some* leaders is preferable to a system in which *all* (or most) firms are laggards. Even if some firms obdurately refused to invest in greater security protection in the face of public and market pressure, terrorists will not inevitably concentrate their attention on less-secure facilities. Many factors go into choosing a target for a terrorist strike, including some that are wholly unrelated to the relative chances that the attack will succeed. It may even be the case that terrorists will prefer in some cases to strike at *better*-defended targets, either because of the importance of those targets to our national identity or because such an attack could demonstrate their ability to strike us at will.[216] Because terrorism risks will not simply be redistributed in light of the implementation of a benchmarking regime, an improvement in the security profiles of a substantial fraction of our critical infrastructure facilities should reduce our net risk exposure.

The laggard problem is moreover amenable to some degree of legislative correction. If it grows acute enough to attract the attention of the gov-

---

[214] *Cf.* Karkkainen, *supra* note 128, at 338–43 (discussing factors that may influence a firm's responsiveness to a benchmarking regime, including the demographics of surrounding communities, its size, whether the firm is publicly traded, the degree to which it relies on advertising or has a negative public image, and its management and governance structures).

[215] *Cf.* Karkkainen, *supra* note 128, at 338 (noting "heterogeneity" of response to TRI); Neil Gunningham & Joseph Rees, *Industry Self-Regulation*, 19 L. & POL'Y 363, 393 (1997) (discussing laggard problem in nuclear safety regulation).

[216] NATIONAL STRATEGY, *supra* note 21, at 72–74.

ernment, a legislative response (whether at the municipal, state, or federal level) could require all firms to meet a baseline level of security. Although the imposition of a mandatory baseline to supplement a voluntary benchmarking regime might frustrate laggards' achievement of some security goals at least cost and may impose some undesirable regulations on leader firms, such a regime might nevertheless capture most of the advantages of a benchmarking approach while mitigating the extent of the laggard problem.

Alternatively, the government might empower DHS to impose fines and even close critical infrastructure facilities that refused to take adequate security measures. INPO's experience on this score is illuminating. The voluntary organization struggled early on with an acute laggard problem that "jeopardized" its authority over the leader firms.[217] INPO's solution was to cooperate with the NRC, which it asked to sanction laggard firms and threaten to shut down their nuclear plants. By flexing the coercive power of the state only *after* voluntary efforts had failed to provoke an adequate response, INPO was able to improve nuclear plant safety at relatively low cost while minimizing the effect of the laggard problem. As two commentators colorfully note, only "the presence of the regulatory gorilla in the closet" assured INPO's success as a voluntary program.[218]

This is, again, not to downplay the laggard problem. Any benchmarking regime must be sensitive to the fact that laggards are sizeable and well-publicized chinks in our defensive armor. The laggard problem is not an insurmountable obstacle, however, and does not by itself justify the abandonment of a benchmarking regime that, on the whole, would reduce our vulnerability to terrorist strikes.

2. *Inadequate Benchmarking?*

One trenchant criticism is that benchmarks of facility-wide security *gaps* are bad proxies for true security *risks*.[219] Benchmarking, in bringing attention to one narrow component of a comprehensive risk assessment (the security profile of a firm), may ignore other important components (including the relative amount of collateral damage of attacks at different facilities and the likelihood that terrorists have the resources or the will to undertake particular kinds of attacks) to such a degree that it will prove next to useless at targeting public oversight to those industries that are truly most vulnerable.[220] On this view, the costs of collecting benchmarked information will outweigh whatever marginal benefit it might otherwise confer.

---

[217] Gunningham & Sinclair, *supra* note 165, at 868–69.

[218] *Id.* at 869.

[219] *Cf.* Karkkainen, *supra* note 128, at 331–35 (arguing that TRI information is a "flawed proxy for environmental performance").

[220] *See supra* note 134 and accompanying text.

The benefits of benchmarking might similarly be thought to be minimal because benchmarking does not allow for comparisons *among* industries. As explained earlier, an effective inspection regime would be likely to limit itself to comparing security vulnerabilities at facilities within a particular industry.[221] Vulnerability rankings will therefore not provide ready answers as to whether facilities in one industry are relatively more secure than facilities in other industries, making it difficult for the public and the market to focus their attention on those facilities that pose the greatest risks.

Although these arguments have some merit, they are overdrawn and are not substantial enough to justify scuttling a benchmarking approach. As an initial matter, risk management should not be continually deferred out of some quixotic devotion to achieving an impossibly comprehensive assessment of our risk profile. We are unlikely ever to have perfect information—or even particularly good information—about how much damage might flow from an attack on part of our critical infrastructure or about the likelihood that terrorists will stage a particular attack. The use of rough and ready proxies will therefore be a necessary component of terrorism regulation.

What is more, a nuanced approach that considers other components of a risk assessment is not inimical to a benchmarking regime. Damage assessments (whether undertaken by governmental agencies or by private researchers) could be overlaid on top of rankings with respect to security gaps. Facilities could be divided into different risk categories—say, "high risk," "moderate risk," and "low risk"—in the same way that EPA categorized the various costs of attacks at different chemical plants based on their location,[222] or that DHS has ranked the "dozen possible strikes it views as most plausible or devastating."[223] That information could then be folded into the public database to allow the public to compare "high risk" facilities against each other and to focus public ire on the appropriate firms. DHS could overlay similar information about the likelihood that terrorists will employ certain tactics. Analogous efforts could be made with respect to industries that pose the most acute risks to homeland security.

Recall, moreover, that firms will respond to benchmarked vulnerability assessments as refracted through the prism of public opinion.[224] Even in the absence of any quantitative information beyond the benchmarks, public watchdog groups and private citizens will be unlikely to mistake benchmarked information for a comprehensive analysis of the risks to our critical infrastructure. They will therefore have to make judgments about

---

[221] *See supra* Part III.A.2.

[222] *See* Grimaldi & Gugliotta, *supra* note 5, at A01 (describing classification of chemical plants into categories based on whether they would expose more than one million, 100,000, or 10,000 people to substantial risks, respectively).

[223] *See* Lipton, *supra* note 11, at A1.

[224] *See supra* Part III.A.3.

which facilities (and, more broadly, industries) they believe warrant additional security. To be sure, as is explored below, those judgments may themselves be skewed and may reflect salient public fears rather than careful considerations of risk. But those judgments may also reflect common sense intuitions about those targets that terrorists will find most appealing. Clashing intuitions between different groups will work themselves out into a more-or-less democratic consensus about the security gaps with which we should be most concerned. Moreover, to the degree that those intuitions are not seen to be helpful guides, the need to carefully develop more information about terrorism risks that can then be folded into the vulnerability database will become apparent.

3. *Lay Perceptions of Terrorism Risk*

In contrast to the regulatory state's typical emphasis on a neutral balancing of the costs and benefits of regulations,[225] benchmarking embraces the mechanism of public opinion to foster performance improvements. As explained earlier, this turn to public perceptions is grounded in a recognition that intractable informational bottlenecks make cost-benefit calculations infeasible in some circumstances.[226]

The downside of this approach, however, is that public perceptions of risks have consistently been shown to diverge, sometimes wildly, from expert opinion.[227] The divergence between expert and lay perceptions of risk can in part be explained with reference to the public's reliance on heuristic devices—rules of thumb for processing information—when making judgments about probability.[228] The most important of these rules of thumb "for the purposes of understanding risk-related law" is the availability heuristic,[229] through which public perceptions of the likelihood of an event will turn on how easy it is to recall a similar event.[230] Naturally, terrorism risks are highly salient and the availability heuristic will predictably cause the public to overstate those risks.[231]

Other contextual factors will also influence lay perceptions of risk. As Richard Pildes and Cass Sunstein explain, those factors include:

---

[225] *See, e.g.*, Exec. Order No. 12,866 § 1(a), 3 C.F.R. §§ 638–639 (1994) (standing executive order governing OMB review of agency rulemaking directing that "[executive] agencies should assess all costs and benefits of available regulatory alternatives").

[226] *See supra* Part II.B.

[227] *See, e.g.*, BREYER, *supra* note 82, at 33.

[228] Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 1–20 (Daniel Kahneman & Amos Tversky eds., 1982).

[229] Cass R. Sunstein, *Probability Neglect: Emotions, Worst Cases, and Law*, 112 YALE L.J. 61, 64 (2002) [hereinafter Sunstein, *Probability Neglect*].

[230] Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1477 (1998).

[231] Cass R. Sunstein, *Terrorism and Probability Neglect*, 26 J. RISK & UNCERTAINTY 121, 121 (2003) [hereinafter Sunstein, *Terrorism*].

> (1) The catastrophic nature of the risk; (2) whether the risk is uncontrollable; (3) whether the risk involves irretrievable or permanent losses; (4) the social conditions under which a particular risk is generated and managed, a point that connects to issues of consent, voluntariness, and democratic control; (5) how equitably distributed the danger is or how concentrated on identifiable, innocent, or traditionally disadvantaged victims, which ties to both notions of community and moral ideals; (6) how well understood the risk process in question is, a point that bears on the psychological disturbance produced by different risks; (7) whether the risk would be faced by future generations; and (8) how familiar the risk is.[232]

Virtually all of these contextual considerations are present in the terrorism context; as Sunstein notes in a recent article on terrorism, these concerns will stoke public fears of terrorism "even if the magnitude of the risk does not justify those changes, and even if statistically equivalent risks occasion little or no concern."[233]

Sunstein identifies a third phenomenon that he calls "probability neglect" that will further serve to skew lay perceptions of terrorism risks. As he describes it, people fall victim to probability neglect "when intense emotions are engaged [and] people tend to focus on the adverse outcome, not on its likelihood. That is, they are not closely attuned to the probability that harm will occur."[234] As a consequence, the public may concentrate on reducing risks beyond the point at which the costs of risk reduction outweigh the benefits.

The confluence of these factors—heuristics, contextual features, and probability neglect—means that the public will predictably "over-react" to terrorism risks, at least as those risks are calculated by experts. Because benchmarking relies on public reaction for its effect, however, equipping laypeople with quantitative assessments demonstrating a host of vulnerabilities could amplify public irrationality and encourage a demand for "legal interventions that might not reduce risks and that might in fact make things worse."[235] The costs of the resulting misallocation of risk-reduction resources would be acute, effectuating what one author has provocatively called "statistical murder" on a grand scale.[236] Perhaps Dorf and Sabel are correct in noting that, when public over-reaction to benchmarks is assured, "confidentiality [may] bree[d] correctives."[237] If so, then

---

[232] Richard H. Pildes & Cass R. Sunstein, *Reinventing the Regulatory State*, 62 U. CHI. L. REV. 1, 57 (1995).

[233] Sunstein, *Terrorism*, *supra* note 231, at 122.

[234] Sunstein, *Probability Neglect*, *supra* note 229, at 62.

[235] Sunstein, *Terrorism*, *supra* note 231, at 122.

[236] John D. Graham, *Legislative Approaches to Achieving More Protection Against Risk at Less Cost*, 1997 U. CHI. LEGAL F. 13, 28.

[237] Dorf & Sabel, *supra* note 105, at 373.

the government could save lives by obscuring our full exposure to terrorism risks lest such exposure provoke a disproportionate public response.

There is no easy answer to this critique. Benchmarking does provide a mechanism for the public to shape regulatory policy—indeed, that is its purpose—and a reliance on inflated lay perceptions of risk is understandably anathema to those who call for the tethering of regulatory policymaking to cost-benefit analysis. But there are at least three reasons to think that we should not be too quick to overstate the degree to which irrationality will characterize a benchmarking regime.

First, the contextual factors that animate many lay perceptions of risk may reflect actual disagreement with the implicit value judgments underlying cost-benefit valuations—what one author has termed the public's "rival rationality."[238] To be sure, the availability heuristic and other inaccurate rules-of-thumb reflect cognitive errors that the regulatory apparatus should make every effort to correct.[239] But relying on the contextual factors that influence lay perceptions of risk is not irrational if that reliance reflects value judgments about which risks we as a society are more or less willing to tolerate.[240] As Pildes and Sunstein argue, "If people do value risks differently depending on these sorts of contextual features, and if these valuations are reasonable, then democratic policy should recognize the relevant contextual differences."[241]

On this view, it would not be irrational for the public to believe that a disproportionate fraction of our risk-reduction dollars should be devoted to terrorism than to, for example, automobile safety. Terrorism risks are catastrophic, uncontrollable, permanently damaging, involuntarily assumed, psychologically distressing, and unfamiliar.[242] The risks of car accidents, in contrast, are well-understood, familiar, voluntarily assumed (at least to some degree), and—however devastating—not catastrophic, in the sense of harming hundreds or thousands of people at once. Differences in risk perception may of course reflect a jumbled mix of cognitive errors and value judgments, and sorting out which is which may prove difficult. But if democracy entails a channeling of value judgments of the polity, then the government should perhaps honor those judgments instead of squelching them as "irrational."

---

[238] PAUL SLOVIC, THE PERCEPTION OF RISK 220–31 (2000). *Contra* BREYER, *supra* note 121, at 35 ("The public's 'nonexpert' reactions reflect not different values but different understandings about the underlying risk-related facts.").

[239] Pildes & Sunstein, *supra* note 232, at 60–61.

[240] *Id.* at 61 (noting cases "in which experts and laypeople *value* differently the same 'objective' risk (understood in terms, say, of aggregate lives at stake) as a result of features of the context that expert decision-theoretic or cost-benefit techniques obscure"). *But see* Sunstein, *Probability Neglect*, *supra* note 229, at 84 (suggesting that "when ordinary people disagree with experts, it is often not because of competing value judgments, but instead because ordinary people are more subject to probability neglect").

[241] Pildes & Sunstein, *supra* note 232, at 58.

[242] *Id.* at 57.

Second, the existence of probability neglect should call attention to the pervasive and acute fears surrounding terrorism risks. That fear, irrational or not, is *itself* a social evil and a legitimate target of governmental regulation.[243] Both Eric Posner and Cass Sunstein have championed the idea that the government has an interest in dampening public fear of terrorist attacks, both because fear is a bad in itself and because that fear may cause "ripple effects" that actually serve to increase risk.[244] That the extent of fear associated with terrorism can be enormous is unquestionable; for example, studies have documented widespread post-traumatic stress disorder in the wake of September 11.[245] There are, however, difficulties with respect to evaluating the cost of fear in a cost-benefit calculation, and agencies consequently "almost never enumerate and price the distressing mental states such as fear, anxiety, worry, panic, or dread."[246] To the extent that benchmarking can channel public feeling to address those fears, benchmarking might prove superior to cost-benefit analysis at achieving an optimal level of fear reduction.

The third response is pragmatic. Given the existence of severe informational bottlenecks and the limited incentives that private firms have to independently reduce their security vulnerabilities, a benchmarking regime may bring us closer than other alternatives to satisfying a cost-benefit test. The current situation, in which very little has been done to minimize grave critical infrastructure risks, almost certainly does not pass cost-benefit muster. Waiting for federal and state governments to allocate the funds needed to secure our critical infrastructure seems impractical if not downright foolish.[247] And, as discussed earlier, command-and-control regulations that are calibrated to highly contingent cost-benefit analyses that are themselves premised on spotty information will be unlikely to provide much risk-reduction bang for the security buck.[248]

Benchmarking is an alternative that, whatever its deficiencies, is at the very least likely to spur some dramatic and yet cost-effective security improvements. If security improvements are appropriate in a post–September 11 world, the fact that benchmarking should provide incentives for firms to improve their security profiles—and continue improving them

---

[243] *See* Posner, *supra* note 3, at 687 ("[T]he experience of fear is a hedonic loss.").

[244] Sunstein, *Terrorism*, *supra* note 231, at 132 ("[F]ear is a real social cost and . . . is likely to lead to other social costs."); Posner, *supra* note 3, at 687–88; *contra* ROSEN, *supra* note 2, at 5–6 (arguing that regulations that are "more concerned about feeling safe than being safe" will only serve to generate irrational "feel good" regulatory responses that are unlikely to provide additional security and yet quite likely to substantially undercut our civil liberties).

[245] ROSEN, *supra* note 2, at 2.

[246] *See generally* Matthew D. Adler, *Fear Assessment: Cost-Benefit Analysis and the Pricing of Fear and Anxiety*, 79 CHI.-KENT L. REV. 977 (2004) (advancing methodologies for measuring fear).

[247] *See id.* at 1005–11.

[248] *See id.* at 1011–24.

into the future—means that it is likely to come closer to the social optimum than clumsy alternatives.

That the public's response to benchmarked information will be infected by a measure of irrationality is of some concern, but industries can attempt to meet that irrationality on the field of open debate—a field in which, because of public-choice pathologies, they will have a tremendous advantage over a dispersed and hard-to-organize public.[249] Certainly our first amendment tradition does not admit of the argument that the possibility of irrational public response justifies withholding information, and indeed the argument is antithetical to a democratic system in which citizens are the wellspring of the government's authority. We have long since abandoned the fiction that our regulatory machinery can run on neutral technocratic expertise, and it seems at least arguable that public participation in making these decisions could prove beneficial inasmuch as it would help improve the accountability and responsiveness of bureaucratic decision-making.[250] In light of the paucity of suggestions as to how to more effectively reduce the vulnerability of our critical infrastructure to terrorist attacks, we would do well not to dismiss a regulatory approach that involved the citizenry in the affirmative promotion of their interests and promised a substantial measure of security improvements.

CONCLUSION

Our nation's vast, interconnected, and mutually dependent critical infrastructure presents a temptation to terrorists and a headache for regulators. Yet no coherent strategy to minimize terrorism-related risks has yet to emerge from the government. Its primary effort to date, the CIIA, is inadequate because it will not promote public-private cooperation; ineffective because it does not provide a mechanism for the thoughtful compilation or analysis of CII; and misguided because its secrecy provisions will ensure regulatory capture and hobble public oversight.

This Article proposes an alternative: an information-oriented approach in which government-trained teams of security experts inspect our critical infrastructure facilities and assess their security profiles against a standardized rubric that will allow for facility-by-facility comparisons of security readiness. When that benchmarked information is published, public and market pressure can be focused on those firms that pose the greatest threats to homeland security, thereby prompting meaningful and continuous reductions in terrorism vulnerabilities while at the same time reserving to firms the choice as to the most cost-effective means of mak-

---

[249] *See generally* MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION (1965).

[250] *See generally* Mark Seidenfeld, *A Civic Republican Justification for the Bureaucratic State*, 105 HARV. L. REV. 1511 (1992) (arguing that civic republicanism, through its focus on citizen involvement and collective decisionmaking, offers the most robust foundation for a bureaucratic state).

ing security improvements. While benchmarking in the terrorism context is not immune to criticism, its capacity to promote a flexible and nimble approach to mitigating security vulnerabilities makes it undeniably attractive as a regulatory option. We thus would do well to consider it as an important weapon in the war against terror.

ARTICLE

TERRORISM AND ASYLUM SEEKERS: WHY THE REAL ID ACT IS A FALSE PROMISE

MARISA SILENZI CIANCIARULO*

The Real ID Act, passed on May 11, 2005, is the first post–September 11 antiterrorism legislation specifically to target a group of vulnerable individuals to whom the United States has historically granted protection: asylum seekers. The passage of the Real ID Act led asylum advocates to wring their hands in despair and immigration restrictionists to clap their hands in glee. This Article argues that both sides of the debate may have been justified in their reactions, but not because of the immediate chilling impact on asylum that they seem to expect. With regard to requirements for establishing asylum eligibility, the Real ID Act, rather than imposing new, onerous restrictions on asylum, codifies case law upon which adjudicators, advocates, and government attorneys have been relying for decades. However, several areas of poor drafting, combined with legislative history mischaracterizing the asylum system as a haven for terrorists and suicide bombers, may result in the denial of bona fide asylum applications. This Article provides concrete guidance for adjudicators, advocates, and government attorneys applying the Real ID Act to asylum cases. It examines the case law upon which some of the provisions are based and offers interpretations for unclear provisions. Overall, this Article emphasizes that it is the duty of adjudicators, advocates, and government attorneys to protect victims of persecution.

I. OVERVIEW

The United States' asylum system has emerged as a new battleground in the "War on Terror." On May 11, 2005, Congress passed the Real ID Act,[1] purported antiterrorism legislation specifically targeting asylum seekers, a group of vulnerable non-citizens fleeing persecution to which the United States has historically offered protection. The purpose of the Real ID Act's asylum provisions,[2] according to its author, House Judiciary Committee Chairman James Sensenbrenner (R-Wis.), is to prevent

* Reuschlein Clinical Teaching Fellow, Villanova Law School. Thanks to Michelle Anderson, Bridgette Carr, Michael Carroll, Chapin Cody, David Everson, Beth Lyon, Dveera Segal, and Amy Spare for their insightful comments. Thanks also to Brendan Wilson and Si Nae Lim for their research assistance. Special thanks to Michele Pistone for her inspiration and guidance, and to Marlena Cianciarulo and Carla Cianciarulo Embrey for their encouragement.

[1] Emergency Supplemental Appropriations Act for Defense, the Global War on Terror, and Tsunami Relief, Pub. L. No. 109-13, Div. B, 119 Stat. 231, 302–23 (2005) [Real ID Act].

[2] The Real ID Act addresses other immigration issues, including withholding of removal, judicial review, border security, and driver's license issuance. *See id.*, §§ 1(b), 102(c)(2), 102, 201. These provisions are beyond the scope of this Article.

terrorists from using the U.S. asylum system to gain lawful immigration status in the United States:

> There is no one who is lying through their teeth that should be able to get relief from the courts, and I would just point out that this bill would give immigration judges the tool to get at the Blind Sheikh who wanted to blow up landmarks in New York, the man who plotted and executed the bombing of the World Trade Center in New York, the man who shot up the entrance to the CIA headquarters in northern Virginia, and the man who shot up the El Al counter at Los Angeles International Airport. Every one of these non–9/11 terrorists who tried to kill or did kill honest, law-abiding Americans was an asylum applicant. We ought to give our judges the opportunity to tell these people no and to pass the bill.[3]

Unfortunately, the law has far more potential to undermine the legitimate goals of the asylum system than it does to strengthen national security.

The fact that Chairman Sensenbrenner targeted non-citizens for anti-terrorism legislation is neither surprising nor uncommon. The magnitude of the September 11 terrorist attacks has clouded the fact that they were not the first incidents of terrorism on U.S. soil. According to the USA PATRIOT Act,[4] terrorism consists of criminal "acts dangerous to human life" intended "to influence the policy of a government by intimidation or coercion."[5] Based on this definition, several acts of terrorism occurred in the years leading up to the September 11 attacks. One such act occurred on a date already somewhat faded from the collective U.S. memory: April 19, 1995. On that day, Timothy McVeigh bombed the Murrah Federal Building in Oklahoma City, Oklahoma, killing 168 people, including 19 children, and injuring more than 500 others.[6] Before the Oklahoma City bombing, the Unabomber killed three people and maimed twenty-nine others over a period of seventeen years.[7] Terrorist Eric Rudolph evaded law enforcement for several years before being arrested, tried and convicted for deadly bombings at multiple abortion clinics and the 1996

---

[3] 151 CONG. REC. H460 (daily ed. Feb. 10, 2005) (statement of Chairman Sensenbrenner). *See also* H.R. REP. NO. 109-72, at H2868 (May 3, 2005) (referencing Sheikh Omar Abdel Rahman ("Blind Sheikh"), Ramzi Yousef (1993 World Trade Center bombing), Ahmad Ajaj (1993 World Trade Center bombing), Mir Aimal Kansi (CIA attack), and Hesham Mohamed Ali Hedayet (El Al Airlines murder)).

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

[5] *Id.* § 802(a)(4) (amending 18 U.S.C. §§ 2331(A) and 2331(B)(ii) (2000)).

[6] Serge F. Kovaleski, *Oklahoma Tries to Get Past the Pain; Rebuilding Is the Easier Part of Recovering from April Bombing*, WASH. POST, Aug. 10, 1995, at A1.

[7] George F. Will, *Sanity and the Unabomber*, WASH. POST, Jan. 8, 1998, at A21.

Summer Olympics in Atlanta.[8] Political or religious ideology motivated the perpetrators of all of these acts of domestic terrorism. None of them was an asylum seeker; indeed, each was a U.S. citizen.

While targeting non-citizens as potential terrorists is commonplace, the Real ID Act is unusual in that it illogically focuses on shoring up an asylum system that already was a difficult and unattractive means of gaining legal status in the United States.[9] The Act squarely targets the well-fortified asylum process while ignoring the myriad other, more likely immigration routes available to non-citizens seeking to harm the United States, including over twenty types of non-immigrant visas,[10] several of which were utilized by the September 11 hijackers.[11] Moreover, application of the Real ID Act's asylum provisions is not limited to asylum seekers who may match the profile of a terrorist.[12] It instead affects all asylum seekers, including those fleeing female genital mutilation, domestic violence, religious persecution, politically based persecution, genocide, and ethnic cleansing. Thus, the Real ID Act has the potential to have a severely negative impact on the U.S. asylum system by making acquisition of asylum even more difficult for those who need it most.

As mentioned above, the modus operandi of this legislation is not novel. Examples abound of imprudent anti-terrorism efforts implemented since September 11, purporting to prevent terrorism but in reality only serving the interests of immigration restrictionists. The controversial National Security Exit Entry Registration System (hereinafter NSEERS) required non-citizens from certain countries, all of which were Arab or Muslim, to register with immigration authorities. This initiative led to thousands of detentions and deportations for immigration violations, but to no terrorism-related convictions.[13] A "voluntary interview" program launched within a month of the September 11 attacks had federal law en-

---

[8] Jay Reeves, *Clinic Bomber Gets 2 Life Sentences: Rudolph is Unrepentant, Says Abortion Must Be Fought "With Deadly Force,"* WASH. POST, July 19, 2005, at A5.

[9] *See infra* Part III (discussing the current asylum system, including significant changes made in 1996).

[10] *See* 8 U.S.C § 1101(a)(15) (2000) (setting forth the nonimmigrant visas available to eligible non-citizens); *see also* 8 U.S.C. § 1101(a)(26) (2000) (defining the term "nonimmigrant visa").

[11] *See* THOMAS R. ELDRIDGE ET AL., 9/11 AND TERRORIST TRAVEL: STAFF REPORT OF THE NATIONAL COMMISSION ON TERRORIST ATTACKS UPON THE UNITED STATES 7–31 (2004) (detailing each of the September 11 hijackers' visa application processes and encounters with U.S. immigration personnel) [hereinafter 9/11 AND TERRORIST TRAVEL]. Most of the hijackers applied for and received tourist visas. *See id.* One applied for and received a student visa after being denied a tourist visa. *See id.* at 13–14.

[12] *See* Pub. L. No. 109-13, 119 Stat. 231, 302–23 (2005).

[13] LEADERSHIP CONFERENCE ON CIVIL RIGHTS EDUCATIONAL FUND & AMERICAN BAR ASSOCIATION COMMISSION ON IMMIGRATION, AMERICAN JUSTICE THROUGH IMMIGRANTS' EYES 107 (2004) [hereinafter AMERICAN JUSTICE THROUGH IMMIGRANTS' EYES]. The government claimed to have gained significant leads in the terrorism investigation but declined to provide any information to the public. *See id. See also* Dalia Hashad, *Stolen Freedoms: Arabs, Muslims, and South Asians in the Wake of Post 9/11 Backlash*, 81 DENV. U. L. REV. 735, 743–44 (2004) (discussing the shortcomings of NSEERS).

forcement agents interviewing thousands of male nationals of Arab and Muslim countries, but did not turn up any significant reported leads in the terrorism investigation.[14] In perhaps the most misguided post–September 11 action prior to the passage of the Real ID Act, the Bush Administration suspended refugee resettlement, stranding thousands of refugees in dangerous, disease-ridden refugee camps, even though none of the September 11 terrorists (or any terrorist in U.S. history) entered the country through the refugee resettlement program.[15] While these measures may have served to provide a sense of security to non-Arab, non-Muslim U.S. citizens, they achieved no apparent or disclosed progress in the War on Terror.

The Real ID Act operates by taking advantage of several prevalent, xenophobic misconceptions held in U.S. society—that all non-citizens are potential terrorists, that an application for asylum is a free and easy pass into the United States, and that U.S. law does not give immigration officials sufficient authority to remove unwanted non-citizens from the country—to pass a restrictionist immigration law that does nothing to strengthen the asylum system against terrorism. Even the name of the Act's section dealing with asylum, "Preventing Terrorists from Obtaining Relief from Removal,"[16] is testament to its alarmist agenda. Amidst that alarmism, several salient facts are obscured. For example, on average, less than thirty percent of asylum claims prevail.[17] Moreover, applicants found to have fabricated asylum claims are banished for life from the United States.[18] Also, all of the terrorists' applications that Chairman Sensenbrenner mentions as evidence of a faulty asylum system[19] were submitted prior to the implementation of stricter asylum provisions contained in the Illegal Immigration Reform and Immigrant Responsibility Act of

---

[14] *See* Susan M. Akram & Maritza Karmely, *Immigration and Constitutional Consequences of Post-9/11 Policies Involving Arabs and Muslims in the United States: Is Alienage a Distinction Without a Difference?*, 38 U.C. DAVIS L. REV. 609, 628–29 (2005) (describing the "voluntary interview" program launched by the Bush Administration); *see also* AMERICAN JUSTICE THROUGH IMMIGRANTS' EYES, *supra* note 13, at 104 (pointing out that the "voluntary interviews" coincided with a massive law enforcement sweep in which over 1,000 nationals of Arab and Muslim countries were arrested for immigration violations).

[15] *See generally*, Marisa S. Cianciarulo, *The W Visa: A Legislative Proposal for Female and Child Refugees Trapped in a Post-9/11 World*, 17 YALE J.L. & FEMINISM (forthcoming Fall 2005) (discussing the suspension of the U.S. refugee resettlement program in the wake of September 11).

[16] Pub. L. No. 109-13, § 101, 119 Stat. 231, 302 (2005).

[17] *See* DEPT. OF HOMELAND SECURITY, OFFICE OF IMMIGRATION STATISTICS, 2003 YEARBOOK OF IMMIGRATION STATISTICS 56 (2004) (reporting that between 1973 and 2003, the U.S. government approved twenty-eight percent of asylum applications) [hereinafter 2003 YEARBOOK].

[18] *See* 8 U.S.C. § 1158(d)(6) (2000); 8 C.F.R. §§ 208.3(c)(5), 1208.3(c)(5) (2005). Terrorists and persons deemed to be a threat to national security are also barred from asylum eligibility. 8 U.S.C. § 1158(b)(2)(A)(iv)–(v) (2000).

[19] *See supra* note 3 and accompanying text.

1996[20] and were denied even under the less strict provisions in place at the time.[21]

In addition, Chairman Sensenbrenner's ill-advised attempt to prevent terrorism at first glance appears merely to codify existing case law governing asylum claims.[22] A closer reading of the Board of Immigration Appeals and Circuit Courts of Appeals cases from which the Real ID Act draws, however, demonstrates subtle but nonetheless significant differences between the language of the Real ID Act and that of the cases that influenced it.[23] Where those cases are thoughtful and thoroughly reasoned, the Real ID Act is careless and rash.[24] Moreover, the previous case law approach allowed for individualized interpretation and evolution of asylum law, whereas the Real ID Act, as a statute, is far more rigid. Improper interpretation of the Real ID Act's language may have devastating consequences for bona fide asylum applicants while providing no additional protection against fraudulent claims.

This Article provides guidance for asylum adjudicators charged with the daunting task of interpreting the Real ID Act. Part II briefly explores the history of U.S. refugee law, from World War II, and the subsequent 1968 ratification of the 1967 Protocol Relating to the Status of Refugees,[25] to the Refugee Act of 1980.[26] Part III describes pre-Real ID Act

---

[20] Omnibus Consolidated Appropriations Act of 1997, Pub. L. No. 104-208, Div. C, 110 Stat. 3009, 3009-546–3009-724 (1996). *See also* 9/11 AND TERRORIST TRAVEL, *supra* note 11, at 47–48 (reporting that World Trade Center bombing perpetrators Ramzi Yousef and Ahmad Ajaj applied for asylum in 1992); *id.* at 51 (stating that Sheikh Omar Abdel Rahman filed an application for asylum on Aug. 27, 1992); *id.* at 215 (stating that Mir Aimal Kansi applied for asylum on Feb. 7, 1992); *id.* at 230 (reporting that Hesham Mohamed Ali Hedayet applied for asylum on Dec. 1, 1992).

[21] *See id.* at 51 (stating that an immigration judge denied Rahman's asylum application on Mar. 16, 1993); *id.* at 48 (reporting that Ajaj's asylum request was denied on Apr. 24, 1993 and that Yousef's application was never adjudicated because he was convicted of carrying out the 1993 World Trade Center bombing and sentenced to 240 years in prison); *id.* at 215 (reporting that Kansi's asylum application was denied); *id.* at 230 (stating that on Mar. 7, 1995, the INS issued a Notice of Intent to Deny Hedayet's asylum application and that Hedayet failed to respond to the notice within thirty days, resulting in the initiation of deportation proceedings against him).

[22] *Compare, e.g.*, Pub. L. No. 109-13, § 101(a)(3)(B)(ii), 119 Stat. 231, 303 (2005) ("Where the trier of fact determines that the applicant should provide evidence that corroborates otherwise credible testimony, such evidence must be provided unless the applicant does not have the evidence and cannot reasonably obtain the evidence.") *with* Matter of S-M-J-, 21 I&N Dec. 722, 725 (B.I.A. 1997).

[23] *See generally* Part IV *infra*.

[24] *Compare, e.g.*, § 101(a)(3)(B)(ii), 119 Stat. at 303 (failing to impose a "reasonableness" requirement) *with* Secaida-Rosales v. INS, 331 F.3d 297, 311 (2d Cir. 2003) (vacating a denial of asylum where, inter alia, the judge's demands for corroboration of testimony were unreasonable), Georgis v. Ashcroft, 328 F.3d 962, 969 (7th Cir. 2003) (same), and Senathirajah v. INS, 157 F.3d 210, 216 (3d Cir. 1998) (vacating a denial of withholding where the judge's demands for corroboration were unreasonable in light of the validity of applicant's testimony).

[25] Jan. 31, 1967, 19 U.S.T. 6223, 606 U.N.T.S. 267 [hereinafter 1967 Protocol].

[26] Pub. L. No. 96-212, 94 Stat. 102 (1980) (codified as amended in scattered sections of 8 U.S.C.) [Refugee Act].

statutory restrictions on asylum and discusses the passage of the Real ID Act itself. Part IV analyzes each asylum provision of the Real ID Act in terms of relevant jurisprudence and regulations and suggests how each provision should be interpreted. This Article concludes that asylum adjudicators have a duty to interpret the Real ID Act in the spirit of the humanitarian treaties and laws upon which the asylum system is based.

II. U.S. ASYLUM AND REFUGEE LAW: 1939–1980

A. *Fleeing Religious Persecution: the Tragedy of the S.S. St. Louis*

The United States has a long history of protecting individuals who are fleeing persecution. The country itself was founded as a shelter from religious persecution, and quickly became home to Quakers, Puritans, Catholics, Huguenots, and other religious denominations unwelcome in seventeenth and eighteenth-century Europe.[27] That tradition of sanctuary has survived several periods of intense xenophobia, racially based exclusionary policies, national security threats, and war.

This history of refugee protection, however, is not unblemished. One of the country's most egregious failures to protect persons fleeing religious and ethnic persecution occurred on June 6, 1939. On that date, the German transatlantic liner *St. Louis* was forced to return to Europe, its 937 passengers having been denied entry to the United States. Most of them were European Jews fleeing Nazi persecution.[28]

At the time, the United States did not have laws specifically permitting refugee admissions. Immigration occurred primarily through a nationality-based quota system; when the allotted number of visas ran out for a particular country or region, applicants had to wait until a visa became available in order to immigrate.[29] At the time of the *St. Louis*'s voyage, the German-Austrian quota had not only been filled but had a waiting list of several years.[30] Entry to the United States would have required an executive order from President Roosevelt, who declined to issue one.[31]

---

[27] *See* International Religious Freedom Act of 1998, 22 U.S.C. § 6401(a)(1) (2000) (noting that "[m]any of our Nation's founders fled religious persecution abroad, cherishing in their hearts and minds the ideal of religious freedom"); WILLIAM CARLSON SMITH, AMERICANS IN THE MAKING: THE NATURAL HISTORY OF THE ASSIMILATION OF IMMIGRANTS 4 (Edward Alsworth Ross ed., 1939) (noting that the United States was "vaunted as a land not only of economic opportunity but also of religious freedom"). *See generally* ROGER DANIELS, COMING TO AMERICA: A HISTORY OF IMMIGRATION AND ETHNICITY IN AMERICAN LIFE 94 (1990).

[28] United States Holocaust Memorial Museum, *Voyage of the "St. Louis,"* http://www.ushmm.org/wlc/en/index.php?ModuleId=10005267 (last visited Nov. 19, 2005) [hereinafter *Voyage of the "St. Louis"*]. The refugees were originally en route to Cuba, but the Cuban government revoked their landing passes and denied them entry. *Id.*

[29] Immigration Act of 1924, ch. 190, 43 Stat. 153, 159–65, *repealed by* Immigration and Nationality Act of 1952, ch. 477, tit. IV, § 403(a)(23), 66 Stat. 163, 279.

[30] *Voyage of the "St. Louis," supra* note 28.

[31] *Id.*

While the *St. Louis* made its way back to Europe, Jewish organizations secured admission for most of the refugees to western European countries. The passengers eventually settled in Belgium, France, Great Britain, and the Netherlands to await their turn to enter the United States.[32] Approximately four months after the *St. Louis*'s return to Europe, World War II began. With the exception of Great Britain, all of the countries to which the *St. Louis* passengers were sent subsequently came under Nazi control. Many of the *St. Louis*'s passengers were forced into hiding, driven into Nazi labor camps, or killed in the Holocaust.[33]

B. The 1967 Protocol Relating to the Status of Refugees

In the aftermath of the Nazi atrocities of World War II, refugee protection gained prominence in the international community. The United Nations General Assembly promulgated the Convention Relating to the Status of Refugees in 1951[34] specifically to provide protection to refugees displaced as a result of World War II.

The United States, however, did not ratify the 1951 Convention, instead choosing to reformulate an independent asylum policy in the Immigration and Nationality Act of 1952.[35] Under the 1952 Act, the Attorney General was given authorization "to withhold the deportation of any alien within the United States to any country in which in his opinion the alien would be subject to physical persecution and for such period of time as he deems to be necessary for such reason."[36]

Refugees gained a more formal immigration status when Congress amended the 1952 Act in 1965,[37] but even that gesture was born largely of the United States' Cold War political concerns rather than humanitarian interests.[38] The 1965 Amendments allowed only refugees from either communist countries[39] or countries in the "general area of the Middle East" to qualify for asylum.[40] Asylum seekers falling within these narrow parameters still had to demonstrate a "clear probability" of persecution (a higher

---

[32] *Id.*

[33] United States Holocaust Memorial Museum, *Wartime Fate of the Passengers of the "St. Louis,"* http://www.ushmm.org/wlc/en/index.php?ModuleId=10005267 (last visited Nov. 19, 2005).

[34] July 28, 1951, 189 U.N.T.S. 137 [hereinafter 1951 Convention].

[35] Pub. L. No. 82-414, 66 Stat. 163 (1952).

[36] *Id.* at 214 (current version at 8 U.S.C. § 1253 (2000)).

[37] An Act to Amend the Immigration and Nationality Act, Pub. L. No. 89-236, 79 Stat. 911, 913 (1965).

[38] *See* Deborah E. Anker & Michael J. Posner, *The Forty Year Crisis: A Legislative History of the Refugee Act of 1980*, 19 SAN DIEGO L. REV. 9, 13–14 (1981) (remarking that Congress's exceptions to immigration policy were strictly responses to Soviet expansionism, and should not be viewed as humanitarian commitments).

[39] *See* An Act to Amend the Immigration and Nationality Act § 203(a)(7) (current version at 8 U.S.C. § 1153 (2000)).

[40] *Id.*

standard than the "reasonable possibility" standard that exists today) before being accepted as refugees.[41] The 1965 Amendments also retained strict numerical limitations.[42]

In 1967, the United Nations updated the 1951 Convention with the Protocol Relating to the Status of Refugees,[43] designed to address any refugee flows arising out of persecution-related events after World War II.[44] In 1968, the United States seemed to align with the international community's refugee policy by signing and ratifying the 1967 Protocol.[45] By acceding to the 1967 Protocol, the United States agreed that "equal status should be enjoyed by all refugees . . . irrespective of the dateline 1 January 1951 . . . ."[46] The Protocol defined "refugee" as any person who:

> [O]wing to well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his [or her] nationality and is unable or, owing to such fear, is unwilling to avail himself [or herself] of the protection of that country; or who, not having a nationality and being outside the country of his [or her] former habitual residence as a result of such events, is unable or, owing to such fear, is unwilling to return to it.[47]

Yet, the United States' assent to the 1967 Protocol did not have a significant effect on asylum processing. From 1968 to 1980, the United States continued to enforce the narrow parameters, low ceiling on approvals, and strict burden of proof mandated by the amended 1952 Act and the courts' interpretation of it.[48]

---

[41] *See, e.g.*, Pierre v. U.S., 547 F.2d 1281, 1289 (5th Cir. 1977) (holding that the burden was on the asylum seeker to show that she was a refugee by a "clear probability" standard of proof); Kashani v. INS, 547 F.2d 376, 379 (7th Cir. 1977) (same); Cisternas-Estay v. INS, 531 F.2d 155, 159 (3d Cir. 1976) (sustaining denial of asylum where applicants failed to demonstrate a "clear probability" of persecution).

[42] *See* An Act to Amend the Immigration and Nationality Act § 203(a)(7) (current version at 8 U.S.C. § 1153 (2000)) (specifying that of the 170,000 visas available yearly, no more than six percent should be granted for asylum applicants).

[43] 1967 Protocol, *supra* note 25.

[44] *See id.* pmbl.

[45] *Id.* (specifying that the Senate ratified the Protocol on Oct. 4, 1968 and the President signed it on Oct. 15, 1968).

[46] *Id.* pmbl.

[47] *Id.* art. 1, ¶ 2.

[48] *See* KAREN MUSALO ET AL., REFUGEE LAW AND POLICY: CASES AND MATERIALS 65 (1997); INS v. Stevic, 467 U.S. 407, 429–30 (1984) (articulating the standard for eligibility for withholding of removal).

C. Refugee Act of 1980

With the Refugee Act of 1980,[49] Congress for the first time passed a law specifically addressing refugees and asylum seekers. By enacting the Refugee Act, Congress sought to give "statutory meaning to our national commitment to human rights and humanitarian concerns."[50] The Refugee Act repealed the 1952 Act's geographical and political limitations on the asylum process,[51] explicitly adopted the 1967 Protocol's definition of "refugee,"[52] formulated a legal right to seek asylum in the United States,[53] and lifted the numerical caps on yearly grants of asylum.[54] In addition, the Refugee Act mandated that the Attorney General establish procedures for asylum processing.[55]

---

[49] Refugee Act, Pub. L. No. 96-212, 94 Stat. 102 (1980).

[50] S. REP. No. 96-256, at 4 (1980), *reprinted in* 1980 U.S.C.C.A.N. 141, 144; *see also* Refugee Act, § 101(a) ("[I]t is the historic policy of the United States to respond to the urgent needs of persons subject to persecution in their homelands . . . .").

[51] Refugee Act, sec. 203(c)(3), § 203(a)(7), 94 Stat. 102. However, the Refugee Act failed to alleviate some of the political biases that had existed in asylum processing before 1980. Between 1984 and 1990, the United States' geopolitical concerns led to disparate treatment of Central American asylum applicants fleeing human rights abuses arising from civil wars in El Salvador, Guatemala, and Nicaragua. During that time period, the United States granted only 2.6% and 1.8% of claims from those fleeing American-backed regimes in El Salvador and Guatemala, respectively, compared to 26% of the asylum requests from those fleeing the communist regime in Nicaragua. *See* Sharon S. Russell, *Migration Patterns of U.S. Foreign Policy Interest*, *in* THREATENED PEOPLES, THREATENED BORDERS 50–67 (Michael S. Teitelbaum & Myron Weiner eds., 1995). That particular bias finally became known during a series of lawsuits spearheaded by the American Baptist Churches against the Immigration and Naturalization Service, the Department of State, the Department of Justice, and the Attorney General. *See* American Baptist Churches v. Thornburgh, 760 F. Supp. 796 (N.D. Cal. 1991); American Baptist Churches v. Meese, 712 F. Supp. 756 (N.D. Cal. 1989); American Baptist Churches v. Meese, 666 F. Supp. 1358 (N.D. Cal. 1987). These lawsuits, which challenged the government's treatment of Salvadoran and Guatemalan asylum seekers and the government's prosecution of those providing sanctuary to them, culminated in a 1991 settlement approved by the District Court for the Northern District of California, in which the Department of Justice conceded, inter alia, that foreign policy, governmental relations with the applicant's country of origin, and the applicant's political beliefs "are not relevant to the determination of whether an applicant for asylum has a well-founded fear of persecution . . . ." *Thornburgh*, 760 F. Supp. at 799.

[52] *See* § 201(a) (defining a refugee as:

> Any person who is outside any country of such person's nationality or, in the case of a person having no nationality, is outside any country in which such person last habitually resided, and who is unable or unwilling to return to, and who is unable or unwilling to avail himself or herself of the protection of, that country because of persecution or a well-founded fear of persecution on account of race, religion, nationality, membership in a particular social group, or political opinion.)

[53] § 208(a) (authorizing "*[a]ny* alien who is physically present in the United States or at a land border or port of entry . . . to apply for asylum . . . .") (emphasis added).

[54] *Id.* (authorizing the Attorney General to grant asylum to any alien who meets the definition of refugee, without any numerical restrictions).

[55] *Id.* The Attorney General issued regulations in 1990 that created a professional corps of asylum officers; vested initial jurisdiction of affirmative asylum claims with the Office of Refugees, Asylum and Parole; established filing procedures for applications for asylum;

The passage of the Refugee Act ushered in a new era of refugee protection. The Supreme Court recognized the implications of the Refugee Act in the groundbreaking case of *INS v. Cardoza-Fonseca*,[56] which articulated a new, lower standard of proof for asylum eligibility, differentiating it from that of withholding of removal.[57] Asylum was no longer an ad hoc, marginal immigration procedure entirely subject to the whims of policy. Over the next twenty-five years, asylum would emerge as a unique, complex body of law and a lightning rod for the national immigration debate, forcing the country to balance traditional humanitarian interests against weighty national security concerns.

III. U.S. ASYLUM AND REFUGEE LAW: 1980–2005

A. *Illegal Immigration Reform and Immigrant Responsibility Act of 1996*

By the mid-1990s, lawmakers were aware of perceived flaws in the U.S. asylum system. Processing delays had led to a backlog of several years,[58] during which time asylum applicants could both legally remain in the United States and apply for immediate work authorization, renewable on a yearly basis until the asylum adjudication was complete.[59] This loophole allowed economic migrants, unscrupulous individuals, or even potential terrorists to avoid deportation and then to abscond as their applications went unexamined.[60]

Congress finally addressed these concerns in the Illegal Immigration Reform and Immigrant Responsibility Act of 1996.[61] The 1996 Act put into effect a number of provisions designed to curtail abuse of the asylum system. The most significant limitations were a one-year deadline on applying for asylum,[62] delay in work authorization eligibility,[63] prompt adjudication of asylum applications,[64] expedited removal,[65] and detention of asylum seek-

---

established interview procedures; set forth eligibility requirements; and established procedures for granting derivative status to immediate family members. *See* INS Asylum Procedures, 8 C.F.R. § 208 (1990).

[56] 480 U.S. 421 (1987).

[57] *Id.* at 430–33.

[58] *See* News Release, U.S. Dep't of Justice, Immigration and Naturalization Service, Asylum Reform: Five Years Later (Feb. 1, 2000), *available at* http://uscis.gov/graphics/publicaffairs/newsrels/Asylum.htm [hereinafter *Asylum Reform*].

[59] *See* 8 C.F.R. § 208.7 (1993).

[60] *See Asylum Reform*, *supra* note 58 ("By 1993, the asylum system was in a crisis, having become a magnet for abuse by persons filing applications in order to obtain employment authorization.").

[61] Omnibus Consolidated Appropriations Act of 1997, Pub. L. No. 104-208, Div. C, 110 Stat. 3009-546 to 3009-724 (1996).

[62] § 604(a), 110 Stat. at 3009-691 (amending 8 U.S.C. § 1158(a)(2)(B) (1994)).

[63] § 604(a), 110 Stat. at 3009-693 (amending 8 U.S.C. § 1158(d)(2) (1994)).

[64] § 604(a), 110 Stat. at 3009-694 (amending 8 U.S.C. § 1158(d)(5)(A)(iii) (1994)) (setting the maximum time for final adjudication at 180 days after application filing).

[65] § 302(a), 110 Stat. at 3009-581 (amending 8 U.S.C. § 1225(b)(1)(B)(iii)(I)–(III) (1994)).

ers.[66] With these provisions in place, the asylum process has become an unlikely choice for an individual seeking an easy, low-profile way to gain lawful immigration status.

1. The One-Year Deadline

As of April 1, 1997, asylum seekers must file their applications for asylum within one year of their entry into the United States.[67] An applicant's failure to prove by clear and convincing evidence that he or she filed within one year of entry bars the applicant from asylum eligibility.[68] Applicants may only overcome the bar if they demonstrate "changed circumstances which materially affect the applicant's eligibility for asylum or extraordinary circumstances relating to the delay in filing an application . . . ."[69] The purpose of this provision is to ensure that individuals applying for asylum do so as the result of an urgent need for protection, rather than as a delay tactic to prolong an unauthorized stay in the United States.

2. Delay in Work Authorization Eligibility and Prompt Adjudication of Asylum Claims

By revoking employment authorization and mandating prompt adjudication of asylum claims, the 1996 Act closed another alleged loophole in the asylum system. The revised provision plainly states that "[a]n applicant for asylum is not entitled to employment authorization."[70] Congress authorized the Attorney General to provide for employment authorization via regulation, but stipulated that such authorization "shall not be granted . . . prior to 180 days after the date of filing of the application for asylum."[71] Therefore, because the 1996 Act also mandates that asylum cases be adjudicated within 180 days of receipt of application,[72] very few asylum seekers will qualify for employment authorization absent a final grant of asylum.

Moreover, the regulations stipulate that "[a]ny delay requested or caused by the applicant shall not be counted as part of" the 180-day delay before eligibility for employment authorization.[73] Thus, even if a denied asylum applicant's appeal puts him or her beyond the typical 180-day mark, he or she remains ineligible for employment authorization during that appellate period.

---

[66] *Id.* (amending 8 U.S.C. § 1225(b)(1)(B)(iii)(IV) (1994)).
[67] 8 U.S.C. § 1158(a)(2)(B) (2000).
[68] *Id.*
[69] *Id.* § 1158(a)(2)(D).
[70] *Id.* § 1158(d)(2).
[71] *Id.*
[72] *Id.* § 1158(d)(5)(A)(iii).
[73] 8 C.F.R. § 208.7(a)(2) (2005).

3. *Expedited Removal*

The expedited removal provisions of the 1996 Act authorize immigration officers at U.S. ports of entry to expel aliens deemed inadmissible for failure to provide valid entry documents.[74] Such removals are referred to as "expedited" because they are not subject to rehearing or review by a judge.[75] An individual who receives an order of expedited removal is barred from reentering the United States for at least five years.[76]

Only those individuals who express a fear of returning to their home country receive an opportunity to avoid being summarily deported. The 1996 Act provides that an asylum officer should interview any such individuals to determine whether the expressed fears are credible.[77] If the asylum officer determines from the "credible fear" interview that the individual has a "significant possibility . . . [of] establish[ing] eligibility for asylum,"[78] the individual may remain in the United States to pursue asylum before an immigration judge.[79] If the asylum officer does not believe the individual has a credible fear of persecution, the individual may be summarily removed.[80]

4. *Detention of Asylum Seekers*

Claiming asylum at a port of entry and even establishing a credible fear of persecution by no means guarantees an easy entry into the United States. Individuals subject to expedited removal for attempting to enter the United States without valid documentation, including those claiming asylum, are subject to mandatory detention under the Act.[81] The Department of Homeland Security usually detains "credible fear" interviewees in immigration detention facilities, or, more commonly, in county jails from which the Department rents bed space.[82] Therefore, many applicants are forced to spend several months or even years in uncomfortable detention quarters, awaiting the adjudication of an asylum application that has less than a thirty percent chance of success.[83]

---

[74] *See* 8 U.S.C. § 1225(b)(1) (2000) (stating that "the officer shall order the alien removed from the United States *without further hearing or review*") (emphasis added); *id.* § 1182(a)(6)(C) (rendering inadmissible persons who attempt to commit fraud to enter the United States); *id.* § 1182(a)(7) (2000) (rendering inadmissible persons who attempt to enter the United States without a visa).

[75] *Id.* §§ 1225(b)(1)(A)(i), 1252(a)(2)(A)(i).

[76] *Id.* § 1182(a)(9)(A)(i).

[77] *Id.* § 1225(b)(1)(A)(ii).

[78] *Id.* § 1225(b)(1)(B)(v); 8 C.F.R. § 208.30(e)(2) (2005).

[79] 8 C.F.R. § 208.30(f) (2005).

[80] 8 U.S.C. § 1225(b)(1)(B) (2000).

[81] *Id.* § 1225(b)(1)(B)(iii)(IV).

[82] AMERICAN BAR ASSOCIATION COMMISSION ON IMMIGRATION, IMMIGRATION DETAINEE PRO BONO OPPORTUNITIES GUIDE 1 (2004).

[83] *See* 2003 YEARBOOK, *supra* note 17, at 56.

B. The Asylum Application Process

An asylum applicant may apply for asylum in one of two ways: affirmatively, by filing an application with U.S. Citizenship and Immigration Services ("CIS"), or defensively, by filing with the immigration court as a defense in removal proceedings.[84] Persons eligible to apply affirmatively for asylum are those who have entered the country lawfully or who have entered illegally but evaded detection.[85] Individuals who request asylum upon entry to the United States, or who are apprehended upon entry for lack of valid entry documents, or who otherwise become subject to the jurisdiction of U.S. Immigration and Customs Enforcement ("ICE") may apply for asylum in front of an immigration judge.[86]

Both affirmative and defensive applicants must undergo identity verification and background checks before being eligible for asylum.[87] The government issues each asylum applicant a file number, or "alien number," which is entered into the Refugees, Asylum and Parole System ("RAPS") database.[88] RAPS interfaces with both the Computer Linked Applicant Information System ("CLAIMS") to identify and update asylum applicants' address changes, and with the Receipt and Alien File Accountability Control System ("RAFACS") to keep track of asylum applicants' files.[89] The asylum office may not grant asylum without first checking the identity of the applicant against all appropriate government databases, including the State Department's Consular Lookout and Support System ("CLASS")[90] and the Department of Homeland Security biometric identification system known as "IDENT."[91]

All affirmative applicants and their dependents must attend a face-to-face interview with a CIS asylum officer,[92] in order to "elicit all useful and relevant information bearing on the applicant's eligibility for asylum." During this "nonadversarial"[93] interview, the applicant "may have counsel or a representative present, may present witnesses, and may submit affidavits of witnesses and other evidence."[94] If the asylum office approves

---

[84] 8 C.F.R. § 208.1(a) (2005).

[85] *Id.* § 208.2(a).

[86] *Id.* §§ 208.2(b), 208.4(b)(3).

[87] *Id.* §§ 208.9(b), 208.10, 240.67, 1240.67.

[88] U.S. DEP'T OF JUSTICE, OFFICE OF THE INSPECTOR GENERAL, IMMIGRATION AND NATURALIZATION SERVICE REFUGEES, ASYLUM AND PAROLE SYSTEM AUDIT REPORT (1998), *available at* http://www.usdoj.gov/oig/reports/INS/a9811.htm.

[89] *Id.*

[90] IRA J. KURZBAN, IMMIGRATION LAW SOURCEBOOK 383 (9th ed. 2004).

[91] *Id.* at 100.

[92] 8 C.F.R. § 208.9(b) (2005). The interview is conducted at one of eight U.S. asylum offices. *See* U.S. Citizenship and Immigration Services, Asylum Offices, *at* http://uscis.gov/graphics/fieldoffices/aboutus.htm#asylum (last visited Nov. 18, 2005).

[93] 8 C.F.R. § 208.9(b) (2005). In the experience of the author, the meaning of "nonadversarial" varies among asylum officers. Some asylum officers conduct interviews in a pleasant, non-threatening manner, while others tend to be aggressive and confrontational.

[94] *Id.*

the case, the applicant will be granted asylum upon completion of the security checks. If the asylum office denies the case, and the applicant is not in lawful status, he or she receives notice of removal proceedings and must appear before an immigration judge.[95]

An asylum applicant in removal proceedings is under the jurisdiction of the Department of Justice Executive Office for Immigration Review (EOIR).[96] Unlike the asylum interview, a removal hearing is an adversarial process. A trial attorney for ICE prosecutes the case.[97] Immigration court respondents may be represented by counsel, but not at government expense.[98] Each side may present documentary and testimonial evidence, conduct direct and cross examination, and object to the evidence and questions of the other side.[99] If an immigration judge declines to grant asylum or other relief, that judge has the authority to issue an order of removal.[100]

Both ICE prosecutors and asylum applicants are entitled to appeal immigration judges' decisions to the Board of Immigration Appeals (BIA).[101] Regulations mandate that the BIA complete appeals within 180 days.[102] During that time, asylum applicants must continue to keep the government apprised of their whereabouts.[103]

C. The Real ID Act's Passage

With the implementation of the laws and procedures discussed above, asylum—despite never having been a particularly popular avenue for terrorists to pursue—had become even more difficult to abuse. Filing deadlines, restrictions on employment authorization, face-to-face interviews with immigration officers, numerous background and identity checks, and mandatory detention combined to create a less-than-ideal environment for an individual seeking to abuse the asylum process.

The 1996 Act, and particularly its expedited removal process, has been praised as "a key tool in overall border and anti-fraud strategy."[104]

---

[95] *Id.* § 208.14(c)(1).

[96] *Id.* § 1003.14(a).

[97] *Id.* § 1003.16(a).

[98] 8 U.S.C. § 1229a(b)(4)(A) (2000); 8 C.F.R. § 1003.16(b) (2005).

[99] 8 U.S.C. § 1229a(b)(4)(B) (2000).

[100] *Id.* § 1229a(c)(1)(A), 8 C.F.R. § 208.14(a) (2005).

[101] 8 C.F.R. §§ 1003.1(b), 1003.38(a) (2005).

[102] *Id.* § 1003.1(e)(8)(i). The 180-day rule applies to those appeals reviewed by a three-member panel of BIA members. Most appeals are subject to single member review, however, and must be adjudicated within ninety days. *Id.*

[103] 8 U.S.C. § 1229(a)(1)(F) (2000).

[104] David A. Martin, *Two Cheers for Expedited Removal in the New Immigration Laws*, 40 VA. J. INT'L L. 673, 675 (2000); *see also* Bo Cooper, *Procedures for Expedited Removal and Asylum Screening Under the Illegal Immigration Reform and Immigrant Responsibility Act of 1996*, 29 CONN. L. REV. 1501, 1524 (1997) (stating that "Congress has created procedures that appear susceptible of fair application meeting the international standards, and the executive branch has taken a number of important steps that go far toward ensuring adequate protection of asylum seekers who fall within those procedures").

Thus, despite no indication that the asylum system needed strengthening beyond the provisions of the 1996 Act, the Real ID Act, authored by House Judiciary Committee Chairman Sensenbrenner, was attached to an emergency Iraq appropriations bill[105] and passed on May 11, 2005.[106] Because legislators considered the appropriations bill to be "must-sign" legislation,[107] the Real ID Act passed without floor debate.[108] This stands in stark contrast to the debate given to most other pre–September 11 anti-terrorism legislation, such as PATRIOT Act,[109] even though both gave law enforcement and other government officials authority to detain and investigate noncitizens.

IV. INTERPRETING THE REAL ID ACT

The Real ID Act addresses three main areas of asylum law. First, the Real ID Act attempts to codify case law on claims in which the persecutor may have been motivated by several factors, one or more of which may not conform to the definition of a refugee.[110] These "mixed motive" cases have generated conflicting jurisprudence,[111] but the Real ID Act's

---

Both David A. Martin and Bo Cooper are former INS General Counsels.

[105] *See* Emergency Supplemental Appropriations Act for Defense, the Global War on Terror, and Tsunami Relief of 2005, Pub. L. No. 109-13, 119 Stat. 231 (2005); Suzanne Gamboa, *More Visas for Nurses, Seasonal Workers Added to Iraq Spending Bill*, AP, May 3, 2005 ("House and Senate negotiators reached agreement Tuesday in the final $82 billion bill devoted primarily to paying for military operations and reconstruction in Iraq and Afghanistan.").

[106] Real ID Act, Pub. L. No. 109-13, Div. B, 119 Stat. 231 (2005).

[107] 151 CONG. REC. S4816, 4837 (daily ed. May 10, 2005) (statement of Sen. Chafee (R-R.I.)) ("By attaching REAL ID to a must pass spending measure, the critical process of vetting the bill in committee was circumvented and an opportunity for discussion and debate, which is essential for effective legislation, was denied."); *id.* at 4831 (Statement of Sen. Obama (D-Ill.)) (stating that the immigration provisions of the Real ID Act will become law "not because it is the right thing to do but because the House majority has abused its privilege to attach this unexamined bill to must-pass legislation").

[108] *See id.* at 4820 (statement of Sen. Byrd (D-W. Va.)) (expressing disappointment that the Real ID Act "was simply grafted onto the emergency supplemental appropriations bill that provides funding for our military operations and our troops, without debate or participation by the conferees"); *id.* at 4826 (statement of Sen. Murray (D-Wash.)) (expressing concern that "far-reaching and unrelated immigration rules got attached to this bill without a vote and without an opportunity for debate"); *id.* at 4831 (statement of Sen. Obama) (expressing concern that, despite the controversial immigration provisions of the Real ID Act, "the Senate did not conduct a full hearing or debate on any one of them").

[109] *See generally* Akram & Karmely, *supra* note 14 (examining the PATRIOT Act and other pre–September 11 policies and legislation targeting Arab and Muslim citizens and noncitizens).

[110] Real ID Act, Pub. L. No. 109-13, Div. B, § 101(a)(3)(B)(i), 119 Stat. 231, 303 (2005).

[111] *Compare, e.g.*, Singh v. Gonzalez, 406 F.3d 191, 197 (3d Cir. 2005) (holding that "an applicant for asylum need not prove that the persecution he or she suffered (or fears suffering in the future) occurred *solely* on account of one of the five grounds enumerated in the INA. Rather, an applicant must show that the persecution was motivated, *at least in part*, by one of the protected characteristics.") *with* Gebremichael v. INS, 10 F.3d 28, 35 (1st Cir. 1993) (holding that the enumerated ground must be at the "root of persecution").

attempt to codify the majority view creates more confusion than it alleviates.

Second, the Real ID Act addresses corroboration requirements. Again, its drafters attempted to codify long-established case law regarding when adjudicators may require additional evidence apart from the applicant's testimony alone.[112] Yet, in doing so, the drafters failed to explicitly include the "reasonableness" requirement found in leading case law.[113] This careless drafting leaves room for adjudicators to abuse their discretion in deciding when to require corroboration by making patently unreasonable corroboration demands on asylum applicants.

Finally, the Real ID Act addresses credibility. In most respects, the credibility provision of the statute codifies case law.[114] Yet, on the matter of immaterial inconsistencies, the Real ID Act's provisions[115] again make a significant departure from case law, United Nations High Commissioner for Refugees guidelines, and Immigration and Naturalization Service guidelines.[116]

The Sections that follow closely examine and analyze each of these provisions. The analysis shows that the law, though potentially detrimental to legitimate asylum seekers, is not necessarily so. If interpreted as suggested below, the damage it may cause to asylum seekers can be minimized.

A. Was It Really Persecution? The Problem of "Mixed Motive" Cases

1. Current Law on Proving that the Harm Suffered Was "on Account of" One of the Five Grounds for Asylum

An applicant for asylum must prove that the harm he or she suffered amounted to persecution. The test for whether harm rises to the level of persecution is threefold. First, the applicant must have suffered harm severe enough to rise to the level of persecution.[117] Second, the harm must

---

The first view is the one adopted by the majority of Courts of Appeals. *See, e.g.*, Lopez-Soto v. Ashcroft, 383 F.3d 228, 236 (4th Cir. 2004); Marku v. Ashcroft, 380 F.3d 982, 988 n.10 (6th Cir. 2004); Girma v. INS, 283 F.3d 664, 667 (5th Cir. 2002); Singh v. Ilchert, 63 F.3d 1501, 1509 (9th Cir. 1995); Osorio v. INS, 18 F.3d 1017, 1028 (2d Cir. 1994).

[112] § 101(a)(3)(B)(ii), 119 Stat. at 303 ("Where the trier of fact determines that the applicant should provide evidence that corroborates otherwise credible testimony, such evidence must be provided unless the applicant does not have the evidence and cannot reasonably obtain the evidence.").

[113] Matter of S-M-J-, 21 I&N Dec. 722, 725 (B.I.A. 1997).

[114] § 101(a)(3)(B)(iii), 119 Stat. at 303.

[115] *Id.* (stating credibility determinations should be made "without regard to whether an inconsistency, inaccuracy, or falsehood goes to the heart of the applicant's claim . . . .").

[116] *See, e.g.*, Singh v. INS, 365 F.3d 1164, 1171 (9th Cir. 2003) ("Minor inconsistencies in the record such as discrepancies in dates which reveal nothing about an asylum applicant's fear for his safety are not an adequate basis for an adverse credibility finding.") (quoting Vilorio-Lopez v. INS, 852 F.2d 1137, 1142 (9th Cir. 1988)).

[117] *See* Mihalev v. Ashcroft, 388 F.3d 722, 730 (9th Cir. 2004) (holding that detention for ten days accompanied by daily beatings and hard labor constitutes persecution, even in

have been committed by a government or an entity that the government is unable or unwilling to control.[118] Third, the harm must have occurred on account of at least one of the five grounds of asylum: race, religion, nationality, political opinion, or membership in a particular social group.[119] The Real ID Act addresses the third component, stating that asylum applicants must prove that one of the five grounds for asylum was or will be "at least one central reason" for the persecution they endured.[120]

Many asylum cases involve "mixed motives," in which persecution may have occurred on account of one or more non-protected grounds, as well as one or more protected grounds. In 1992, the Supreme Court held in *INS v. Elias-Zacarias*, that an applicant must provide "*some* evidence . . . direct or circumstantial" of the persecutor's motive.[121] The Court further specified that establishing asylum eligibility does *not* require "direct proof of [the] persecutors' motives."[122] Moreover, "an applicant does not bear the unreasonable burden of establishing the exact motivation of a 'persecutor' where different reasons for actions are possible."[123]

The importance of these holdings cannot be overstated, because, as discussed more fully in Part III.B, *infra*, persecutors generally do not provide their victims with evidence of, insights into, or discussion about the atrocities they commit. Requiring asylum applicants to prove the mental states of their persecutors to an adjudicator would be an almost impossible burden for applicants to meet.[124] Moreover, according to the Board of Immigration Appeals,[125] "[s]uch a rigorous standard would largely render nugatory the Supreme Court's decision in *INS v. Cardoza-Fonseca*, 480 U.S. 421 (1987), and be inconsistent with the 'well-founded fear' standard embodied in the 'refugee' definition."[126]

---

the absence of serious physical injury); Ouda v. INS, 324 F.3d 445, 454 (6th Cir. 2003) (finding that threats and beatings combined with deprivation of livelihood and ability to leave home amount to persecution); Andriasian v. INS, 180 F.3d 1033, 1042 (9th Cir. 1999) (holding that persistent death threats and assaults on one's family constitute persecution).

[118] 8 U.S.C. § 1101(a)(42)(A) (2000). *See also, e.g.*, Galicia v. Ashcroft, 396 F.3d 446, 448 (1st Cir. 2005); Abdulrahman v. Ashcroft, 330 F.3d 587, 592 (3d Cir. 2003); Llana-Castellon v. INS, 16 F.3d 1093, 1097–98 (10th Cir. 1994).

[119] 8 C.F.R. § 208.13 (b)(2)(A) (2005).

[120] Real ID Act, Pub. L. No. 109-13, Div. B, § 101(a)(3)(B)(i), 119 Stat. 231, 303 (2005).

[121] 502 U.S. 478, 483 (1992).

[122] *Id.*

[123] Shoafera v. INS, 228 F.3d 1070, 1076 (9th Cir. 2000) (quoting Matter of Fuentes, 19 I&N Dec. 658, 662 (B.I.A. 1988)); *see also* Romilus v. Ashcroft, 385 F.3d 1, 7 (1st Cir. 2004) ("[N]or is [the asylum applicant] required to establish [the persecutors'] exact motivations.").

[124] *See, e.g.*, Zubeda v. Ashcroft, 333 F.3d 463, 474 (3d Cir. 2003) ("[R]equiring an alien to establish the specific intent of his/her persecutors could impose insurmountable obstacles. . . .").

[125] Matter of S-P-, 21 I&N Dec. 486, 489–90 (B.I.A. 1996).

[126] *Id. See also* INS v. Cardoza-Fonseca, 480 U.S. 421 (1987) (holding that asylum seekers must prove only a reasonable possibility of persecution in order to establish a well-founded fear).

2. *The Real ID Act: "At Least One Central Reason"*

The Real ID Act states that asylum applicants must prove that "race, religion, nationality, membership in a particular social group, or political opinion was or will be at least one central reason for persecuting the applicant."[127] Although, as discussed below, this language can and should be interpreted as a codification of *Matter of S-P-*, the use of the term "central" creates opportunity for adjudicators to require more proof of causation than *Elias-Zacarias* and *Matter of S-P-* permit.

The Real ID Act does not define "centrality," but it appears to have adopted the term "central" from proposed INS regulations issued in December 2000[128] in which centrality was a major theme. In those regulations, the INS proposed that "[i]n cases involving a persecutor with mixed motivations, the applicant must establish that the applicant's protected characteristic is central to the persecutor's motivation to act against the applicant."[129] In explaining the reasoning behind this proposed regulation, the INS cited two cases which purportedly represented "con-flicting interpretations of the extent to which the persecutor's motivation must relate to a protected characteristic":[130] the Ninth Circuit case *Singh v. Ilchert*[131] and the First Circuit case *Gebremichael v. INS.*[132] In *Singh*, the Ninth Circuit held that "persecutory conduct may have more than one motive, and so long as one motive is one of the statutorily enumerated grounds, the requirements have been satisfied."[133] In its decision in *Gebremichael*, the First Circuit stated that an asylum applicant must prove that one of the five statutory grounds is "at the root of persecution."[134]

Despite the INS's reasoning that these opinions were at odds with each other,[135] a more careful reading of the cases shows that the INS misread the First Circuit case and that no split ever existed. In using the phrase "root of persecution," the First Circuit simply borrowed language from its own prior decision in *Ananeh-Firempong v. INS.*[136] In that case, the court quoted the following language from a United Nations Asylum Handbook:[137]

---

[127] Real ID Act, Pub. L. No. 109-13, Div. B, § 101(a)(3)(B)(i), 119 Stat. 231, 303 (2005).
[128] Asylum and Withholding Definitions, 65 Fed. Reg. 76,588, 76,598 (Dec. 7, 2000).
[129] *Id.* at 76,598.
[130] *Id.* at 76,592.
[131] 63 F.3d 1501 (9th Cir. 1995).
[132] 10 F.3d 28 (1st Cir. 1993).
[133] *Singh*, 63 F.3d at 1509.
[134] *Gebremichael*, 10 F.3d at 35.
[135] *See supra* notes 128–130 and accompanying text.
[136] 766 F.2d 621 (1st Cir. 1985).
[137] OFFICE OF THE UNITED NATIONS HIGH COMMISSIONER FOR REFUGEES, HANDBOOK ON PROCEDURES AND CRITERIA FOR DETERMINING REFUGEE STATUS, U.N. Doc. HCR/IP/4/Eng/Rev.1 (1992) [hereinafter UNHCR HANDBOOK].

> Membership [in] such a particular social group may be at the root of persecution because there is no confidence in the group's loyalty to the Government or because the political outlook, antecedents or economic activity of its members or the very existence of the social group as such, is held to be an obstacle to the Government's policies.[138]

The UNHCR Handbook was used merely to define the term "social group," rather than to address "centrality" of motive. The First Circuit even noted as such in *Ananeh-Firempong*, stating that the *UNHCR Handbook* is a "'useful tool' for interpreting the phrase 'social group' as it appears in the U.N. Protocol . . . ."[139] Thus, the court, in using the language "root of persecution," did not intend to propound a bright-line rule for determining the motivation of a persecutor.[140]

INS's reliance on the *Gebremichael* "root of persecution" formula is even more problematic when one considers that the First Circuit was the only circuit even arguably to use such an analysis when determining the persecutor's motive. The Second, Eighth, and Ninth Circuits were united in the prevailing view that persecution on account of a protected ground is established when the applicant proves that the persecutor's motives were based *in part* on a protected ground.[141]

Moreover, the legislative history of the Real ID Act clearly indicates the legislature's support for the prevailing view among the Courts of Appeals regarding motive. First, the conference committee replaced the phrase "a central reason" in the House of Representatives version of the bill[142] with the phrase "at least one central reason."[143] This substitution emphasizes that more than one motive may prompt a persecutor to cause harm and that the asylum seeker need not prove that the protected ground was foremost in the persecutor's mind.

---

[138] 766 F.2d at 626 (quoting UNHCR HANDBOOK at ¶ 78).

[139] *Id.*

[140] Even if the First Circuit's use of the term "root of persecution" were to be construed as a means of analyzing the motivation of a persecutor, it properly applies only in the context of withholding of removal under 8 U.S.C. § 1253(h) (2000), which carries a substantially higher burden of proof than asylum. *See* INS v. Cardoza-Fonseca, 480 U.S. 421, 430–33 (1987).

[141] *See, e.g.*, Chang v. INS, 119 F.3d 1055, 1065 (3d Cir. 1997) (granting pettition for review where persecutor's motivation was based in part on a protected ground); Singh v. Ilchert, 63 F.3d 1501, 1509 (9th Cir. 1995) (holding that "an applicant is not required to provide direct proof of her persecutor's motives but rather some evidence of it, direct or circumstantial"); Osorio v. INS, 18 F.3d 1017, 1028 (2d Cir. 1994) (holding that persecution on account of a protected ground "does not mean persecution *solely* on account of the victim's political opinion. That is, the conclusion that a cause of persecution is economic does not necessarily imply that there cannot exist other causes of the persecution."). This view was later adopted by most other circuits. *See* Lopez-Soto v. Ashcroft, 383 F.3d 228, 236 (4th Cir. 2004); Marku v. Ashcroft, 380 F.3d 982, 988 n.10 (6th Cir. 2004); Girma v. INS, 283 F.3d 664, 667 (5th Cir. 2002).

[142] H.R. 418, 109th Cong. (1st Sess. 2005).

[143] H.R. Rep. No. 109-72 (2005) (Conf. Rep.).

Second, the Act's drafters intended that "[t]he central reason standard will . . . require aliens who allege persecution because they have been erroneously identified as terrorists *to bear the same burden as all other asylum applicants . . . in accordance with Supreme Court precedent.*"[144] While this distinction did not appear in the legislation itself, it indicates that the rephrased centrality requirement was not intended to impose a *higher* bar for all asylum applicants beyond what the Supreme Court had already enunciated; rather, it was simply imposed to prevent applicants accused of being terrorists from getting a *lower* bar.

The only sensible interpretation of the centrality provision of the Real ID Act, therefore, is that adjudicators should continue to rely on the standard set forth in the Supreme Court's *Elias-Zacarias* decision and the BIA case *Matter of S-P-* for the "nexus" determination. That is, an adjudicator must conclude whether, based on a reasonable interpretation of the facts and the evidence, the persecution occurred at least in part because of one of the five grounds.[145] This analysis also comports with what the INS suggested in the preamble to the 2000 proposed regulations: that nexus is not established "if the protected characteristic was incidental or tangential to the persecutor's motivation."[146] There is no evidence, then, suggesting that drafters intended a heightened nexus requirement for asylum applicants.

3. Analyzing Nexus Under the Real ID Act: "Miguel's Case"

The case of Miguel,[147] an asylum seeker from Colombia, illustrates the analysis set forth above. Miguel owned a small factory in Bogotá, Colombia. He belonged to a small local branch of the national conservative party called the National Salvation Movement (Movimiento de Salvación Nacional, "MSN"). As a factory owner, Miguel sought to employ poor rural youth who, because of their socioeconomic status, are at high risk for recruitment by a guerrilla group called the Armed Revolutionary Forces of Colombia (Fuerzas Armadas Revolucionarias de Colombia, "FARC"). As a member of the MSN, Miguel donated building materials for civic projects in the poverty-stricken rural areas surrounding Bogotá. He also supported political candidates who reflected his conservative, anti-guerrilla values and political viewpoint.

---

[144] H.R. REP. NO. 109-72, at 165 (emphasis added).

[145] *See* INS v. Elias-Zacarias, 502 U.S. 478, 483 (1992); *Matter of S-P-*, 21 I&N Dec. 486, 494 (B.I.A. 1996).

[146] Asylum and Withholding Definitions, 65 Fed. Reg. 76,588, 76,592 (Dec. 7, 2000).

[147] "Miguel" is the pseudonym of an actual client of the Villanova Clinic for Asylum, Refugees, and Emigrant Services (CARES) who, along with his family, was granted asylum. The names of CARES's clients are kept anonymous due to the highly personal nature of asylum claims. Documentation is on file with the author.

In May 1999, a group of armed men robbed Miguel's factory at gunpoint. Miguel was not present, but his wife, sister-in-law, and several employees were. The armed men identified themselves as guerrillas and demanded that the captives turn over their jewelry and cash. They also forced Miguel's wife to turn over the keys to one of the company SUVs. Before leaving, they locked everyone in a room and warned them not to report the incident to the police. After escaping, Miguel's wife promptly filed a police report. Shortly thereafter, Miguel received a menacing letter from a person claiming to be a FARC commander and threatening harm against the family in retaliation for Miguel's wife's reporting the robbery to the police.

From those facts alone, the only indication that the robbery and the threat were politically motivated is that the robbers and letter writer identified themselves as guerrillas. The main purpose for their attack on Miguel's factory seems to have been to obtain cash and a vehicle; the main purpose of the letter seems to have been to retaliate against or intimidate the family for reporting the robbery to the police. Thus, on these facts alone, it would be difficult for Miguel to prove that his membership in an anti-guerrilla political party had much, if anything, to do with the robbery and threatening letter. His political opinion may have in part motivated the guerrillas to steal from him rather than from someone who supports their cause, but there is no evidence of that in this account. These facts would indicate that the motivation on account of a protected ground was tangential or incidental at best, but would not necessarily satisfy the centrality requirement.

Miguel's case, however, contains additional facts that would demonstrate a sufficient nexus between the persecution suffered and one of the five grounds for asylum. Shortly following the robbery, the family began receiving threatening phone calls from a person claiming to be a FARC commander, demanding that Miguel discontinue his "activities." In addition, two men who identified themselves as guerrillas briefly abducted Miguel's eldest daughter at gunpoint and ordered her to tell her father to stop his activities. The robbery, when viewed in conjunction with these additional incidents, takes on new meaning: this family was a specific target of a guerrilla group attempting to curtail Miguel's political activities.

Moreover, Miguel's student representatives provided a number of supporting documents demonstrating that political opinion and membership in a particular social group, rather than financial gain or general violence, motivated the guerrillas. An expert in Miguel's case testified that the guerrillas, who purport to follow strict Marxist tenets, target small business owners because they represent the capitalism and free market economy that Marxists oppose. A State Department Report on Human Rights in Colombia, as well as reports from several respected human rights groups, indicated that FARC guerrillas were responsible for numerous attacks

against individuals whom they perceived as ideologically and politically opposed to FARC's goals.

This evidence, together with the family's testimony, was sufficient to prove centrality under the Real ID Act and allow Miguel to be granted asylum. Although no direct evidence of the persecutor's exact motivation exists, it is clear that the protected grounds—political opinion and membership in a particular social group—were not merely incidental or tangential motivations for the persecution. FARC may have targeted Miguel and his family in part for pecuniary gain, but Miguel's opposition to FARC, manifested in his business and political activities, formed at least part of the motivation for the attacks on Miguel's family.

B. Corroborating Evidence: How Much Is Enough?

1. Corroboration of Persecution Generally

Corroborating asylum claims presents significant challenges, especially in terms of logistics and authentication. Obviously, most asylum seekers will not come to court equipped with notarized affidavits from their persecutors stating, "I, Joe Persecutor, beat and tortured your client on three occasions between December 1999 and August 2003 on account of her political opinion against our oppressive but beloved dictator. Her political opinion was foremost in my mind when this occurred."[148] Moreover, many asylum seekers arrive from countries that lack infrastructure, adequate communication systems, and sometimes even a functioning government.[149] Obtaining documents, even ones as relatively common as a birth certificate or medical report, can therefore involve logistical impediments that often prove insurmountable.

Additionally, persons escaping persecution may leave behind important documents (such as identity cards, birth certificates, medical records, etc.) when fleeing their countries, either in haste or in an attempt to conceal their identities from persecutors.[150] By attempting to obtain the documents later, an asylum seeker risks interception of his or her mail, poten-

---

[148] *See* Bolanos-Hernandez v. INS, 767 F.2d 1277, 1285 (9th Cir. 1984) ("Persecutors are hardly likely to provide their victims with affidavits attesting to their acts of persecution."); *see also generally* Virgil Wiebe et al., *Asking for a Note from Your Torturer: Corroboration and Authentication Requirements in Asylum, Withholding and Torture Convention Cases*, 01-10 IMMIGR. BRIEFINGS 1 (Oct. 2001) (discussing in detail the corroboration requirements for asylum seekers).

[149] *See* United Nations High Commissioner for Refugees, *Refugees: 2004 Year in Review*, REFUGEES MAG., Jan. 1, 2005, at 8–12 (describing conditions in refugee-producing countries around the world), *available at* http://www.unhcr.ch/cgi-bin/texis/vtx/publ/opendoc.pdf?tbl=PUBL&id=41e3a9fc4.

[150] *See* Michele R. Pistone, *The New Asylum Rule: Improved but Still Unfair*, 16 GEO. IMMIGR. L.J. 1, 8 (2001) (explaining that "records may take months or years to compile because refugees usually leave them behind, and the documents may be available only in the country from which the refugee has fled.").

tially exposing family and friends to harassment by the persecuting entity.[151] Even documentation of physical trauma itself can be difficult to obtain, such as in rape cases, with often little if any physical evidence.[152] In many cases, therefore, the more legitimate the persecution, the less likely it is that the asylum seeker will have the required proof. As such, establishment of an accurate but not unduly burdensome corroboration process can be very difficult.

2. *The Courts on Corroboration*

Courts have recognized the unique challenges discussed above. In 1987, the Board of Immigration Appeals decided in *Matter of Mogharrabi* that, due to the difficulty asylum seekers often face in obtaining corroborating evidence, "the applicant's testimony [alone] will suffice if it is credible, detailed and specific."[153] Several Courts of Appeals adopted this reasoning,[154] and it eventually made its way into the Code of Federal Regulations.[155]

Two years later, in *Matter of Dass*, the Board clarified its holding in *Matter of Mogharrabi* and articulated a general rule for corroboration: where corroborating evidence is available, the applicant should present it;

---

[151] *See id.* (stating that "[e]ven if friends or family members can obtain copies of the documents, hostile governments may intercept international mail. Therefore, asylum applicants may hesitate for a long time before asking others to put themselves at risk by requesting corroborating records.").

[152] *See* PHYSICIANS FOR HUMAN RIGHTS, EXAMINING ASYLUM SEEKERS: A HEALTH PROFESSIONAL'S GUIDE TO MEDICAL AND PSYCHOLOGICAL EVALUATIONS OF TORTURE 54–61 (2001). The Guide states:

> In the majority of political asylum applicants who allege sexual assault during torture, the traumatic event(s) will have occurred months or years before the medical examination. Therefore, most individuals will not have physical signs at the time of the examination. . . . Even on examination of the female genitalia immediately after rape there is identifiable damage in less than 50% of cases. Anal examination of males and females after anal rape shows lesions in less than 30% of cases.

[153] 19 I&N Dec. 439, 444 (B.I.A. 1987) (relying on Cardoza-Fonseca v. INS, 767 F.2d 1448, 1458 (9th Cir. 1985)).

[154] *See* Cordon Garcia v. INS, 204 F.3d 985, 992–93 (9th Cir. 2000) (holding that due to "the serious difficulty with which asylum applicants are faced in their attempts to prove persecution . . . this court does not require corroborative evidence" from asylum applicants who have testified credibly); Gumbol v. INS, 815 F.2d 406, 412 (6th Cir. 1987) (citing Youkhanna v. INS, 749 F.2d 360, 362 (6th Cir. 1984)) (holding that an asylum seeker must present some specific facts, *either* through objective evidence or through persuasive credible testimony, to show that his fear of persecution is well-founded); Ganjour v. INS, 796 F.2d 832, 837 (5th Cir. 1986) (holding that an asylum applicant "must present *specific* facts, through objective evidence if possible, or through his or her own persuasive, credible testimony, showing actual persecution or detailing some other good reason to fear persecution . . . ." (quoting Carvajal-Munoz v. INS, 743 F.2d 562, 576 (7th Cir. 1984)).

[155] 8 C.F.R. §§ 208.13(a), 208.16(b), 1208.13(a), 1208.16(b) (1990).

when unavailable, the applicant should explain why.[156] The Board also distinguished between a claim "focused on specific events involving the [applicant] personally" and a "sweeping and general" claim involving an elaborate background context.[157] The Board noted that in order to better establish the credibility of the applicant's story, more corroborating background evidence would be necessary in the latter.[158]

The Board further refined this holding in *Matter of S-M-J-*,[159] clarifying that in cases where corroborating evidence is *reasonably* expected, it *should* be provided.[160] The Board went on to say that if the applicant fails to present such evidence, it "*can* lead to a finding that [the] applicant has failed to meet her burden of proof."[161] However, the Board noted that "specific documentary corroboration of an applicant's particular experiences is not required unless the supporting documentation is of the type that would normally be created or available in the particular country and is accessible to the alien, such as through friends, relatives, or co-workers."[162]

Matter of S-M-J- also provides examples of the types of facts "easily subject to verification"[163] for which adjudicators may reasonably expect corroborating evidence. Those examples include "evidence of [the applicant's] place of birth, media accounts of large demonstrations, evidence of a publicly held office, or documentation of medical treatment."[164]

Importantly, the BIA placed the burden of providing evidence of general country conditions on the adjudicator[165] and the government attorneys,[166] as well as on the asylum applicant. Thus, most asylum cases pre-

---

[156] 20 I&N Dec. 120, 124–25 (B.I.A. 1989).

[157] *Id.* at 125.

[158] *Id.*

[159] 21 I&N Dec. 722 (B.I.A. 1997).

[160] *See id.* at 725 (holding that "[u]nreasonable demands are not placed on an asylum applicant to present evidence to corroborate particular experiences (e.g., corroboration from the persecutor). However, where it is reasonable to expect corroborating evidence for certain alleged facts pertaining to the specifics of an applicant's claim, such evidence should be provided.").

[161] *Id.* at 726 (emphasis added).

[162] *Id. See also* Matter of M-D-, 21 I&N Dec. 1180 (B.I.A. 1998), *rev'd sub nom.* Diallo v. INS, 232 F.3d 279 (2d. Cir. 2000) (upholding the BIA's determination that asylum applicants should provide corroborating evidence when it is available).

[163] 21 I&N Dec. at 725.

[164] *Id.*

[165] *See* Matter of S-M-J-, 21 I&N Dec. at 727 (stating that:

> Although the burden of proof is not on the Immigration Judge, if background information is central to an alien's claim, and the Immigration Judge relies on the country conditions in adjudicating the alien's case, the source of the Immigration Judge's knowledge of the particular country must be made part of the record . . . . Thus, the statute specifically recognizes that the presentation of evidence is a proper function of an Immigration Judge.).

[166] *See id.* (holding that "[a]s a general matter, therefore, we expect the Service to introduce into evidence current country reports, advisory opinions, or other information readily available from the Resource Information Center").

sumably have at least a State Department Report in the record, often supplied by the government. This aspect of *Matter of S-M-J-* is especially relevant in cases involving unrepresented, detained clients with limited or no access to resources or legal aid.[167]

The opinion's admonitions to adjudicators and government attorneys aside, *Matter of S-M-J-* is not particularly sympathetic to the plight of asylum seekers. A document that is easily subject to verification may still not be easily obtainable by an individual fleeing persecution. A refugee from war-torn Somalia who lived for years in a Kenyan refugee camp, for example, is unlikely to have or be able to obtain a document as common and presumably simple as a birth certificate. Moreover, what an adjudicator considers reasonable to provide may differ from a persecution victim's perception of what is reasonable to provide, due to the fears discussed above: interception by government officials, danger to family members, and repercussions for colleagues.

In *Ladha v. INS*,[168] the Ninth Circuit explicitly rejected the corroboration rule of *Matter of S-M-J-*, basing its holding on three lines of Ninth Circuit cases. The first line of cases "emphasizes the difficulty of proving specific threats by persecutors, and emphasizes that credible testimony as to a threat is sufficient to prove that the threat was made."[169] The second line of cases "emphasizes that not only specific threats but also other facts that serve as the basis for an asylum or withholding claim can be shown by credible testimony alone if corroborative evidence is 'unavailable.'"[170] Finally, the third line of cases "make[s] clear that when an alien credibly testifies to certain facts, those facts are deemed true and the question re-

---

[167] *See* Michele R. Pistone, *Justice Delayed Is Justice Denied: A Proposal for Ending the Unnecessary Detention of Asylum Seekers*, 12 HARV. HUM. RTS. J. 197, 219–20 (1999) [hereinafter *Justice Delayed is Justice Denied*]. The author noted:

> Detained asylum seekers usually do not have access to relevant legal or background resource materials. Indeed, until recently, the INS did not have any standards concerning the content of legal libraries at detention facilities. Even though these standards are in place, they do not apply to non-INS facilities such as local, county, and city jails where many asylum seekers are detained; therefore, the contents fall far short of alleviating concerns about the availability of sufficient corroborative materials. The standards do not require libraries to maintain up-to-date information about country conditions . . . . Even with all the relevant legal and country condition resource materials necessary to present a claim for asylum, only the minority of asylum seekers fluent in English are able to use them.

[168] 215 F.3d 889, 899 (9th Cir. 2000).

[169] *Id.* (citing Lopez-Reyes v. INS, 79 F.3d 908, 912 (9th Cir. 1996); Artiga Turcios v. INS, 829 F.2d 729, 723 (9th Cir. 1987); and Bolanos-Hernandez v. INS, 767 F.2d 1277, 1285, 1288 (9th Cir. 1984)).

[170] Ladha, 215 F.3d at 900 (citing Castillo v. INS, 951 F.2d 1117, 1121 (9th Cir. 1991); Limsico v. INS, 951 F.2d 210, 212 (9th Cir. 1991); Estrada-Posadas v. INS, 924 F.2d 916, 918–19 (9th Cir. 1991); Aguilera-Cota v. INS, 914 F.3d 1375, 1378 (9th Cir. 1990); Blanco-Comarribas v. INS, 830 F.2d 1039, 1042–43 (9th Cir. 1987); Del Valle v. INS, 776 F.2d 1407, 1411 (9th Cir. 1985); Cardoza-Fonseca v. INS, 767 F.2d 1448, 1453 (9th Cir. 1985)).

maining to be answered becomes whether these facts, and their reasonable inferences, satisfy the elements of the claim for relief."[171] Based on the reasoning in these three lines of cases, in which the court recognized the difficulties that asylum seekers face in attempting to corroborate their claims, the *Ladha* court held that "[t]o the extent that decisions such as *Matter of S-M-J-* . . . establish a corroboration requirement for credible testimony, they are disapproved."[172]

Ladha's reasoning, though arguably most in line with the spirit of refugee protection, was not favored among other Courts of Appeals.[173] As such, there remains a split between the Circuit Courts of Appeals as to whether *Matter of S-M-J-* is good law.

3. *Corroboration Under the Real ID Act*

The Real ID Act permits an asylum seeker to corroborate his or her claim solely with his or her own testimony, but stipulates that such testimony must be "credible, . . . persuasive, and refer[ ] to specific facts sufficient to demonstrate that the applicant is a refugee."[174] However, even if that testimony is sufficient to establish asylum eligibility, the Real ID Act permits adjudicators to require corroborating evidence: "Where the trier of fact determines that the applicant should provide evidence that corroborates otherwise credible testimony, such evidence must be provided unless the applicant does not have the evidence and cannot reasonably obtain the evidence."[175] Therefore, that clause directly refutes *Ladha's* ruling that no corroboration should be required for credible testimony, and for the most part, codifies *Matter of S-M-J-*.[176]

---

[171] *Ladha*, 215 F.3d at 900–01 (citing Khourassany v. INS, 208 F.3d 1096, 1100 (9th Cir. 2000); Yazitchian v. INS, 207 F.3d 1164, 1168 (9th Cir. 2000); Duarte de Guinac v. INS, 179 F.3d 1156, 1159 (9th Cir. 1999); Del Carmen Molina v. INS, 170 F.3d 1247, 1249 & n.2 (9th Cir. 1999); Campos-Sanchez v. INS, 164 F.3d 448, 451 n.1 (9th Cir. 1999)).

[172] *Ladha*, 215 F.3d at 901.

[173] *See* Gontcharova v. Ashcroft, 384 F.3d 873, 877 (7th Cir. 2004) (declining to follow the Ninth Circuit's rejection of *Matter of S-M-J-* but requiring that immigration judges explain their application of *Matter of S-M-J-*); El-Sheikh v. Ashcroft, 388 F.3d 643, 647 (8th Cir. 2004) (quoting Diallo v. INS, 232 F.3d 279, 287 (2d Cir. 2000) and Abdulai v. Ashcroft, 239 F.3d 542, 554 (3d Cir. 2001)) (declining to adopt the Ninth Circuit's rejection of *Matter of S-M-J-* but adopting the Second and Third Circuits' requirement that immigration judges and the BIA "(1) rule explicitly on the credibility of an applicant's testimony; (2) explain why it was reasonable to expect additional corroboration; and (3) assess the sufficiency of the applicant's explanations for the absence of corroborating evidence"); Dorosh v. Ashcroft, 398 F.3d 379, 382 (6th Cir. 2004) (expressly rejecting the Ninth Circuit's reasoning and adopting that of the Second and Third Circuits with regard to *Matter of S-M-J-*)). *See generally* Brian P. Downey & Angelo A. Stio, III, *Of Course We Believe You, But . . . : The Third Circuit's Position on Corroboration of Credible Testimony*, 48 VILL. L. REV. 1281 (2003).

[174] Real ID Act, Pub. L. No. 109-13, Div. B, § 101(a)(3)(B)(ii), 119 Stat. 231, 303 (2005).

[175] *Id.*

[176] *See* H.R. REP. No. 109-72, at H2870 (2005) (Conf. Rep.).

That the drafters of the Real ID Act chose to eviscerate *Ladha* was unfortunate, but not completely unexpected, particularly when most other Circuit Courts of Appeals backed the *Matter of S-M-J-* ruling. Yet, the Real ID Act's corroboration provisions are particularly troubling in that they differ from even the less desirable approach sanctioned by *Matter of S-M-J-* and the Courts of Appeals that approved of it. Those decisions required that adjudicators act reasonably in requesting additional corroboration and explain their rationale for the request.[177] The Real ID Act's provision, on the other hand, does not hold specifically adjudicators to a standard of reasonableness when determining whether corroboration is necessary or whether the corroboration provided is sufficient.

The only sensible conclusion to draw from the Real ID Act's drafters' failure to include a reasonableness provision is that Congress intended adjudicators to rely on the guidance found in *Matter of Mogharrabi*, *Matter of Dass*, and *Matter of S-M-J-* to determine when and what kind of corroboration should be expected. This is supported by the fact that the Act's sponsors stated in the Conference Report that credibility determinations "must be reasonable and take into consideration the individual circumstances of the specific witness and/or applicant."[178] That language evinces a Congressional intent for the statute to be read in light of the Board's "reasonableness" standard. Moreover, as a practical matter, failure on the part of appellate courts and immigration judges to read a reasonableness requirement into the Act could lead to abuse of discretion, inconsistent application of the law, and the denial of valid asylum claims.[179]

4. Analyzing the Real ID Act's Corroboration Requirement: The Case of Marie and Paul

The case of Marie and Paul,[180] a married couple from Haiti, illustrates how adjudicators should apply the Real ID Act's corroboration requirement. Marie and Paul fled Haiti with their two young daughters in the summer of 2003, during the rule of Jean-Bertand Aristide and his politi-

---

[177] *See* Secaida-Rosales v. INS, 331 F.3d 297, 311 (2d Cir. 2003) (vacating a denial of asylum where, inter alia, the judge's demands for corroboration of testimony were unreasonable); Georgis v. Ashcroft, 328 F.3d 962, 969 (7th Cir. 2003) (same); Senathirajah v. INS, 157 F.3d 210, 216 (3d Cir. 1998) (vacating a denial of withholding where the judge's demands for corroboration were unreasonable in light of the validity of applicant's testimony).

[178] H.R. REP. No. 109-72, at H2870 (2005) (Conf. Rep.).

[179] *See* Qiu v. Ashcroft, 329 F.3d 140, 153–54 (2d Cir. 2003) (holding that "[u]nless the BIA anchors its demands for corroboration to evidence which indicates what the petitioner can reasonably be expected to provide, there is a serious risk that unreasonable demands will inadvertently be made. . . . What is (subjectively) natural to demand may not . . . be (objectively) reasonable.").

[180] "Marie" and "Paul" are pseudonyms of actual clients of the Villanova Clinic for Asylum, Refugee and Emigrant Services (CARES) who were granted asylum. The names of CARES's clients are kept anonymous due to the highly personal nature of asylum claims. Documentation is on file with the author.

cal party, the Fanmi Lavalas ("Lavalas"). Paul's family opposed the Lavalas and he, as well as his brother, uncle, father, and the rest of the family, refused to take part in what they considered a corrupt political process. An armed, pro-Lavalas gang called the Chimeres attacked the family several times for their refusal to support the Lavalas. Still, the family refused to submit. One day, Paul's uncle and brother went missing. Their house had been vandalized and they were never found. Paul's father, fearing for his family members' lives, raised money for Paul to flee Haiti with Marie and the children.

Upon arriving in the United States, Marie, Paul, and the children were placed in a family detention center. They applied for asylum but were detained for the duration of their proceedings. Fortunately, they secured representation through the Villanova Clinic for Asylum, Refugee and Emigrant Services.

Paul's representatives submitted several pieces of evidence in support of his claim. First was Paul's own affidavit. They also submitted a number of newspaper articles, the 2003 State Department Country Report on Human Rights for Haiti, several reports from reputable non-governmental organizations, and the affidavit of an expert on Haiti. All of these materials corroborated Paul's written and oral testimony and supported his claim that he would suffer beatings, torture, or even death if he were to return to Haiti.

Yet other materials that an adjudicator could reasonably expect were noticeably absent, such as affidavits from Paul's family members, including his father, who had arranged for the trip. This is where the case-by-case analysis prescribed in *Matter of S-M-J-* becomes crucial. Marie and Paul were detained and destitute. They came from a poor village, and they, as well as their family members, had limited education. There were no telephones in their village. Paul and his representatives could have tried to send a letter, but even if the mail service had been adequate, the family members were not literate. Yet the general reports of Haitian country conditions—together with Paul's consistent, detailed, and persuasive testimony—provided sufficient corroboration to establish a claim, and the immigration judge granted the family asylum.

Authorizing adjudicators to require corroboration in any form, for any fact, in any case, would serve only to imperil bona fide asylum seekers like Marie and Paul. It is therefore improper under the humanitarian spirit of the 1967 Protocol[181] to impose such a severe and unrealistic burden of corroboration on asylum seekers. Rather, adjudicators should continue to follow the *Matter of S-M-J-* reasonableness test[182] in determining when it is appropriate to require corroboration and what kind of corroboration it is appropriate to require.

---

[181] *See* 1967 Protocol, *supra* note 25, at pmbl.

[182] 21 I&N Dec. 722, 725 (B.I.A. 1997).

C. Credibility: Is the Devil Now in the Details?

Credibility is arguably the most crucial aspect of any asylum case. As discussed above, specific corroboration is difficult, if not impossible, in many cases. An asylum applicant's testimony is often the most probative evidence available. The credibility of that testimony therefore becomes critical.

Courts have endeavored to strike a balance between protecting the asylum system from fraud and accepting that certain minor factors may adversely impact credibility. Current case law stipulates that asylum adjudicators take into account the totality of the circumstances when making a credibility determination, including such factors as demeanor,[183] plausibility,[184] and factual inconsistencies and omissions.[185] In addition, current regulations direct the BIA to apply a "clearly erroneous" standard of review with regard to credibility, thus affording great deference to immigration judges' credibility determinations.[186] Thus, because of the discretion involved in the adjudicator's role and the deference accorded to it, incorrect interpretations of the Real ID Act's credibility provisions may harm legitimate victims of persecution. It is therefore crucial to examine the particular factors that influence credibility determinations so as to assist judges in acting justly within their broad discretion.

---

[183] *See* Cordero-Trejo v. INS, 40 F.3d 482, 487 (1st Cir. 1994) (holding that credibility findings based on demeanor deserve more deference that those based on testimonial analysis); Sarvia-Quintanilla v. INS, 767 F.2d 1387, 1395 (9th Cir. 1985) (holding that an immigration judge is in the unique position to observe the alien's tone and demeanor, to explore inconsistencies in the testimony, and to determine whether the testimony has "the ring of truth"); Kokkinis v. Dist. Dir., 429 F.2d 938, 941–42 (2d Cir. 1970) (holding that "great weight" should be afforded to the findings of the special inquiry officer who conducted the deportation hearing, because, inter alia, he had the opportunity to observe the respondent's demeanor); Matter of V-T-S-, 21 I&N Dec. 792, 796 (B.I.A. 1997) (recognizing the immigration judge's "advantage of observing the alien as he testifies").

[184] *See* Salaam v. INS, 229 F.3d 1234 (9th Cir. 2000) (holding that findings of implausibility cannot be based upon unsupported assumptions); Matter of B-, 21 I&N Dec. 66, 71 (B.I.A. 1995) (holding that consistent, sufficiently detailed, and unembellished testimony may provide a plausible and coherent account of the basis for the fear of persecution, without corroborating evidence); Matter of Dass, 20 I&N Dec. 120, 124 (B.I.A. 1989) (holding that the court is to determine whether the alien's testimony is believable, consistent, and sufficiently detailed to provide a plausible and coherent account of the basis for his alleged fear); *see also* UNHCR HANDBOOK, *supra* note 137, at ¶ 204 (stating that "[t]he applicant's statements must be coherent and plausible, and must not run counter to generally known facts").

[185] *See In re* A-S-, 21 I&N Dec. 1106, 1109 (B.I.A. 1998) (refusing to overturn an immigration judge's adverse credibility determination based on inconsistencies and omissions, because the record revealed that "(1) the discrepancies and omissions described by the Immigration Judge are actually present; (2) these discrepancies and omissions provide specific and cogent reasons to conclude that the alien provided incredible testimony; and (3) a convincing explanation for the discrepancies and omissions had not been supplied by the alien").

[186] 8 C.F.R. § 1003.1(d)(3)(i) (2005).

1. Factors Impacting Credibility

a. Psychological Effects of Trauma

An asylum applicant may have one or more psychological disorders that affect credibility. Post-Traumatic Stress Disorder (PTSD) affects survivors of or witnesses to severely traumatic events, such as rape, murder of loved ones, and physical or psychological torture.[187] Sufferers of PTSD and other trauma-related disorders may have one or more of the following symptoms: flat affect when relaying traumatic events, inability to remember dates or sequences of events correctly, dissociation, and avoidance.[188] All of these symptoms affect the way an adjudicator perceives the veracity of an applicant's statements. An applicant with a flat affect rather than an emotional display while recounting her rape may appear incredible. An applicant who cannot recall the precise date on which he witnessed the massacre of his family may cause an adjudicator to doubt his credibility. Similarly, applicants who respond vaguely to direct questions about crucial elements of their claims may be dissociating or avoiding, but may appear to an adjudicator to be lying.[189] Therefore, those individuals appearing most incredible will, in fact, often have suffered the most serious trauma.

b. Cultural Differences

Cultural differences may also have a strong impact on an adjudicator's perception of an applicant's credibility.[190] Meeting the eyes of an au-

---

[187] AMERICAN PSYCHIATRIC ASS'N, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS 424–25 (4th ed. 1994).

[188] *Id.*; *see also* Memorandum from Jeff Weiss, Acting Dir., Office of Int'l Affairs, to Asylum Officers, Immigration Officers, and Headquarters Coordinators (Dec. 10, 1998), at 14, *available at* http://uscis.gov/graphics/lawsregs/handbook/10a_ChldrnGdlns.pdf (stating that "[q]uestionable demeanor can be the product of culture or trauma rather than a lack of credibility. . . . Symptoms of trauma can include depression, indecisiveness, indifference, poor concentration, long pauses before answering, as well as avoidance or disassociation") [hereinafter Children's Guidelines].

[189] *See* James C. Hathaway & William S. Hicks, *Is There a Subjective Element in the Refugee Convention's Requirement of "Well-Founded Fear?,"* 26 MICH. J. INT'L L. 505, 519–20 (2005) ("Individuals suffering from PTSD may be among the most fearful asylum applicants, yet they are acutely disadvantaged in their ability to communicate that trepidation to decisionmakers."); Michael Kagan, *Is Truth in the Eye of the Beholder? Objective Credibility Assessment in Refugee Status Determination*, 17 GEO. IMMIGR. L.J. 367, 396 (2003) ("The psychological effects of traumatic events can hamper a refugee's ability to communicate why he or she is afraid to return home, and make it difficult for some of the most vulnerable refugees to establish claims that would give them legal protection."); *Justice Delayed Is Justice Denied*, *supra* note 167, at 221 (explaining that the effects of PTSD "can lead to an adverse assessment of the asylum seeker's credibility on the witness stand").

[190] *See* Beate Anna Ort, *International and U.S. Obligations Toward Stowaway Asylum Seekers*, 140 U. PA. L. REV. 285, 309–10 (1991) (explaining how cultural differences with regard to conceptions of family, geography, time, country, common sense, and "verbal behavioral cues" affect credibility determinations in asylum cases); *see also* Kagan, *supra* note 189, at 379 ("The nonverbal cues that people tend to rely on to decide if another person is

thority figure and addressing him or her in a confident tone of voice may be an inconceivable notion to an applicant from a culture in which downcast eyes and soft speech are required to show respect.[191] In the United States and other Western cultures, however, failing to meet an adjudicator's eyes and speak firmly may to some observers indicate evasiveness and deceit.[192] Similarly, a rape victim from a culture in which the victim is often blamed, ostracized or even severely punished for being raped may fail to reveal that she has been raped until significant legal and psychological counseling have taken place. Her failure to do so from the time she first encounters a U.S. immigration official, however, may indicate to an adjudicator that she fabricated the rape, and thus harm her asylum claim.

c. Circumstances of Flight

The circumstances of an asylum applicant's flight from his or her own country may also impact credibility. Upon entering the United States, most asylum applicants encounter uniformed, armed, sometimes curt government officials. Many applicants' experiences with government officials have not been pleasant. Government officials may have been responsible for years of oppression, persecution and killings in applicants' home countries. Some asylum seekers may have had to lie to government officials in the past to save their lives or avoid torture, or to prevent the persecution of friends, family, or colleagues. The impact of such experiences remains upon arrival at American gates. Certain asylum applicants may therefore lie or avoid revealing important details to government officials, especially those whom they meet upon first entering the country.[193] They may also have the same initial mistrust of their lawyers, which may harm their ability to have effective representation. Both issues may harm their asylum claims.

---

telling the truth vary widely from culture to culture."); Children's Guidelines, *supra* note 188, at 14 (providing various examples of cultural differences that asylum officers will likely encounter).

[191] Children's Guidelines, *supra* note 188, at 14 (noting that downcast eyes are a signal of respect to authority in certain Asian cultures).

[192] *See, e.g.*, Balasubramanrim v. INS, 143 F.3d 157, 161 (3d Cir. 1998).

[193] *See* Children's Guidelines, *supra* note 188, at 7 ("Officers must be culturally sensitive to the fact that every asylum applicant is testifying in a foreign environment and may have had experiences which give him or her good reason to distrust persons in authority. . . . A fear of encounters with government officials in countries of origin may carry over to countries of reception.").

2. *Current Law on Credibility*

a. Deference to Adjudicator: "Clearly Erroneous" Standard of Review

Immigration judges' credibility determinations receive a great deal of deference from reviewing courts.[194] The Immigration and Nationality Act authorizes Courts of Appeals to reject an immigration judge's credibility determination only if a "reasonable adjudicator would be compelled" to do so.[195] Similarly, the Board of Immigration Appeals may only overturn an immigration judge's credibility determination if the decision is "clearly erroneous."[196] In practice, so long as an adverse credibility determination is based on more than bare speculation,[197] the Courts of Appeals and the Board will generally uphold it.[198]

b. Inconsistencies and Omissions

An adverse credibility determination based on inconsistencies and omissions will pass the "clearly erroneous" test if based on evidence in the record. In *Matter of A-S-*,[199] the Board of Immigration Appeals set out the criteria for determining whether an adverse credibility determination

---

[194] Credibility determinations based on demeanor receive particular deference because of the immigration judge's opportunity to observe the applicant's testimony. *See* Singh-Kaur v. INS, 183 F.3d 1147, 1149–51 (9th Cir. 1999) (affording great deference to credibility determination based on observation of demeanor); Kokkinis v. District Director, 429 F.2d 938, 941–42 (2d Cir. 1970) (holding that "great weight" should be afforded to the adjudicator who conducted the hearing because he had the opportunity to observe the applicant's demeanor); Matter of V-T-S-, 21 I&N Dec. 792, 796 (B.I.A. 1997) (recognizing that an immigration judge has the advantage of observing an applicant as he or she testifies).

[195] 8 U.S.C. § 1252(b)(4)(B) (2000).

[196] 8 C.F.R. § 1003.1(d)(3)(i) (2005).

[197] *See, e.g.*, Dia v. Ashcroft, 353 F.3d 228, 249 (3d Cir. 2003) (en banc) (overturning an immigration judge's credibility determination based on "speculation and conjecture"); Unase v. Ashcroft, 349 F.3d 1039, 1042 (7th Cir. 2003) (finding that an immigration judge's adverse credibility determination was unsupported by the record when the immigration judge relied on speculation and tenuous logic).

[198] *See* Kalitani v. Ashcroft, 340 F.3d 1, 4–5 (1st Cir. 2003) (upholding an immigration judge's credibility determination based upon discrepancies in the applicant's testimony regarding who procured the documents allowing her to enter the United States, inconsistencies regarding the applicant's identity, and perceived implausibility in the applicant's account); Wu Biao Chen v. INS, 344 F.3d 272, 274–75 (2d Cir. 2003) (upholding an adverse credibility determination based on the applicant's hesitant and unconvincing testimony as well as several inconsistencies in his testimony); Krouchevski v. Ashcroft, 344 F.3d 670, 673 (7th Cir. 2003) (finding that an applicant's assertions that the inconsistencies present in his testimony were the result of translation errors and misunderstandings were insufficient to overcome the "clearly erroneous" standard of review); Matter of R-S-H-, 23 I&N Dec. 629, 637 (B.I.A. 2003) (upholding an immigration judge's adverse credibility finding based on the "clearly erroneous" standard).

[199] 21 I&N Dec. 1106, 1110 (B.I.A. 1998) (noting that an individual fleeing persecution may have difficulty "remembering exact dates when testifying before an immigration judge").

based on inconsistencies and omissions is supported by the record. First, the discrepancies and omissions must actually be present in the record.[200] Second, the discrepancies and omissions must provide specific and cogent reasons to conclude that the applicant provided incredible testimony.[201] Finally, the applicant must have had an opportunity to explain the discrepancies and omissions and must have failed to do so.[202]

Most Courts of Appeals have held that discrepancies and omissions that do not go to the heart of the claim are not an appropriate basis for an adverse credibility determination.[203] In *Uwase v. Ashcroft*,[204] for example, the immigration judge had denied asylum because the applicant, a survivor of the 1994 Rwandan genocide, provided inconsistent testimony regarding her means of supporting herself in the United States. The applicant's student visa paperwork indicated that financial support for her stay in the United States would come from Rwanda; her testimony in court, however, was that her father's friend in the United States was supporting her.[205] The Seventh Circuit held that, because the testimony regarding the facts of her asylum claim was credible, and that the testimony regarding her means of support did not go to the heart of her claim, the Board of Immigration Appeals was incorrect to sustain the immigration judge's negative credibility determination.[206] Requiring that inconsistencies go to the claim's core is an important safeguard against unjust denials of asylum.

c. Airport Statements

Current regulations permit asylum adjudicators to take into account "any oral or written statement that is material and relevant to any issue in the case previously made by the respondent or any other person during any investigation, examination, hearing, or trial."[207] These include statements that U.S. Immigration and Customs Enforcement and U.S. Customs and Border Protection officers elicit from asylum applicants whom they apprehend attempting to enter the United States without proper documenta-

---

[200] *Id.* at 1109.

[201] *Id.*

[202] *Id.*

[203] *See* Kondakova v. Ashcroft, 383 F.3d 792, 796 (8th Cir. 2004), *cert. denied*, 125 S.Ct. 894 (2005) ("While minor inconsistencies and omissions will not support an adverse credibility determination, inconsistencies or omissions that relate to the basis of persecution are not minor but are at the heart of the asylum claim."); Singh v. INS, 365 F.3d 1164, 1171 (9th Cir. 2003) ("Minor inconsistencies in the record such as discrepancies in dates which reveal nothing about an asylum applicant's fear for his safety are not an adequate basis for an adverse credibility finding.") (quoting Vilorio-Lopez v. INS, 852 F.2d 1137, 1142 (9th Cir. 1988)); *see also* Leia v. Ashcroft, 393 F.3d 427, 436 (3d Cir. 2005); Sylla v. INS, 388 F.3d 924, 926 (6th Cir. 2004); Capric v. Ashcroft, 355 F.3d 1075 (7th Cir. 2004).

[204] 349 F.3d 1039 (7th Cir. 2003).

[205] *Id.* at 1042–43.

[206] *Id.*

[207] 8 C.F.R. § 1240.7(a) (2005).

tion. As discussed in Section b, *supra*, such statements may contain discrepancies and omissions that become evident after an asylum applicant has submitted an asylum application and provided oral testimony.

Some courts have recognized that, due to trauma and cultural differences, statements taken upon an asylum applicant's entry into the United States may not be the most reliable indicators of credibility. In *Balasubramanrim v. INS*,[208] the Third Circuit held that the immigration judge and the Board of Immigration Appeals placed undue weight on a statement provided by a Sri Lankan asylum applicant to immigration officers upon his arrival at JFK Airport in New York.[209] The interview took place without a translator and the only record of the interview consisted of twenty-five handwritten questions and answers.[210] During the interview, Mr. Balasubramanrim stated that he had been arrested by a Sri Lankan rebel group and detained for ten days.[211] He did not mention any other arrest or mistreatment. In his asylum application, he listed eight incidents of arrest, several of which entailed physical violence and torture.[212] The immigration judge and Board denied Balasubramanrim's case, finding that the discrepancies between his airport statement and his subsequent written and oral testimony rendered him incredible. The Third Circuit reversed, holding that discrepancies in airport statements alone are an inappropriate basis for an adverse credibility determination.[213]

As such, current case law on credibility suggests that a judge's application denial, if based on inconsistencies and discrepancies "going to the heart of the claim," will likely not be reversed as "clearly erroneous" unless those inconsistencies arose from an airport statement.

3. The Real ID Act's Impact on Credibility Determinations

The Real ID Act codifies the long-established prescription that adjudicators weigh the totality of the circumstances when making credibility determinations.[214] Yet, the Real ID Act departs from established case law regarding whether adjudicators should take into account minor inconsistencies and omissions by stating that immigration judges may base a credibility determination on, *inter alia*, inconsistencies, inaccuracies, or falsehoods "without regard to whether an inconsistency, inaccuracy, or falsehood goes to the heart of the applicant's claim . . . ."[215] As discussed in Section 2.b, *supra*, most circuits have held differently, as do as the INS

---

[208] 143 F.3d 157 (3d Cir. 1998).
[209] *Id.* at 159–60.
[210] *Id.* at 162.
[211] *Id.*
[212] *Id.* at 159.
[213] *Id.* at 164.
[214] Real ID Act, Pub. L. No. 109-13, Div. B, § 101(a)(3)(B)(iii), 119 Stat. 231, 302–23 (2005).
[215] *Id.*

Guidelines, which state that "[m]inor inconsistencies, misrepresentations, or concealment in a claim should not lead to a finding of incredibility where the inconsistency, misrepresentation or concealment is not material to the claim."[216]

This provision of the Real ID Act, however, is not a license to base a negative credibility finding in whole or in any significant part upon inconsistencies regarding immaterial facts. It merely permits immaterial inconsistencies to be considered as part of the totality of the circumstances.[217] This is clear because the conference committee expressly rejected language in the House of Representatives version of the bill that would have allowed adjudicators to dispense with a reasoned totality of the circumstances analysis and make negative credibility determinations based on "*any such factor*, including . . . any inaccuracies or falsehoods . . . without regard to whether an inconsistency, inaccuracy, or falsehood goes to the heart of the applicant's claim."[218]

Therefore, the Real ID Act specifies that immaterial discrepancies should be factored in, but conclusively denies them controlling weight. Because the Real ID Act is otherwise silent on the weight of such factors, adjudicators should look to established case law, such as *Balasubramanrim* and *In re A-S-*, as well as the UNHCR Handbook, in deciding how much weight to give to discrepancies that do not go to the heart of the claim.[219]

That being said, legislators should seriously consider repealing this portion of the Real ID Act. Permitting adjudicators to take into account minor inconsistencies that do not go to the heart of the claim will not prevent terrorism any better than the pre-Real ID system did. This is in part because the asylum system simply did not need strengthening.[220]

One might argue that the absence of signs of successful abuse does not mean the system is perfect. Yet, even if the asylum system could be strengthened, the focus on minor inconsistencies would be an ineffective

---

[216] *INS Supplementary Refugee/Asylum Adjudication Guidelines, reprinted in* 67 INTERPRETER RELEASES, 101–03 (Jan. 22, 1990).

[217] *See* § 101(a)(3)(B)(iii), 119 Stat. at 303–04 (listing the numerous factors that adjudicators may consider while employing a "totality of the circumstances" analysis).

[218] H.R. 418, 109th Cong. § 101(a)(3)(B)(iii) (2005) (emphasis added).

[219] *See* 151 CONG. REC. S4816, 4838 (daily ed. May 10, 2005) (statement of Sen. Brownback (R-Kan.)) (clarifying that in applying the Real ID Act, "[i]t would not be reasonable to find a lack of credibility based on inconsistencies, inaccuracies or falsehoods that do not go to the heart of the asylum claim without other evidence that the asylum applicant is attempting to deceive the trier of fact").

[220] *See supra* notes 20–21 and accompanying text (discussing the failure of numerous terrorists to gain entry through the asylum process even under the more lax system in place before the pre-1996 reforms). *See also* 151 CONG. REC. H453, 468 (daily ed. Feb. 9, 2005) (statement of Rep. Zoe Lofgren (D-Cal.)) ("[W]e have heard references to those who came prior to the first World Trade Center bombing. We made changes in the law subsequent to that. That fix has already been done. We do not need to do what is before us today."); 151 CONG. REC. S4816, 4838 (daily ed. May 10, 2005) (statement of Sen. Brownback (R-Kan.)) ("[The Real ID Act's] language was added based on a claim that our asylum system can be used by terrorists to enter the country. This is not the case.").

means of doing so; a terrorist bent on gaining access to lawful immigration status will likely not make mistakes in his or her asylum proceedings, but rather will be well-rehearsed and thoroughly coached.[221] The focus on minor inconsistencies very well may, however, prevent bona fide asylum seekers from accessing the protection to which they are entitled by law and treaty. A bona fide asylum seeker, particularly one who has survived torture or other trauma, often will encounter or trigger the credibility issues discussed above. As a result, the provision will hurt bona fide asylum seekers without helping national security.

V. IMPLICATIONS FOR SPECIAL CASES UNDER THE REAL ID ACT: GENDER ASYLUM AND CHILDREN'S ASYLUM CLAIMS

The carelessness of the drafters of the Real ID Act is most evident with respect to gender-based asylum claims and the asylum claims of children. Women and especially children tend to be more at risk for persecution than adult males.[222] Moreover, given that U.S. law enforcement officials focused their immediate post–September 11 investigative measures largely on males between the ages of eighteen and thirty-five, it is clear that the government recognizes that women and children are also the least likely to present a security risk to the United States.[223] Yet, the drafters of the Real ID Act failed to distinguish gender-based and children's asylum claims from general asylum claims. Although a full analysis of gender-based and children's claims under the Real ID Act is beyond the scope of this Article, it is important to discuss briefly the potential impact that the Real ID Act may have on those claims.

A. Gender Claims

Gender claims generally involve women fleeing persecution stemming from cultural, religious, and social subjugation and subordination. Examples include women who refuse to conform to their countries' strict interpretation of Islamic law;[224] women who have been subjected to or are

---

[221] *See generally* 9/11 AND TERRORIST TRAVEL, *supra* note 11 (detailing the modus operandi of foreign terrorists who have used subterfuge to enter and remain in the United States).

[222] *See* Jonathan Todres, *Women's Rights and Children's Rights: A Partnership with Benefits for Both*, 10 CARDOZO WOMEN'S L.J. 603, 605 (2004) (noting that "political obstacles, such as not having the right to vote, and developmental issues, such as the more limited verbal skills of younger children, make children more susceptible to exploitation"). Professor Todres also identifies a number of practices to which women are vulnerable due to their gender, such as "domestic violence, incest, rape, trafficking and forced prostitution, child marriages, dowry-related violence, and female genital mutilation." *Id.* at 606.

[223] *See* Akram & Karmely, *supra* note 14, at 629–32 (discussing several anti-terrorism measures that targeted almost exclusively Muslim and Arab males ages eighteen and older).

[224] *See, e.g.*, Fatin v. INS, 12 F.3d 1233 (3d Cir. 1993). Parastoo Fatin fled Iran and

in danger of being subjected to female genital mutilation;[225] and women fleeing life-threatening domestic violence.[226] Gender claims have had mixed results in the courts; the law, particularly with regard to claims based on domestic violence, is in a state of transition.[227]

The most challenging issue for gender-based asylum claims is proving that the motivation of the persecutor was one of the five grounds for asylum.[228] This is particularly true for claims based on domestic violence, in which adjudicators often find that the persecution occurred at the hands of a single private actor on account of motivations not supported by the asylum statute.[229] Thus, one might assume that the Real ID Act's centrality provision would affect gender claims most. As discussed above, however, the centrality provision, although its language derives from the proposed gender regulations, does not much alter the extent to which an asylum applicant must prove the motivation of his or her persecutor. Therefore,

---

sought asylum because of her beliefs in equal rights for women and her refusal to wear a veil as prescribed by Iranian law. Because of her views, Ms. Fatin faced a jail sentence or public whipping or stoning. *Id.* at 1236.

[225] *See, e.g.*, Matter of Kasinga, 21 I&N Dec. 357 (B.I.A. 1996). Fauziya Kasinga fled from Togo to the United States to avoid being forced to submit to her tribe's ritual of female genital mutilation. *Id.* at 358. The BIA overturned an immigration judge's denial of asylum, finding that Ms. Kasinga belonged to a particular social group and qualified for asylum, even if the persecutors lacked malignant intent. *Id.* at 365.

[226] *See, e.g.*, Matter of R-A-, 22 I&N Dec. 906 (B.I.A. 1999). The applicant in *Matter of R-A-* was a Guatemalan woman named Rodi Alvarado, whose husband subjected her to severe abuse between 1984 and 1994. Her husband raped and sodomized her, "broke windows and mirrors with her head, dislocated her jaw," pistol-whipped her, terrorized her with a machete, and kicked her violently in the spine while she was pregnant. Ms. Alvarado repeatedly attempted to flee her husband and to obtain protection, to no avail. The authorities "refused to intervene because it was a 'domestic matter.'" Ms. Alvarado's husband threatened to "cut off her arms and legs, and . . . leave her in a wheelchair, if she ever tried to leave him." *Id.* at 908–10. *See also* Center for Gender and Refugee Studies, *Rodi Alvarado's Story*, http://sierra.uchastings.edu/cgrs/campaigns/update.php (last visited Nov. 19, 2005).

[227] *See generally* Leslye E. Orloff & Janice v. Kaguyutan, *Offering a Helping Hand: Legal Protections for Battered Immigrant Women: A History of Legislative Responses*, 10 AM. U.J. GENDER SOC. POL'Y & L. 95 (2001); ELIZABETH M. SCHNEIDER, BATTERED WOMEN AND FEMINIST LAWMAKING (2000). *See also, e.g.*, Matter of R-A-, 22 I&N Dec. at 918 (reversing the immigration judge's grant of asylum and holding that the social group consisting of "Guatemalan women who have been involved intimately with Guatemalan male companions who believe that women are to live under male domination" was not "recognized and understood to be a societal faction or otherwise a recognizable segment of the population"). As discussed in Part IV.A, *supra*, Attorney General Janet Reno issued proposed regulations recognizing gender-related persecution claims in 2000, but they have never become final. 65 Fed. Reg. 76,588 (proposed Dec. 7, 2000).

[228] *See* Victoria Neilson, *Homosexual or Female? Applying Gender-Based Asylum Jurisprudence to Lesbian Asylum Claims*, 16 STAN. L. & POL'Y REV. 417, 422–25 (2005) (discussing the problems women face in convincing courts that their persecution was based on a protected ground and, in particular, membership in a particular social group).

[229] *See* Laura S. Adams, *Fleeing the Family: A Domestic Violence Victim's Particular Social Group*, 49 LOY. L. REV. 287, 287 (2003) ("[O]ne of the primary arguments against granting refugee status to domestic violence victims is that domestic violence is private in nature and therefore is not the type of politically motivated harm entitled to international protection under refugee law.").

the Real ID Act's centrality provision appears neither to help nor to harm gender-based asylum applicants.

The provision regarding credibility, however, may prejudice gender claims more than it harms non-gender claims. As discussed in Part IV.C, *supra*, cultural issues, past experiences with government officials, and trauma-related psychological ailments have a tremendous impact on credibility determinations. This impact is especially significant in many gender claims, in which cultural stigmas associated with disclosing incidents of rape, female genital mutilation, and domestic violence prevent many bona fide asylum applicants from being forthright about the persecution they suffered.[230] Usually, victims of gender-based persecution must first receive psychological counseling and establish trusting relationships with their legal counsel before revealing their experiences fully.[231] For many victims, however, particularly those in immigration detention, neither psychological nor legal counsel is available. Thus, the Real ID's credibility provisions may result in the inappropriate denial of gender-based asylum claims.

The case of a woman from Albania provides an example of a victim of gender-based persecution for whom cultural differences had devastating consequences.[232] The twenty-nine-year-old woman, whose name has been kept confidential per her request, was married at the age of sixteen by arrangement of her family. She lived with her husband's family in the mountains of northern Albania, kept house, and tended livestock. When the Albanian communist regime ended, her husband took a job with the new authoritarian government. Eventually, southern Albanians rebelled against that regime. The woman's husband, a government worker, fled into the mountains after several groups of armed supporters of the regime tried forcibly to recruit him to fight against the southern rebels. Shortly after he disappeared, a group of armed, masked men came to the house, separated the woman from the rest of the family, and gang-raped her.

Four days later, the woman fled to the United States with a false passport. U.S. officials took her into custody and placed her in a detention center. She received a credible fear interview with a female asylum officer, but the interpreter was an Albanian male and she was too embarrassed to describe the rape in front of him. The asylum officer found that the woman's fear was not credible, and the woman appealed the decision to an immi-

---

[230] *See* LAWYERS COMMITTEE FOR HUMAN RIGHTS, REFUGEE WOMEN AT RISK: UNFAIR U.S. LAWS HURT ASYLUM SEEKERS 6 (2002).

[231] *Id.* at 3–4.

[232] Celia W. Dugger, *In New Deportation Process, No Time, or Room, for Error*, N.Y. TIMES, Sept. 20, 1997, at A1; Celia W. Dugger, *Albanian Seeking Asylum Is Allowed to Return to U.S.*, N.Y. TIMES, Jan. 14, 1998 at B5; *see also* Human Rights First, formerly Lawyers Committee for Human Rights, *Is This America? The Denial of Due Process to Asylum Seekers in the United States* (Oct. 2000), *available at* http://www.humanrights-first.org/refugees/reports/due_process/due_process.htm (profiling the case of the Albanian woman).

gration judge. During the hearing, the woman revealed that she had been raped. The immigration judge did not believe her, because she had not mentioned the rape in her credible fear interview. The woman was deported to Albania but could not return to her village and family because of the perceived shame and loss of honor resulting from the rape. Fortunately, a reporter from the *New York Times* gave her case media attention, and her lawyers located a witness who could corroborate a portion of her claim. In a highly unusual turn of events, the U.S. government had her returned to the United States for a full asylum hearing.[233]

As this example shows, adjudicators of gender-based cases must be vigilant in deciding how much weight, if any, to give minor discrepancies and omissions, especially when those discrepancies and omissions are based on airport interviews. The risk, then, is that abuse of the Real ID Act's more stringent credibility provision will greatly increase the number of victims of gender-based persecution who, like the woman from Albania, do not receive the refugee protection they deserve. The odds are that diligent reporters will not search for and fight for the return of those who are wrongfully deported due to officers' unreasonable expectations of immediate and total candor.

B. Children's Claims

Child asylum applicants stand to lose the most from the passage of the Real ID Act if adjudicators do not interpret it appropriately. The Real ID's Act's corroboration requirements and credibility criteria, if not applied more liberally to child asylum applicants, could result in many bona fide and eligible child asylum applicants having their claims denied. Such a result would be contrary to the United States' obligations under the 1967 Protocol[234] and the International Covenant on Civil and Political Rights.[235] It would also be far astray of the ideals espoused in the Convention on the Rights of the Child.[236] Fortunately, the INS issued "Guide-

---

[233] Dugger, *Albanian Seeking Asylum Is Allowed to Return to U.S.*, *supra* note 232.

[234] *See supra* note 25 and accompanying text.

[235] International Covenant on Civil and Political Rights, Dec. 16, 1966, 999 UNTS 171 (entered into force Mar. 23, 1976). Parties to the Covenant agree that "[e]very child shall have, without any discrimination as to race, colour, sex, language, religion, national or social origin, property or birth, the right to such measures of protection as are required by his status as a minor, on the part of his family, society and the State." *Id.* art. 24(1).

[236] G.A. Res. 44/25, U.N. Doc. A/RES/44/25 (Dec. 5, 1989) (entered into force Sept. 2, 1990). The Convention on the Rights of the Child obligates parties to "take appropriate measures to ensure that a child who is seeking refugee status or who is considered a refugee in accordance with applicable international or domestic law and procedures shall, whether unaccompanied or accompanied by his or her parents or by any other person, receive appropriate protection and humanitarian assistance." *Id.* art. 22(1). However, "because the U.S. has signed but not ratified the CRC, its provisions . . . provide guidance only and are not binding on adjudicators. Having signed the CRC, however, the United States is obligated under international treaty law to refrain from acts which would defeat the object and purpose of the Convention." Children's Guidelines, *supra* note 188, at 2.

lines for Children's Asylum Claims"[237] that provide comprehensive guidance on adjudicating children's claims. These guidelines can assist adjudicators in reading the Real ID Act consistently with the aforementioned international obligations. If adjudicators adhere to those guidelines faithfully, impact on children's claims should be minimal.

1. The INS Guidelines for Children's Asylum Claims

The Children's Guidelines provide guidance to asylum officers on procedural matters, including interviewing techniques, factors affecting credibility,[238] and cross-cultural skills. The Children's Guidelines also address in great detail the legal analysis of children's asylum claims, including determining whether harm rises to the level of persecution, establishing nexus, and requiring corroborating evidence. The Children's Guidelines emphasize that asylum officers must have special consideration for child asylum applicants in how they conduct the asylum interview,[239] what expectations to have regarding testimony and corroboration, and how they evaluate the asylum claim.[240] In developing this guidance, the INS relied on several international documents, including the Universal Declaration of Human Rights,[241] United Nations Executive Committee of the Office of the United Nations High Commissioner for Refugees conclusions,[242] UNHCR policies on refugee children,[243] and Canadian children's guidelines issued

---

[237] Children's Guidelines, *supra* note 188.

[238] *See id.* at 9 (cautioning asylum officers to be vigilant for nonverbal indications of confusion and discomfort, such as "a puzzled look, knitted eyebrows, downcast eyes, long pauses and irrelevant responses").

[239] *See id.* at 7–9 (providing detailed suggestions for establishing a child-friendly atmosphere and building rapport with a child asylum applicant).

[240] *See id.* at 14–27 (outlining the most appropriate factors and circumstances to consider and ask about).

[241] Children's Guidelines, *supra* note 188, at 2 (citing Universal Declaration of Human Rights, G.A. Res. 217A, at 71, U.N. GAOR, 3d Sess., 1st plen. mtg., U.N. Doc A/810 (Dec. 12, 1948)).

[242] Children's Guidelines, *supra* note 188, at 3 (citing UNHCR, Executive Committee of the UNHCR Programme, *Conclusions on the International Protection of Refugees*, No. 47 (XXXVIII) on Refugee Children (1987) (condemning the exposure of refugee children to physical violence and human rights violations and calling for international action to assist the child victims); UNHCR, Executive Committee of the UNHCR Programme, *Conclusions on the International Protection of Refugees*, No. 59 (XL) on Refugee Children (1989) (providing examples of how the needs of refugee children could be assessed, monitored, and met); and UNHCR, Executive Committee of the UNHCR Programme, *Conclusions on the International Protection of Refugees*, No. 84 (XLVIII) on Refugee Children and Adolescents (1997) ("reaffirming the 'best interests of the child' principle")).

[243] Children's Guidelines, *supra* note 188, at 3–4 (citing UNHCR, Sub-Comm. of the Whole on Int'l Prot., *Policy on Refugee Children*, U.N. Doc. EC/SCP/82 (Aug. 6, 1993) (noting that refugee children have different needs and potentials than adults due to age-related developmental differences); UNHCR, REFUGEE CHILDREN: GUIDELINES ON PROTECTION AND CARE (1994) (incorporating international norms relevant to the protection and care of refugee children); and UNHCR, GUIDELINES ON POLICIES AND PROCEDURES IN DEALING WITH UNACCOMPANIED CHILDREN SEEKING ASYLUM (Feb. 1997) (emphasizing the unique needs of unaccompanied refugee children)).

in 1996.[244] The INS also consulted with UNHCR, U.S. nongovernmental organizations, and asylum and refugee experts while formulating the Children's Guidelines.[245] Therefore, they should be entitled to substantial weight in asylum proceedings.[246]

2. *Centrality, Corroboration, and Credibility Under the Children's Guidelines*

The Children's Guidelines provide concrete guidance to judges on how to determine whether the persecutor of a child applicant was motivated by a protected ground. The Guidelines point out that "a child may express fear or have experienced harm without understanding the persecutor's intent."[247] With regard to mixed motive cases, the Guidelines caution that "the child may be unable to identify all relevant motives, but a nexus can still be found if the objective circumstances support the child's claim that the persecutor targeted the child based on one of the protected grounds."[248] This comports with the Real ID Act's mandate that an applicant prove that a protected ground was at least one central reason for the persecution.[249]

The Children's Guidelines are also essential to the proper evaluation of a child's corroboration requirements under the Real ID Act. As discussed above, the drafters of the Real ID Act failed to impose an explicit reasonableness requirement on adjudicators with regard to when and how much corroborative evidence to require from asylum applicants.[250] Nonetheless, adjudicators should follow the Children's Guidelines suggestions on how to determine reasonably whether a child asylum applicant must corroborate his or her testimony. First, the Guidelines note that children "may lack the necessary documents to establish their race, nationality, or religion, and may have more limited access to these documents than a similarly situated adult. . . ."[251] Second, the Guidelines remind adjudicators that credible, consistent, and sufficiently detailed testimony may obviate the need for corroboration, and that in determining whether the child's testimony is sufficient to meet this standard, adjudicators should take into account the child's age, maturity level, and emotional state.[252] Third, the

---

[244] Children's Guidelines, *supra* note 188, at 4 (citing CANADIAN IMMIGRATION AND REFUGEE BOARD, CHILD REFUGEE CLAIMANTS: PROCEDURAL AND EVIDENTIARY ISSUES (1996)).

[245] Children's Guidelines, *supra* note 188.

[246] *See* Chevron U.S.A., Inc. v. Natural Resources Defense Council, Inc., 467 U.S. 837, 844 (1984) (stating that "considerable weight should be accorded to an executive department's construction of a statutory scheme it is entrusted to administer, and the principle of deference to administrative interpretations.").

[247] *Id.* at 21.

[248] *Id.*

[249] *See supra* Part IV.A (analyzing the centrality provision of the Real ID Act).

[250] *See supra* Part IV.B (analyzing the corroboration requirements of the Real ID Act).

[251] Children's Guidelines, *supra* note 188, at 22.

[252] *Id.* at 26.

Guidelines instruct adjudicators to take into account whether the child has legal counsel and whether the child has been in contact with his or her family in determining whether to require corroboration.[253]

Finally, the Guidelines provide essential guidance on determining whether a child asylum applicant is credible. The Guidelines advise that numerous indicators of unreliability may actually be the result of cultural differences, distrust of authority figures, and trauma.[254] If asylum officers encounter vagueness, inconsistencies, inappropriate laughter, or hesitation, they must not assume unreliability but rather "remember the possible developmental or cultural reasons" that may have caused it.[255] The Guidelines also warn that children are more susceptible to coercion by adults to tell a fabricated story, and advise adjudicators to explore the claim in depth, should a child begin to tell a seemingly fabricated story.[256] These considerations continue to apply under the Real ID Act, which, as discussed in Part IV.C.3, *supra*, expressly adopted a totality of the circumstances test for determining credibility. Therefore, in several areas, the Children's Guidelines offer pertinent and considered advice on the adjudication of children's asylum claims.

VI. CONCLUSION

This Article has examined the recently enacted Real ID Act and has proposed interpretations of its three major asylum provisions. The analysis reveals that the Real ID Act is a codification—albeit a vague and poorly drafted one—of existing case law, regulations, and agency guidance. Despite the assertions of the Real ID Act's supporters that the legislation is designed to repair a broken asylum system, the Real ID Act makes very few substantive changes to asylum law. Moreover, the changes that it does make will serve only to weaken the asylum system.

Applicants for asylum have always had the burden of proof that the Real ID Act reiterates: to prove that the persecution they suffered was on account of one of the five grounds for asylum; to provide corroborating evidence when it would be reasonable to expect it; and to establish credibility. The Real ID Act creates potentially dangerous ambiguity in these crucial areas of asylum law. That it does so in the name of protecting the United States from terrorists demonstrates a profound ignorance of the current asylum system and the protections against fraud and abuse already built into it.

Overall, the Real ID Act creates more problems and confusion than it purports to alleviate. Even more unfortunate is that it does so against a

---

[253] *Id.* at 27.

[254] *See supra* Part IV.C.1 (discussing factors that may affect an adjudicator's perception of an asylum applicant's credibility).

[255] Children's Guidelines, *supra* note 188, at 14.

[256] *Id.*

background of misconceptions about the asylum system and of misplaced wariness about asylum seekers. Fortunately, existing case law, international and agency guidelines, and simple common sense still have a vital role in asylum adjudications. Asylum adjudicators, as executors of the asylum laws and the treaties that generated them, have a duty to rely on these guiding principles in applying the Real ID Act to the claims that come before them. If they fulfill that duty, the damage that the Real ID Act does to the asylum system need not be extensive.

ARTICLE

RFRA, CHURCHES AND THE IRS: RECONSIDERING THE LEGAL BOUNDARIES OF CHURCH ACTIVITY IN THE POLITICAL SPHERE

CHRIS KEMMITT*

The author argues that the Internal Revenue Code section governing the tax-exempt status of religious organizations infringes upon the free exercise rights guaranteed by the Religious Freedom Restoration Act. By threatening to revoke the tax-exempt status of any religious group that engages in partisan political activity, the government has effectively penalized churches for their expression of religiously held beliefs. In an attempt to overcome this problem, the author proposes a new, bright-line standard for the regulation of political activity by religious groups. Under the new standard, groups would be permitted to engage in partisan behavior but forbidden from using tax-exempt funds to support such activities. Adopting the new standard, the author concludes, would have the dual effect of facilitating IRS enforcement of the provision and avoiding the infringement of churches' free exercise rights.

Churches, along with other charitable organizations, have long been exempt from federal taxes.[1] However, this exemption is not automatic, and is predicated upon churches' abstention from certain activities within the political sphere.[2] Most notably, the Internal Revenue Code ("I.R.C.") prohibits churches from involvement in any activity that could be construed as supporting a political candidate's campaign.[3] Churches that violate this provision are subject to the revocation of their tax-exempt status: church income is no longer exempt from taxation and individuals may no longer deduct their charitable contributions to those churches.[4]

Despite the harsh potential penalties, allegations of improper political actions by churches abounded during the 2004 presidential campaign.[5] Numerous churches were accused of providing assistance both to Senator Kerry's and President Bush's campaigns, and both campaigns actively

* J.D. Yale Law School, 2005. Law Clerk to the Honorable Judge Nancy Gertner, United States District Court for the District of Massachusetts. I'd like to extend my thanks to Kimberley Zelnick, Josh Kelner, and Allon Kedem for their thoughtful comments and helpful suggestions.

[1] 26 U.S.C. § 501(c)(3) (2000).

[2] *See infra* note 54 and accompanying text.

[3] *See infra* note 54 and accompanying text.

[4] *See infra* notes 49–52 and accompanying text.

[5] *See infra* notes 96–115 and accompanying text.

solicited such assistance.[6] The frequency with which the campaign activity prohibition was, at least allegedly, violated highlights the ineffectiveness of the current prohibition. As currently formulated, the Internal Revenue Code lacks a bright-line standard for determining whether or not actions taken by religious organizations are considered partisan, and hence unacceptable, or non-partisan and therefore permissible.[7] Consequently, the Internal Revenue Service ("IRS") is charged with enforcing a statute that requires a highly subjective interpretation of religious activity and that threatens to impinge upon fundamental First Amendment liberties. The combination of interpretive indeterminacy and potential infringement of constitutional rights has led to a regime in which statutory enforcement is sporadic at best and nonexistent at worst. Many churches still refrain from partisan activity—indeed, some are chilled from even participating in permissible political activity because of the guidelines' ambiguity—but the actions of others remain largely unchecked.

And enforcement concerns are just the tip of the iceberg. A more serious, if much ignored, problem with the prohibition is its infringement upon the free exercise rights guaranteed by the Religious Freedom Restoration Act ("RFRA"). RFRA ensures that free exercise rights are not undermined by the government unless the government possesses a compelling interest for the infringement.[8] With regard to the campaign activity prohibition, the government inaccurately bifurcates church practices into two separate spheres: the political and the religious. In so doing, the government ignores churches' history of religiously compelled political involvement and inaccurately portrays partisan activity as beyond the bounds of religious endeavor. Having thus isolated certain church practices as non-religious, the government then penalizes churches for engaging in them by withholding the tax-exemption given to all charitable organizations.[9] The burden this imposes is impermissible under RFRA, because the government lacks a compelling interest sufficient to justify such a broad-based prohibition on church activity.[10]

While the government does possess a compelling interest in prohibiting churches from spending tax-exempt funds on partisan activity—in fact, the Establishment Clause requires as much—the current prohibition is not narrowly tailored to advance that interest.[11] Instead, it bans broadly an entire class of protected activity, whether or not the activity implicates the relevant government interest. Consequently, the campaign activity prohibition should be held inapplicable to all church activity that does not involve the use of tax-exempt monies. This reformulation of the prohibition

---

[6] *See infra* notes 96–115 and accompanying text.
[7] *See infra* notes 84–85 and accompanying text.
[8] *See infra* notes 150–152 and accompanying text.
[9] 26 U.S.C. § 501(c)(3) (2000).
[10] *See infra* Part IV.B.ii.
[11] *See infra* Part II.C.

would also provide the additional benefit of a bright-line standard that would help both the IRS in its enforcement of the provision and churches in their adherence to it.

I. HISTORICAL OVERVIEW

A. Tax Exemptions

1. Pre-American Experience

Churches have received preferential tax treatment since at least the time of Constantine,[12] and some scholars have argued that the privilege can be traced back as far as ancient Egypt[13] or even Sumeria.[14] Additional examples can be found in the early Muslim conquests and medieval Europe.[15] But the American tradition of benevolent tax treatment for churches can be traced directly to the English experience.

In England, laws regarding tax exemptions for religious institutions were originally derived from two somewhat discordant roots: the common law and the equity law traditions.[16] Under the common law tradition, churches and other charitable trusts were granted exemptions because they disposed of certain responsibilities that would otherwise fall to the government.[17]

Much of English common law was established during the English Reformation in the sixteenth century.[18] At this time, the Tudors "consolidated their authority over religion and the church and subjected them to comprehensive ecclesiastical laws enforceable by both common law and com-

---

[12] Robert E. Rodes, Jr., *The Last Days of Erastianism: Forms in the American Church-State Nexus*, 62 HARV. THEOLOGICAL REV. 301, 317 (1969); *see also* MARTIN A. LARSON & C. STANLEY LOWELL, THE CHURCHES: THEIR RICHES, REVENUES, AND IMMUNITIES 19 (1969). After Constantine's conversion to Christianity, he granted a series of preferences to his newfound religion, including total exemption from all taxes. *See id.* And by the time he died in 337 A.D., the Christian Church had been the recipient of several forms of tax exemptions. While later rulers revoked some of these privileges, others remained part of Roman law. *See* ALFRED BALK, THE FREE LIST: PROPERTY WITHOUT TAXES 21 (1971).

[13] *See* Vaughn E. James, *Reaping Where They Have Not Sowed: Have American Churches Failed To Satisfy the Requirements for the Religious Tax Exemption?*, 43 CATH. LAW. 29, 32–35 (2004) (arguing that several examples of tax exemption for churches and clerics can be seen in pre- and post-Exodus Hebrew history).

[14] John W. Whitehead, *Tax Exemption and Churches: A Historical and Constitutional Analysis*, 22 CUMB. L. REV. 521, 524 (1992). Note that the ancient Sumerian practices are more difficult to categorize because of the overlap between religious and civil institutions.

[15] *See id.* at 529–31.

[16] *See* John Witte, Jr., *Tax Exemption of Church Property: Historical Anomaly or Valid Constitutional Practice?*, 64 S. CAL. L. REV. 363, 368 (1991).

[17] This was not a full exemption in the case of many churches. Only state-established churches qualified for the exemption, the exemption did not include all taxes, and the exemption could be lifted in times of crisis. Whitehead, *supra* note 14, at 531–32. *See also* Witte, *supra* note 16, at 368.

[18] Witte, *supra* note 16, at 369.

missary courts."[19] The common law included a substantial body of fundamentally religious law that governed everything from orthodox doctrine and morality to parish boundaries and church locations.[20] The church was an agency of the state and, as such, was regulated by, and intended to serve, the state. "By devoting their properties to the religious uses prescribed by the common law, church corporations and their clergy were discharging the state's responsibility for the established religion."[21] And in return for providing important governmental functions, the church was rewarded with tax support and tax exemptions like other agencies of the state.[22]

The church received tax exemptions from the law of equity as well, which provided tax benefits to organizations that dispensed certain social benefits.[23] Churches were excluded from taxes not because of their religious nature, but rather because of their characterization as charitable organizations. Parliament passed the Statute of Charitable Uses in 1601, providing the first working definition of charity. While the statute's provisions were largely ignored in practice,[24] its preamble has proven to be hugely influential for the Internal Revenue Code's understanding of charitable organizations.[25] According to the preamble, activities worthy of state support included:

> [R]elief of aged, impotent and poor people, . . . maintenance of sick and maimed soldiers and mariners, schools of learning, free schools, and scholars in universities, . . . repair of bridges, ports, havens, causeways, churches, . . . education and preferment of orphans, . . . relief, stock or maintenance for houses of correction, . . . and . . . relief or redemption of prisoners or captives.[26]

The statute did not list religion aside from a passing reference to the "repair of churches." Sir Francis Moore later wrote that this omission stemmed not from a desire to exclude religious activity, but instead from the hope that such an exclusion would protect church property from government confiscation.[27] At this point in English history, the Chantries Act

---

[19] *Id.*

[20] *Id.* at 370–71.

[21] *Id.* at 375.

[22] *See id.*

[23] *See id.* at 368.

[24] Whitehead, *supra* note 14, at 532. *See also* Ann M. Murphy, *Campaign Signs and the Collection Plate: Never the Twain Shall Meet?*, 1 PITT. TAX REV. 35, 42 (2003).

[25] Christine Roemhildt Moore, *Religious Tax Exemption and the "Charitable Scrutiny" Test*, 15 REGENT U. L. REV. 295, 298 n.19 (2003) ("Notice the correlation between the organizations listed [in the Statute of Charitable Uses] and those qualifying under 26 U.S.C. § 501 (2001): churches, charities, scientific, literary and educational organizations, . . . ").

[26] 2 RESTATEMENT (SECOND) OF TRUSTS § 368 cmt. a (1959) (quoting Statute of Charitable Uses, 1601, 43 Eliz., c. 4 (Eng.)).

[27] Whitehead, *supra* note 14, at 533.

of 1547 allowed for the confiscation of all property employed in "superstitious uses."[28] Because England's established religion periodically changed between Catholicism and Protestantism, religious property was subject to the vicissitudes of the monarch's religious affiliation. During a period of adverse religious affiliation, property employed for religious purposes by the out-of-favor denomination could be seized by the Crown under the Chantries Act as property used for superstitious purposes. Religious uses were thus:

> [O]n purpose omitted in the penning of the [Statute of Charitable Uses] . . . lest the gifts intended to be employed in purposes grounded on charity might, in change of time, contrary to the mind of the giver, be confiscated into the king's treasury; for religion, being variable according to the pleasures of succeeding princes, that which at one time is held for orthodox may, at another, be accounted superstitious; and then such lands are forfeited as appears in [the Chantries Act].[29]

However, despite attempts to exclude "religious uses" from the category of charitable purposes, the phrase was included no later than 1639.[30] It has subsequently been deemed to comprise one of the four principal divisions of charity in English law.[31]

2. The American Experience

The Supreme Court has recognized that "[f]ew concepts are more deeply embedded in the fabric of our national life, beginning with pre-Revolutionary colonial times, than for the government to exercise at the very least [a] kind of benevolent neutrality towards churches and religious exercise generally . . . ."[32] This benevolent neutrality is visible from the earliest stages of colonial existence. At first, American practices largely mimicked the English tradition of tax exemptions, though the specifics varied from colony to colony.[33] Like the English model, tax advantages were available exclusively to members of the established church. Colonies without established churches did not provide tax exemptions for any church.

Prior to the Revolutionary War, nine of the thirteen colonies had established churches.[34] In these colonies, the established church received

---

[28] Witte, *supra* note 16, at 376 n.49.

[29] *Id.* (quoting Moore, *Readings upon the Statute of 43 Elizabeth*, *in* G. DUKE, LAW OF CHARITABLE USES 131–32 (R. W. Bridgman ed. 1805)).

[30] *See id.*

[31] Comm'rs v. Pemsel, 1891 A.C. 531, 574 (H.L).

[32] Walz v. Tax Comm'n, 397 U.S. 664, 676–77 (1970).

[33] *See* James, *supra* note 13, at 38.

[34] *Id.* at 38.

government assistance, either in the form of taxes levied to support the church and church personnel, or through exemptions granted to otherwise applicable taxes.[35] But around the time of the Revolution, American politico-religious policy shifted, moving away from established churches and toward the separation of church and state.[36] Several colonies disestablished their churches either at the inception of, or shortly following, the Revolution, and with the adoption of the First Amendment in 1791, disestablishment became constitutionally mandated, at least at the federal level.

But disestablishment did not prove fatal to churches' tax-exempt status. Though neither the federal constitution nor any of the state constitutions provided any basis for granting tax exemptions,[37] both the federal government and various state governments soon began to recognize such exemptions.[38] Pennsylvania led the way on the state level, adopting an amendment that specifically protected church property from taxation.[39] Other states, including Virginia (now over its anti-clerical phase), would follow Pennsylvania's lead.[40]

Early federal tax statutes also included provisions granting limited exemptions to churches and other charitable organizations. The first such example occurred in 1802 when the Seventh Congress enacted a taxing statute for the County of Alexandria, which specifically exempted churches from taxation.[41] Another exemption followed in 1812 when Congress refunded import duties to religious organizations involved in the importation of religious articles.[42] And Congress in 1815 granted exemption from a tax on household furniture to charitable, religious and literary organizations.[43]

The exemption of charitable organizations, including churches, remained a part of federal tax policy fifty years later, when the government passed the first federal income tax during the Civil War.[44] And in 1894, tax exemptions for religious organizations were formally recognized for the first time in the Wilson Tariff Act of 1894. The Tariff Act provided an exemption for "corporations, companies, or associations organized and conducted solely for charitable, religious, or educational purposes . . . [and] stocks, shares, funds, or securities held by any fiduciary or trustee for charitable, religious, or educational purposes."[45] Although the tariff was held unconstitutional by the Court a year later in *Pollock v. Farmer's Loan &*

---

[35] *Id.*
[36] *Id.* at 39–40.
[37] *See* D. B. ROBERTSON, SHOULD CHURCHES BE TAXED? 69 (1968).
[38] *See* James, *supra* note 13, at 40.
[39] *See* ROBERTSON, *supra* note 37, at 69.
[40] *See id.*
[41] *See* Walz v. Tax Comm'n, 397 U.S. 664, 677 (1970).
[42] *See id.*
[43] Act of Jan. 18, 1815, ch. 23, § 14, 3 Stat. 186, 190 (1815).
[44] *See* James, *supra* note 13, at 41.
[45] Tariff Act of 1894, ch. 349, § 32, 28 Stat. 509, 556 (1894).

government assistance, either in the form of taxes levied to support the church and church personnel, or through exemptions granted to otherwise applicable taxes.[35] But around the time of the Revolution, American politico-religious policy shifted, moving away from established churches and toward the separation of church and state.[36] Several colonies disestablished their churches either at the inception of, or shortly following, the Revolution, and with the adoption of the First Amendment in 1791, disestablishment became constitutionally mandated, at least at the federal level.

But disestablishment did not prove fatal to churches' tax-exempt status. Though neither the federal constitution nor any of the state constitutions provided any basis for granting tax exemptions,[37] both the federal government and various state governments soon began to recognize such exemptions.[38] Pennsylvania led the way on the state level, adopting an amendment that specifically protected church property from taxation.[39] Other states, including Virginia (now over its anti-clerical phase), would follow Pennsylvania's lead.[40]

Early federal tax statutes also included provisions granting limited exemptions to churches and other charitable organizations. The first such example occurred in 1802 when the Seventh Congress enacted a taxing statute for the County of Alexandria, which specifically exempted churches from taxation.[41] Another exemption followed in 1812 when Congress refunded import duties to religious organizations involved in the importation of religious articles.[42] And Congress in 1815 granted exemption from a tax on household furniture to charitable, religious and literary organizations.[43]

The exemption of charitable organizations, including churches, remained a part of federal tax policy fifty years later, when the government passed the first federal income tax during the Civil War.[44] And in 1894, tax exemptions for religious organizations were formally recognized for the first time in the Wilson Tariff Act of 1894. The Tariff Act provided an exemption for "corporations, companies, or associations organized and conducted solely for charitable, religious, or educational purposes . . . [and] stocks, shares, funds, or securities held by any fiduciary or trustee for charitable, religious, or educational purposes."[45] Although the tariff was held unconstitutional by the Court a year later in *Pollock v. Farmer's Loan &*

---

[35] *Id.*
[36] *Id.* at 39–40.
[37] *See* D. B. ROBERTSON, SHOULD CHURCHES BE TAXED? 69 (1968).
[38] *See* James, *supra* note 13, at 40.
[39] *See* ROBERTSON, *supra* note 37, at 69.
[40] *See id.*
[41] *See* Walz v. Tax Comm'n, 397 U.S. 664, 677 (1970).
[42] *See id.*
[43] Act of Jan. 18, 1815, ch. 23, § 14, 3 Stat. 186, 190 (1815).
[44] *See* James, *supra* note 13, at 41.
[45] Tariff Act of 1894, ch. 349, § 32, 28 Stat. 509, 556 (1894).

Trust Co.,[46] for reasons unrelated to the tax exemption, it would serve as a precedent for federal tax legislation in the coming decades.

After the Sixteenth Amendment effectively overruled *Pollock* and ensured the constitutionality of federal income taxes, Congress passed new tax legislation that provided tax advantages to charities, including religious organizations.[47] Individuals were still not allowed to deduct their charitable contributions,[48] but this oversight was partially corrected in the Revenue Act of 1917.[49] And in the Revenue Act of 1921, Congress decreed that all contributions "to or for the use of corporations, community chests, funds, or foundations organized and operated exclusively for charitable, etc., purposes were deductible."[50] Still revealing its English antecedents, the Revenue Act's original list of qualified tax-exempt organizations was taken from the English common law of charitable uses,[51] though the current list has been expanded to encompass organizations not originally included under the common law.[52]

B. The Campaign Activity Prohibition

Today, I.R.C. § 501(a) provides a federal tax exemption for charitable organizations, § 170 permits individuals to deduct contributions to charitable organizations, and § 501(c)(3) enumerates the types of organizations that qualify for tax-exempt treatment under § 501(a). According to § 501(c)(3), the following organizations are exempt: "Corporations, and any community chest, fund, or foundation, organized and operated exclusively for *religious*, charitable, scientific, testing for public safety, literary, or educational purposes, or to foster national or international amateur sports competition . . . or for the prevention of cruelty to children or animals . . . ."[53]

In its enumeration of qualified organizations, § 501(c)(3) largely tracks the text of earlier Revenue Acts, but the section ends with a more recent addition. After listing the various organizations that qualify for tax-exempt

---

[46] 157 U.S. 429 (1895), *modified by* 158 U.S. 601 (1895) (rejecting as unconstitutional "direct taxes" such as the Wilson Tariff).

[47] *See* Murphy, *supra* note 24, at 45.

[48] Massachusetts Congressman John Jacob Rogers (R) suggested as an amendment to the Tariff Act of 1913 that individuals should be able to deduct contributions made to charitable organizations, but his suggestion was rejected by Congress. *See* Carol A. Jones, *Hernandez v. Commissioner: The Supreme Court Forces a Square Peg into a Round Hole*, 25 WAKE FOREST L. REV. 917, 924 (1990).

[49] *See* Murphy, *supra* note 24, at 45.

[50] Rockefeller v. Comm'r, 76 T.C. 178, 185 (1981), *aff'd*, 676 F.2d 35 (2d Cir. 1982). The Revenue Act of 1921 largely paralleled the Act of 1917, but also allowed for the deduction of indirect contributions such as gifts in trust for the benefit of charities. *See* Murphy, *supra* note 24, at 46.

[51] *See* Murphy, *supra* note 24, at 46.

[52] *See id.*

[53] 26 U.S.C. § 501(c)(3) (2000) (emphasis added).

treatment, Congress limits the exemption, requiring organizations to make certain that:

> [N]o substantial part of [their] activities . . . is carrying on propaganda, or otherwise attempting, to influence legislation . . . and *which does not participate in, or intervene in (including the publishing or distributing of statements), any political campaign on behalf of (or in opposition to) any candidate for public office.*[54]

The prohibitions on lobbying and campaign activities both arose as Senate floor amendments, and neither was subject to congressional hearings.[55] Senator David Reed introduced the lobbying prohibition, passed by Congress in 1934, and then-Senator Lyndon Baines Johnson (D-Tex.) introduced the campaign activity prohibition passed two decades later in 1954.[56] Because the campaign prohibition was raised as a floor amendment and not subject to debate, the legislative record is essentially silent.[57]

In the absence of a legislative record, scholars continue to debate the impetus behind the campaign prohibition. While there is no clear consensus regarding Johnson's motivation,[58] four theories dominate the discussion, three of which focus on the Johnson-Dougherty primary in 1954.[59] These three theories put forth different permutations of the argument that Johnson wanted to stop his opponent Dudley Dougherty from receiving either financial assistance or other, non-monetary aid from various chari-

---

[54] *Id.* (emphasis added).

[55] *See* Steffen Johnson, *Of Politics and Pulpits: A First Amendment Analysis of IRS Restrictions on the Political Activities of Religious Organizations*, 42 B.C. L. REV. 875, 880 (2001).

[56] *See id.*

[57] *See* 100 CONG. REC. 9604 (1954) (statement of Sen. Johnson) (stating only that the amendment's aim was to deny tax-exempt status to "not only those people who influence legislation but also to those who intervene in any political campaign on behalf of any candidate for office"). It has been alleged that Johnson intended that the prohibition be accepted without legislative history in order to conceal his self-interested motives behind the legislation. *See* D. Benson Tesdahl, *Intervention in Political Campaigns After the Pickle Hearings—A Proposal for the 1990's*, 4 EXEMPT ORG. TAX REV. 1165, 1178 n.26, 1179 n.38 (1991). According to this account, Lawrence M. Woodworth, a Johnson staffer, stated that Johnson was upset about support that a political opponent was receiving from a charitable organization and requested that Woodworth draft the language that Johnson proposed later that day. *Id.* Woodworth also stated that Johnson did not want there to be any legislative history for the prohibition. *Id.*

[58] *Compare* Murphy, *supra* note 24, at 46–63 (arguing that the preponderance of the evidence suggests that Johnson proposed the amendment in an attempt to moderate a more far-reaching proposal from the previous day), *with* Patrick L. O'Daniel, *More Honored in the Breach: A Historical Perspective of the Permeable IRS Prohibition on Campaigning by Churches*, 42 B.C. L. REV. 733 (2001) (arguing that Johnson proposed the amendment to stop tax-exempt, charitable organizations from aiding his opponent in a political campaign).

[59] Judith E. Kindell & John Francis Reilly, *Election Year Issues* 114 (IRS publication), *available at* http://www.irs.gov/pub/irs-tege/eotopici02.pdf (last visited Nov. 23, 2005).

table organizations and used the prohibition as his means of doing so.[60] The fourth theory disputes the notion that Johnson acted out of political self-interest and suggests that, on the contrary, Johnson's proposal was a response to an intemperate, alternative proposal motivated by anti-Communist sentiment in Congress.[61] According to this theory, Congress was moving to pass a more restrictive prohibition on the activities of charitable organizations, motivated by a fear that charitable foundations may be used as a vehicle for Communists or Communist sympathizers, who hoped to use putatively charitable foundations to poison American political discourse with leftist propaganda.[62]

Whatever the reason for Johnson's action, it is clear that his motivation was unrelated to the activities of religious institutions.[63] "Indeed, Senator Johnson did not hesitate to coordinate support from churches when it was to his own political advantage,"[64] and the language of the amendment did not single out religious organizations in any way.[65] Churches were simply included under the broad rubric of charitable organizations. Moreover, George Reedy, Johnson's chief aide in 1954, later stated that to the best of his recollection, "Johnson would never have sought restrictions on religious organizations . . . ."[66] The ban on political activity by churches is apparently the result of historical happenstance and was neither the result of any Congressional intent nor the reflection of any overarching political goal.[67]

II. THE IRS GUIDELINES AND THEIR PROBLEMS

As currently formulated, the Internal Revenue Code lacks a bright-line standard for determining whether actions taken by religious organizations are partisan, and hence unacceptable, or non-partisan, and therefore permissible. As a result, the IRS is forced to attempt highly subjective interpretations of religious activity which threaten to impinge upon fundamental First Amendment liberties. This combination of interpretive indetermi-

---

[60] One theory is that Johnson was concerned that funds from a charitable organization were being funneled to Dougherty's campaign. *See* BRUCE R. HOPKINS, THE LAW OF TAX-EXEMPT ORGANIZATIONS 504 (7th ed. 1998). Two other Dougherty-supporting organizations, Facts Forum, which produced both television shows and a national periodical, and the Committee for Constitutional Government, may have triggered Johnson's ire, and thus his proposal. *See* Kindell & Reilly, *supra* note 59, at 448–49.

[61] *See* Murphy, *supra* note 24, at 49–55.

[62] *See id.*

[63] *See id.* 46–63. *See generally* O'Daniel, *supra* note 58.

[64] Johnson, *supra* note 55, at 881.

[65] Nor were any of the charitable organizations that Johnson may have been targeting—namely, the groups he believed to be supporting his political opponents—affiliated with any religion.

[66] Deirdre Dessingue, *Prohibition in Search of a Rationale: What the Tax Code Prohibits; Why; To What End?*, 42 B.C. L. REV. 903, 917 n.51 (2001).

[67] Though, to be fair, Congress has never repealed the ban for religious groups despite several attempts by House members to pass such legislation.

nacy and potential infringement of a constitutional right have led to a regime in which only the most brazen of violators are punished, if anyone is punished at all.[68] For example, the frequency with which the campaign activity prohibition was, at least allegedly, violated in 2004 and the lack of attendant sanctions from the IRS highlights the untenable nature of the current prohibition. A number of churches still refrain from partisan activity, but the actions of many others remain largely unaffected.

A. The Law

As discussed above, churches, like other charitable organizations, are exempted from taxation by the Internal Revenue Code as long as they refrain from engaging in certain activities, including "participat[ing] in, or interven[ing] in (including the publishing or distributing of statements), any political campaign on behalf of (or in opposition to) any candidate for public office."[69] But churches are not treated identically to other charitable organizations by the Internal Revenue Service. Unlike other tax-exempt organizations, churches do not need to apply to the IRS to obtain an advance determination that they satisfy the requirements for tax exemption under § 501(c)(3).[70] Instead, churches may simply hold themselves out as tax-exempt to congregants, and the congregants may deduct any charitable contributions they choose to make on the assumption that their church qualifies under § 501(c)(3).[71]

The IRS keeps track of which organizations have received a ruling or determination letter verifying their tax-deductible status in the periodically updated "Publication No. 78."[72] This listing can be used by donors to determine whether or not an organization has been extended tax-exempt status and assure them that their contributions are deductible under § 170(a). Donations to a church that has not been subject to a formal ruling by the IRS are deductible; however, in the event of an audit, the taxpayers claiming the deductions will be required to prove that the church met the requirements of § 501(c)(3).[73]

The unique tax treatment of churches is also visible in the special restrictions placed on the IRS's ability to investigate the tax-exempt status of churches.[74] The Church Audit Procedures Act established both the circumstances under which the IRS may investigate a church and the means by which it may proceed.[75] "Upon a 'reasonable belief' by a high-level

---

[68] *See* Branch Ministries v. Rossotti, 211 F.3d 137, 139 (D.C. Cir. 2000) (describing the brazen violation of a full page anti-Clinton advertisement in the *Washington Post*).

[69] 26 U.S.C. § 501(c)(3) (2000).

[70] 26 U.S.C. §§ 508(a), (c)(1)(A).

[71] *Id.*

[72] *See* Rev. Proc. 82-39, 1982-2 C.B. 759, §§ 2.01, 2.03.

[73] *See Branch Ministries*, 211 F.3d at 139.

[74] *Id.*

[75] 26 U.S.C. § 7611 (2000).

Treasury official that a church may not be exempt from taxation under section 501, the IRS may begin a 'church tax inquiry,'"[76] which is defined as:

> [A]ny inquiry to a church (other than an examination) to serve as a basis for determining whether a church—
>
> (A) is exempt from tax under section 501(a) by reason of its status as a church, or
>
> (B) is . . . engaged in activities which may be subject to taxation . . . .[77]

If the IRS is unable to make a determination on this first pass, it may opt for a more probing inquiry, a "church tax examination," in which the IRS may examine its records and/or activities "to determine whether [the] organization claiming to be a church is a church for any period."[78]

In making its determination, the IRS need not consider whether the activity in dispute constituted a substantial part of the church's actions, because the campaign activity prohibition is absolute.[79] As stated by the Seventh Circuit, the exemption is lost "by participation in any political campaign on behalf of any candidate for public office. It need not form a substantial part of the organization's activities."[80] After determining that a charitable organization has violated the campaign activity prohibition, the IRS has two weapons at its disposal: the revocation of tax-exempt status and/or the levying of § 4955 excise taxes, which are taxes levied against the church in response to political participation.[81] Although Congress primarily intended § 4955 to serve as an additional penalty to revocation,[82] the IRS also has the option of using that section by itself as an intermediate sanction in cases where the expenditure was unintentional and involved only a small amount and where the organization subsequently adopted procedures to assure that similar expenditures would not be made in the future.[83]

Participation or intervention in political activity is never defined by the IRS,[84] but it includes, though it is not limited to, "the publication or distribution of written or printed statements or the making of oral state-

---

[76] *See* Branch Ministries v. Rossotti, 211 F.3d 137, 140 (D.C. Cir. 2000).

[77] 26 U.S.C. § 7611(h)(2) (2000).

[78] *Id.* § 7611(b)(1)(A),(B).

[79] *See* Kindell & Reilly, *supra* note 59, at 352.

[80] U.S. v. Dykema, 666 F.2d 1096, 1101 (7th Cir. 1981). *See also* Ass'n of the Bar of New York v. Comm'r, 858 F.2d 876, 877 (2d Cir. 1988).

[81] *See* 26 U.S.C. § 4955 (2000).

[82] *See* HOUSE BUDGET COMM. REP., H.R. Rep. No. 100-391, at 1623–24 (1987).

[83] 59 Fed. Reg. 64,359, 64,360 (Dec. 14, 1994).

[84] *See* Kindell & Reilly, *supra* note 59, at 344.

ments on behalf of or in opposition to a candidate for public office."[85] As such, any endorsement of a candidate is strictly prohibited and even the rating of candidates on a non-partisan basis is not allowed.[86] Distributing partisan campaign literature, providing or soliciting any form of support for a political campaign or candidate, and establishing a political action committee (PAC) are similarly disallowed.[87] Purely non-partisan activities like voter registration drives and issue ads are explicitly permitted as long as they are in keeping with the organization's tax-exempt purpose and are not a backhanded attempt to engage in partisan activity.[88] But, "[i]n situations where there is no explicit endorsement or partisan activity, there is no bright-line test for determining if the I.R.C. § 501(c)(3) organization participated or intervened in a political campaign. Instead, all the facts and circumstances must be considered."[89] As a result, the IRS lacks the specific statutory guidance necessary to help agents determine when, and when not, to intervene in arguably impermissible church activities.[90]

B. Abuses

One frequent and arguably impermissible church activity is the endorsement of candidates in political campaigns. The campaign activity prohibition has oft been honored in the breach. Both major political parties have tested its limitations. In *Branch Ministries v. Rossotti*, the D.C. District Court cited no fewer that sixty-five examples of candidates campaigning in various churches and synagogues, including episodes involving Jesse Jackson, Al Gore, Bill Clinton, Oliver North, and Rudolph Giuliani.[91] And this figure is likely dwarfed by the number of political endorsements given from the pulpit by various religious leaders.

In one particularly egregious example from Rev. Jesse Jackson's first presidential campaign, Jackson called upon black churches to provide manpower and facilities for his campaign fundraising efforts.[92] Leading up to Super Bowl Sunday, the Jackson campaign distributed flyers and encouraged church members to bring donations for Jackson's campaign to church

---

[85] Treas. Reg. § 1.501(c)(3)-1(c)(3)(iii) (as amended in 1990).

[86] *See Ass'n of the Bar of New York*, 858 F.2d at 878.

[87] *See* Kindell & Reilly, *supra* note 59, at 344.

[88] *See id.*

[89] *See id.*

[90] *See id.*

[91] 40 F. Supp. 2d 15, 21–22 (D.D.C. 1999). While politicians are technically allowed to address churches and synagogues, the invitation by churches to campaigning politicians smacks of intervention in a political campaign, whether or not the church's endorsement is explicit. This interpretation is supported by the heavy reliance made by many candidates on speeches from the pulpit during campaign season.

[92] *The D.C. Circuit Review August 1999–July 2000: Recent Decisions of the United States Court of Appeals for the District of Columbia Circuit: Constitutional Law*, 69 GEO. WASH. L. REV. 554, 567 (2001).

that Sunday.[93] Campaign offerings were then collected separately from regular church donations by many churches.[94] The IRS did not even investigate this potential violation of § 501.

Campaign politicking increased markedly in the 2000 election, with one observer citing at least eighteen examples of questionable behavior by churches and synagogues.[95] And in 2004, politicking by churches reached an all-time high.[96] According to Americans United for the Separation of Church and State ("Americans United"), more potential violations occurred in 2004 alone than in the preceding five years combined.[97] In the first two weeks of June, Bush staffers set out to identify 1600 friendly congregations in Pennsylvania for campaign assistance.[98] President Bush also visited the Pope and reportedly complained to Cardinal Angelo Sodano, the Vatican Secretary of State, that "not all American bishops are with me."[99] Additionally, twenty members of the House of Representatives attempted to slip a "Safe Harbor for Churches" provision, which would have allowed churches to support candidates, into a jobs bill.[100] This flurry of activity caused Rev. Barry Lynn, Executive Director of Americans United, to comment that "[t]his is the most concentrated dose of religion and politics I have ever seen. And it looks like it's full steam ahead."[101]

The Reverend's prediction proved accurate. Veiled or outright endorsements were made in churches from Cincinnati[102] to Philadelphia[103]

---

[93] *Id.*

[94] *Id.*

[95] *See* O'Daniel, *supra* note 58, at 736.

[96] Telephone Interview with Dohnya Khalili, Spokesperson, Americans United for the Separation of Church and State (Dec. 2, 2004).

[97] *See id.*

[98] *See* Alan Cooperman, *Churchgoers Get Direction from Bush Campaign*, WASH. POST, July 1, 2004, at A6.

[99] Don Lattin, *Politics and the Church: Bush Woos Faithful with a Religious Fervor*, S.F. CHRON., June 21, 2004, at A1.

[100] *See* American Jobs Creation Act of 2004, H.R. 4520, 108th Cong., § 692 (2004). The Safe Harbor provision would have allowed pastors to engage in political activity and endorse candidates so long as they made clear that they were acting as individuals and not as representatives of the church and did not make statements in church publications, at church functions, or using church funds. The clergy would also have been allowed three unintentional violations of the statute per year, with an increased penalty for each violation. The first violation would have entailed a corporate tax on one week's revenues, the second, a corporate tax on half of the church's annual revenues, and the third, a corporate tax on a year's worth of revenue. A fourth violation would have led to the full revocation of the church's tax-exempt status. *See also* Lattin, *supra* note 99; Alan Cooperman, *Speaker Pushes Jobs Bill Provision: Religious Leaders Would Be Allowed More Freedom to Participate in Partisan Politics*, WASH. POST, June 9, 2004, at A19.

[101] Lattin, *supra* note 99.

[102] *See* Edward E. Plowman, *Pulpit Politics*, WORLD MAG., Nov. 6, 2004. At Allen Temple AME church, the minister, Donald H. Jordan stated, "I'm not worried about the law; I'm asking you to support him," after Senator Edwards had spoken. *Id.*

[103] *Id.* At the Mt. Airy Church, Pastor Ernest C. Morris followed Sen. Kennedy to the pulpit and declared, "I can't tell you who to vote for, but I can tell you what my mamma told me last week: 'Stay out of the Bushes.'" *Id.*

to Aspen.[104] Jerry Falwell publicly admitted to supporting President Bush from the pulpit.[105] And pulpit appearances by campaigning politicians on both sides of the aisle were just as numerous as clerical endorsements.[106] Perhaps most troubling were the attempts by the Bush-Cheney campaign to use churches as recruiting grounds for its re-election campaign. In July, the Republican National Committee asked Roman Catholics who supported Bush to provide copies of their parish directories to the campaign.[107] The month before, the Bush-Cheney campaign had sent a detailed list of instructions to its religious volunteers across the country.[108] Religious Bush supporters had a list of twenty-two objectives with deadlines ranging from July 31 to October 31.[109] These objectives included, for example, instructions to send church directories to the Bush-Cheney state headquarters; to talk to pastors about holding Voter Registration Drives;[110] and, as the election neared, to host campaign-related potluck dinners, distribute voter guides, and call members of the church.[111]

Last but not least on the list of questionable church behavior was the distribution of the Christian Coalition's putatively non-partisan voter guides. This past summer, the Christian Coalition estimated that by Election Day it would have given out approximately thirty million voter guides,[112] most of which were disseminated through churches. The pamphlets selected fifteen politically charged issues and detailed each candidate's position on each issue. "[In 2004], Bush was portrayed as opposing 'unrestricted abortion on demand,' 'federal firearm registration' and United Nations command of U.S. soldiers, while Kerry was listed as offering 'no response' on each of these concerns that aroused the passion of social conservatives."[113] The Christian Coalition has distributed such voter guides through

---

[104] *See* Deborah Frazier, *Pulpit Politics Irk Parish; Aspen Priest Angers Flock with Voting Instructions*, ROCKY MT. NEWS, Oct. 28, 2004, at 30A.

[105] *See* David Kirkpatrick, *Citing Falwell's Endorsement of Bush, Group Challenges His Tax-Exempt Status*, N.Y. TIMES, July 16, 2004, at A16.

[106] *See, e.g.*, David Karp, *IRS Tells Churches: No Politics*, ST. PETERSBURG TIMES, Sept. 15, 2004, at 1B (discussing campaign stops by John Kerry and Terry McAuliffe in Orlando and Miami).

[107] *See* Frances Grandy Taylor, *Politics Pushes at the Door of Religion; Campaign Tactics Blur Church-State Line*, HARTFORD COURANT, July 24, 2004, at A1.

[108] *See* Cooperman, *supra* note 98; *see also* David Kirkpatrick, *Bush Appeal to Churches Seeking Help Raises Doubts*, N.Y. TIMES, July 2, 2004, at A15. Technically, individuals may support candidates, even occasionally using the church as a medium (as long as they are not the minister or rabbi), but this extensive commingling of churches and partisan activity appears to violate at least the spirit of the prohibition, if not the letter of it. *See* Kindell & Reilly, *supra* note 59, at 366. Perhaps more importantly, it reveals the substantial ambiguity and enforcement difficulties caused by the current terms of the prohibition.

[109] Cooperman, *supra* note 98.

[110] *Id.* (internal quotations omitted).

[111] *Id.* (internal quotations omitted).

[112] *See* David Lightman, *GOP Hopes Religion Sways Midwest*, HARTFORD COURANT, Oct. 29, 2004, at A1.

[113] Walter Shapiro, *Vigorous Efforts Attempt To Turn Up Voter Turnout*, USA TODAY,

churches since the 1992 elections.[114] Moreover, no church has ever lost its tax-exempt status for assisting the Christian Coalition—despite the Internal Revenue Code's prohibition on distributing partisan campaign literature and the fact that former Coalition executive director Ralph Reed has boasted of the political advantage that the guides provide for Republican candidates.[115] This non-enforcement should not be surprising. In fact, no church has ever lost its tax-exempt status for distribution of any of the aforementioned literature.

C. Explanations

Despite the number of alleged violations over the years, IRS auditing activity has been extremely limited and tax reprisals by the IRS have been essentially nonexistent.[116] In the fifty-four years following the passage of the prohibition, only two churches have ever lost their tax-exempt status and only two others have been required to pay excise taxes.[117] The most recent church to lose its tax-exempt status demonstrates just how entangled with politics a church must become before losing its tax-exempt status.[118]

Four days before the presidential election of 1992, the Church at Pierce Creek, a Christian Church in Binghamton, New York, placed full-page advertisements in both the *Washington Post* and *USA Today*. The advertisements proclaimed, "Christians Beware. Do not put the economy ahead of the Ten Commandments."[119] The ads then proceeded to detail the ways in which then-governor Clinton's positions on various issues ran afoul of Biblical precepts[120] and concluded with the question, "How then can we vote for Bill Clinton?"[121] Small print at the bottom of the ad explained that, "This advertisement was co-sponsored by The Church at

---

Oct. 25, 2004, at 10A.

[114] *See* Peter Brien, *Voter Pamphlets: The Next Best Step in Election Reform*, 28 J. LEGIS. 87, 89 (2002).

[115] *See, e.g.*, Steven B. Imhoof, *The Politics of Politicking Under I.R.C. § 501(c)(3): A Guide for Politically Active Churches*, 5 NEXUS 102 n.48 (2000); Shannon L. Race, *The Christian Coalition As a Tax-Exempt Organization: Federal Income Tax Recommendations for the Politically Active TEO*, 43 WAYNE L. REV. 1931, 1957 n.113 (1997).

[116] *See* Murphy, *supra* note 24, at 68; Randy Lee, *When a King Speaks of God; When God Speaks to a King: Faith, Politics, Tax-Exempt Status, and the Constitution in the Clinton Administration*, 63 LAW & CONTEMP. PROBS. 391, 424 (2001).

[117] *Review of Internal Revenue Code Section 501(c)(3) Requirements for Religious Organizations: Hearing on H.R. 2357 Before the Subcomm. on Oversight of the H. Comm. on Ways & Means*, 107th Cong. 80 (2002) (statement of Rep. Karen L. Thurman, Member, Subcomm. on Oversight). Two religious organizations that were not churches also lost their tax-exempt status during this time period. *See* Branch Ministries v. Rossotti, 40 F. Supp. 2d 18, 19 (D.D.C. 1999).

[118] *See Branch Ministries*, 40 F. Supp. 2d at 17–19.

[119] *Id.* at 17.

[120] *Id.*

[121] *Id.*

Pierce Creek, Daniel J. Little, Senior Pastor, and by churches and concerned Christians nationwide. *Tax-deductible donations for this advertisement gladly accepted.* Make donations to: The Church at Pierce Creek,"[122] presumably to ensure that the IRS was compelled to take action.

Naturally, most churches that choose to engage in political activity employ more subtle means of communicating their messages. And the IRS's actions regarding the Church at Pierce Creek have not stopped churches from signaling support for candidates in numerous other ways.[123] Nor has the Church at Pierce Creek incident spurred the IRS to punish similarly clear, if more minor, violations of the statute, such as political endorsements made by ministers. There are several reasons for the continued abuses and the IRS's reluctance. First, the IRS is unwilling to punish churches severely for behavior that has historically been considered acceptable, and lacking a feasible alternative to complete revocation, the most acceptable option is complete—or near complete—abeyance from regulation. The IRS appears loath to punish churches for minor infractions. For example, no church has ever lost its tax-exempt status for endorsing a candidate or allowing a candidate to speak from the church's pulpit.[124]

Second, because the IRS does not have a bright line test,[125] punishment of infractions is necessarily fact-intensive and highly subjective. Most cases of church intervention are not as obvious as that of the Church at Pierce Creek and do not involve any of the handful of acts that are deemed unacceptable.[126] Given the importance of the First Amendment freedoms at stake, the IRS appears hesitant to revoke the tax-exempt status of a church in any situation which could conceivably be construed as non-partisan. One need only consider the following ambiguities with which the IRS must contend to understand the breadth of the problem. First, the IRS must determine the difference between issue and candidate advocacy. This is significant because the issue/candidate advocacy distinction determines whether the IRS has the authority to revoke tax-exempt status: while the IRS permits charitable organizations to engage in issue advocacy, candidate advocacy is cause for revocation. IRS enforcement is difficult because there is no clear line between advocacy in support of an issue and that in support of a candidate. Consequently, churches may advocate their positions on an issue—even during a campaign period—but the IRS may theoretically revoke their tax-exempt status if it feels that the church is using a "code word" such as "pro-life"

---

[122] *Id.* (emphasis added).

[123] Alan L. Feld, *Rendering unto Caesar or Electioneering for Caesar? Loss of Church Tax Exemption for Participation in Electoral Politics*, 42 B.C. L. REV. 931, 939 (2001).

[124] *See* Branch Ministries v. Rossotti, 40 F. Supp. 2d 18, 21–22 (D.D.C. 1999).

[125] *See* Kindell & Reilly, *supra* note 59, at 344.

[126] *See id.*

or "pro-choice" to stand in for a candidate.[127] Thus, the IRS is left with the unenviable task of determining when a church is speaking out on a particular issue and when its speech is really a proxy for impermissible candidate advocacy.[128] The IRS has had an understandably difficult time making that distinction.

Third, the IRS must consider the individual/organization distinction. While charitable organizations are barred from engaging in partisan political activities, the individual members of such organizations, even the heads of such organizations, are not so barred. In some circumstances, however, the acts of an individual may be imputed to his organization.[129] As a result, if the IRS wants to enforce the statute, it must differentiate between actions by individuals and actions by an organization. It must also determine in which circumstances to impute individuals' actions to their organizations. Such judgments can be nearly impossible to make, involving fact-intensive and subjective analyses that the IRS is hesitant to undertake.

Fourth, the IRS must confront the ability of churches to circumvent the prohibition on 501(c)(3) political activity by creating 501(c)(4) organizations which establish political action committees (PACs).[130] Charitable organizations are expressly forbidden from establishing PACs.[131] However, a charitable organization may establish a related, but independent § 501(c)(4) organization.[132] Section 501(c)(4) organizations are still exempt from taxes but are not tax-deductible to donors.[133] Such organizations cannot conduct partisan political activity either, but they may establish their own PACs.[134] Therefore, a § 501(c)(3) organization cannot establish a PAC, but it can establish a § 501(c)(4) arm that can establish a PAC. In so doing, the § 501(c)(3) organization will still maintain its tax-exempt status, as long as the activities of the downstream organizations cannot be attributed to the § 501(c)(3).[135] While this could, in theory, appear to impose a substantial burden on churches, the IRS's requirements are more bark than bite. Generally, as long as the organizations are separately incorporated and keep records sufficient to prove that tax-deductible contributions to the § 501(c)(3) organization are not being used to pay for the activities of the other organizations, the IRS will not attribute the acts of one organization to the other.[136]

---

[127] *See id.* at 345.
[128] *See id.* at 346.
[129] *Id.* at 363–64.
[130] *See* Kindell & Reilly, *supra* note 59, and accompanying text.
[131] *See id.*
[132] *Id.* at 367.
[133] *See* 26 U.S.C. § 170(c) (2000).
[134] Kindell & Reilly, *supra* note 59, at 367.
[135] *Id.*
[136] *Id.* at 367.

Fifth, the IRS must discern between non-partisan, and partisan, and thereby impermissible, voter guides published by churches and religious organizations. Charitable organizations are prohibited from distributing voter education material produced by a candidate or PAC.[137] However, they are not prohibited from distributing voter guides, as long as those guides are non-partisan.[138] To determine whether or not the guides have a partisan agenda, the IRS must inquire whether a "wide variety of issues" are covered in the guide and whether the questions "indicate a bias toward the organization's preferred answer."[139] Because these hazy standards are so subjective, the IRS is rarely presented with a case in which the offending party has committed an act that is dispositively, inarguably partisan.[140] As a result, the IRS almost never finds churches guilty of infractions.[141]

III. The Free Exercise Clause

The discernment and enforcement difficulties that the IRS confronts in its efforts to implement the guidelines as they exist currently create a regime where the power of the IRS is undermined. But practical enforcement and policy concerns are just a fraction of the problem: a more serious concern stems from the prohibition's infringement of the free exercise rights guaranteed by the Religious Freedom Reformation Act ("RFRA"). RFRA ensures that free exercise rights may not be limited by the government unless the government possesses a compelling interest for the limitation. The campaign activity prohibition fails RFRA's test, as the government inaccurately bifurcates church practices into two separate spheres: the political and the religious. In so doing, the government ignores churches' history of religiously compelled political involvement and inaccurately portrays partisan activity as beyond the bounds of religious endeavor. Having thus defined certain church practices as non-religious, the government then penalizes churches for engaging in them by withholding the

---

[137] *Id.* at 370.

[138] *Id.*

[139] *Id.* at 371–72. One example that illustrates both a question indicating a bias toward the organization's preferred answer and the difficulties facing IRS enforcement is this year's Christian Coalition Voter Guide, which refers to "unrestricted abortion on demand" instead of simply asking whether a candidate supports abortion. *See* Shapiro, *supra* note 113. The wording shows a clear bias toward the pro-life view associated with the Republican Party, but could still be described as an issue in the election. The Christian Coalition guide also demonstrates the difficulty of enforcing the tax laws in another way. In order to satisfy IRS requirements, candidate questionnaires such as the Christian Coalition Voter's Guide must send a questionnaire to all candidates and publish the unedited responses of all candidates. *See* Kindell & Reilly, *supra* note 59, at 371–72. Because Kerry did not respond to the Christian Coalition, he is listed as having "no response" to a variety of issues to which Kerry did, in fact, have a position. This discrepancy seems to serve as thinly veiled candidate endorsement.

[140] *Cf.* Branch Ministries v. Rossotti, 40 F. Supp. 2d 18, 21 (D.D.C. 1999).

[141] *See id.*

tax-exemption given to all charitable organizations. This burden is impermissible under RFRA, because the government lacks a compelling interest sufficient to justify such a broad-based prohibition on church activity.

The government does possess a compelling interest in prohibiting churches from spending tax-exempt funds on partisan activities; in fact, the Establishment Clause demands such a ban.[142] However, the current prohibition is not tailored to address that concern. Instead, it bans broadly an entire class of protected activity, whether or not the activity implicates the relevant government interest. Consequently, the campaign activity prohibition should be held inapplicable to all church activity that does not involve the use of tax-exempt monies. This reformulation of the prohibition would also provide a bright-line standard that would help both the IRS in its enforcement of the provision and churches in their adherence to it.

A. Applicable Law

1. RFRA

RFRA emerged from a battle between Congress and the Supreme Court over the meaning of the Free Exercise Clause of the First Amendment.[143] Prior to 1990, individuals whose free exercise rights had been burdened by the government were constitutionally entitled to relief so long as the government lacked a compelling state interest for burdening the right.[144] Then came *Employment Division v. Smith*,[145] which marked a significant turning point in the Supreme Court's Free Exercise Clause jurisprudence. *Smith* established the rule that individuals whose religious freedoms were abridged by the government were left without constitutional remedy so long as the offending law was deemed "generally applicable" and otherwise "valid and neutral."[146] In the wake of *Smith*, the government could more easily burden free exercise rights, a possibility which angered many members of Congress. In 1993, Congress responded by enacting the Religious Freedom Restoration Act (RFRA), which re-established the Court's pre-*Smith* compelling interest test.[147] A scant four years later, the Supreme Court offered its retort, ruling RFRA unconstitutional as applied to the states because of the limitations of sections of the Fourteenth Amendment.[148] Today RFRA remains inapplicable to the states, but it does apply to the federal government.[149]

---

[142] *See infra* notes 207–217 and accompanying text.

[143] *See* U.S. CONST. amend. I.

[144] *See* Sherbert v. Verner, 374 U.S. 398 (1963); Wisconsin v. Yoder, 406 U.S. 205 (1972).

[145] 494 U.S. 872 (1990).

[146] *Id.* at 878.

[147] *See* 42 U.S.C. §§ 2000bb to 2000bb-1 (2000).

[148] *See* City of Boerne v. Flores, 521 U.S. 507 (1997).

[149] *Boerne* merely overruled RFRA's application to the states under section 5 of the

According to its text, the purposes of RFRA are:

> (1) to restore the compelling interest test as set forth in *Sherbert v. Verner*, 374 U.S. 398 (1963) and *Wisconsin v. Yoder*, 406 U.S. 205 (1972) and to guarantee its application in all cases where free exercise of religion is substantially burdened; and
>
> (2) to provide a claim of relief or defense to persons whose religious exercise is substantially burdened by government.[150]

RFRA continues to state that:

> (a) In General: Government shall not substantially burden a person's exercise of religion even if the burden results from a rule of general applicability, except as provided in subsection (b).
>
> (b) Exception: Government may substantially burden a person's exercise of religion only if it demonstrates that application of the burden to the person—
>
> (1) is in furtherance of a compelling governmental interest; and
>
> (2) is the least restrictive means of furthering that compelling governmental interest.[151]

Finally, RFRA applies to "all Federal law, and the implementation of that law, whether statutory or otherwise, and whether adopted before or after the enactment of this Act."[152]

14th Amendment. *Id.* at 527–29. But RFRA's applicability to the federal government is derived from Congress's Article I powers, specifically the Necessary and Proper Clause embodied in Article I, Section 8 of the Constitution. *See* O'Bryan v. Bureau of Prisons, 349 F.3d 399, 401 (7th Cir. 2003). Every federal Court of Appeals that has squarely addressed this issue has decided that RFRA applies to the federal government. *See id.* at 400–01; Guam v. Guerrero, 290 F.3d 1210, 1221 (9th Cir. 2002); Henderson v. Kennedy, 265 F.3d 1072, 1073 (D.C. Cir. 2001); Kilkumura v. Hurley, 242 F.3d 950, 958 (10th Cir. 2001); Christians v. Crystal Evangelical Free Church, 141 F.3d 854, 856 (8th Cir. 1998). Moreover, RFRA raises the interesting question of whether one Congress can bind a later Congress in matters of statutory interpretation. *See, e.g.*, Community-Service Broadcasting of Mid-America, Inc. v. FCC, 593 F.2d 1102, 1116 (D.C. Cir. 1978) (Congress is "generally free to change its mind; in amending legislation Congress is not bound by the intent of an earlier body.") However, since RFRA was passed long after the pertinent IRC provisions, that issue does not arise here.

[150] 42 U.S.C. § 2000bb(b) (2000).

[151] *Id.* at §§ 2000bb(a)–(b).

[152] *Id.* at §§ 2000bb-3(a). Note that the question of binding future Congresses does not exist here. *See supra* note 149.

2. *Reading RFRA*

RFRA's commands are quite direct. The statute requires that courts apply the compelling interest test as set forth in both *Sherbert* and *Yoder*.[153] Despite this, courts have rarely applied the *Sherbert* test in its substance to the tax law.[154] Still, since RFRA was enacted in 1993, five federal Courts of Appeals have reviewed RFRA challenges to tax laws.[155] In each case, the court applied a watered-down version of the compelling interest test, drawn from *Lee v. United States*[156] or *Hernandez v. Commissioner*.[157] The courts then held that the government was not required to accommodate the free exercise right in question.[158]

On the surface, the compelling interest tests detailed in *Lee* and *Hernandez* are quite similar to those in *Sherbert* and *Yoder*.[159] But a closer examination reveals an important point of difference. Neither *Lee* nor *Hernandez* requires the government to make an affirmative showing that the disputed free exercise burden is in furtherance of a compelling government interest.[160] And the progeny of *Lee* have gone a step further, relying upon dicta from *Lee* to establish an unequivocal rule that the government interest in collecting taxes trumps all conflicting free exercise rights.[161] By subtly reducing the government's burden in this way, the tests in *Lee* and *Hernandez* are unfaithful to RFRA.

When applying RFRA, this marked divergence from the *Sherbert/Yoder* compelling interest test is unacceptable for two reasons. First, this interpretation directly conflicts with RFRA's statutory language.[162] The statutory text specifically names the *Sherbert* test and requires the government to "demonstrate" that it possesses a compelling interest and that the law in question is the least restrictive means of furthering that interest. Neither *Lee* nor *Hernandez* is consistent with that test. Second, the *Lee/Hernandez*

---

[153] *See* Sherbert v. Verner, 374 U.S. 398 (1963); Wisconsin v. Yoder, 406 U.S. 205 (1972). The compelling interest test in *Sherbert* and *Yoder* is the test discussed immediately above in the RFRA excerpt.

[154] *See* Michelle O'Connor, *The Religious Freedom Restoration Act: Exactly What Rights Does It "Restore" in the Federal Tax Context?*, 36 ARIZ. ST. L.J. 321, 329 (2004).

[155] *Id.* at 363. *See* United States v. Indianapolis Baptist Temple, 224 F.3d 627 (7th Cir. 2000); Branch Ministries v. Rossotti, 211 F.3d 137 (D.C. Cir. 2000); Browne v. United States, 176 F.3d 25 (2d Cir. 1999); Adams v. Comm'r, 170 F.3d 173 (3d Cir. 1999); Droz v. Comm'r, 48 F.3d 1120 (9th Cir. 1995).

[156] 455 U.S. 252 (1982).

[157] 490 U.S. 680 (1989).

[158] *See* O'Connor, *supra* note 154, at 362.

[159] Compare *Sherbert*, 374 U.S. at 403 (requiring an affirmative showing of a "compelling state interest" in order for the state to infringe upon free exercise rights) and *Yoder*, 406 U.S. at 221–22 (same) with *Hernandez*, 490 U.S. at 699–700 (requiring that a "substantial government interest" justify any burden placed on free exercise) and *Lee*, 455 U.S. at 258 (holding that "the state may justify a limitation on religious liberty by showing that it is essential to accomplish an overriding governmental interest").

[160] *See Lee*, 455 U.S. at 258; *Hernandez*, 490 U.S. at 700.

[161] *See* O'Connor, *supra* note 154, at 362.

[162] *See id.* at 377.

approach "thwarts the purpose for which RFRA was enacted—namely, to afford additional protection to free-exercise rights."[163] By allowing the government's interest in collecting taxes to serve as an unassailable trump card, the courts reduce the protections afforded by RFRA to something more akin to *Smith*, the decision that prompted Congress to pass RFRA in the first place. Consequently, any analysis of the tax laws' congruence with RFRA should utilize the test espoused in *Sherbert* and *Yoder*, as Congress has mandated.[164]

B. Applying RFRA to the Campaign Prohibition

In order to apply the compelling interest test to the political campaign prohibition we must make three inquiries: first, whether the statute in question substantially burdens the free exercise of religion; second, if so, whether the government has a compelling interest that is furthered by the statute; and third, if so, whether the statute is the least restrictive means by which the government could accomplish its stated compelling interest.

1. Substantial Burden Inquiry

As the Supreme Court noted many years ago, "the power to tax the exercise of a privilege is the power to control or suppress its enjoyment."[165] Moreover, the government may not "deny a benefit to a person because

---

[163] *Id.*

[164] Even if courts were not required to apply the *Sherbert* test in this context, the application of the *Lee* or *Hernandez* standards to the campaign activity prohibition might still sound the prohibition's death knell. In *Hernandez*, the Court explained that it need not examine the substantiality of the burden at issue because "our decision in *Lee* establishes that even a substantial burden would be justified by the 'broad public interest in maintaining a sound tax system,' free of 'myriad exceptions flowing from a wide variety of religious beliefs.'" Hernandez v. Comm'r, 490 U.S. 680, 699–700 (1989). Essentially, the Court adopted a position that would end most challenges to the tax laws before they even began. By asserting that the government's interest in preserving a sound tax system outweighed the substantial burdens imposed on individuals' free exercise rights, the Court obviated the need to further consider the infringed free exercise rights. But the present case can be easily distinguished from more paradigmatic *Lee* or *Hernandez* case. In most cases, the free exercise complaint seeks remediation in the form of an exemption from a preexisting tax obligation. Allowing such exceptions for every burdened religion might, as *Lee* suggests, lead to an inefficient system plagued with myriad exceptions that could greatly hinder administration. But challenges to the political campaign prohibition raise an entirely different sort of free exercise claim. Instead of seeking a tax exemption, this challenge seeks an exemption from a law prohibiting certain behaviors. Thus, unlike *Hernandez* or *Lee*, the accommodation of the free exercise rights in question would not require the sort of exceptions to the tax code that the Court feared. The administration of the tax system would remain exactly as is and no new duties would be imposed on the government. As a result, any such challenge to the campaign activity prohibitions would require a Court to engage with the claim's merits instead of summarily rejecting it for fear of complicating the administration of the tax code.

[165] Jimmy Swaggart Ministries v. Bd. of Equalization, 493 U.S. 378 (1990) (quoting Murdock v. Pennsylvania, 319 U.S. 105 (1943)) (internal quotations omitted).

he exercises a constitutional right."[166] By denying tax exemptions to churches because they speak out on political matters, the campaign activity prohibition punishes them for the free exercise of religion. The prohibition is thereby the functional equivalent of fining churches for religious activity.[167]

The question, then, is whether the campaign activity prohibition substantially concerns the free exercise of religion. It does. The campaign prohibition creates a false dichotomy, separating the words and actions of churches into two spheres: the purely political and the purely religious. In so doing, the tax laws fundamentally misrepresent what constitutes religious activity. Religion is not, and has never been, a purely academic exercise, focused only on scriptural exegeses. It is, and remains, a socially oriented endeavor, which many sects believe requires certain social obligations. German theologian Johann Baptist Metz described the social obligation that undergirds the church's political involvement as follows: the "eschatological promises of biblical tradition—liberty, peace, justice, reconciliation—cannot be made private. They force one ever anew into social responsibility."[168] This description is congruent with the historical interaction between churches and American society: "[a]s long as anyone can remember, churches have raised society's consciousness regarding political issues. They comment on the culture, rebuke its leaders, and boldly denounce its mores, as they deem necessary."[169] In the past two centuries alone, churches have played vital roles in myriad political struggles, including slavery, taxation, women's suffrage, prohibition, civil rights, war, weapons of mass destruction, capital punishment, and abortion.[170]

Different religious traditions interpret their social mandates in different ways. Some religious traditions shun involvement in the secular world. Some believe that they bear social and political obligations, but eschew involvement in partisan politics.[171] Still others clearly compel their leaders and adherents to involve themselves in the political realm. Consider the examples of the Presbyterian Church (USA) and Black Churches in America.[172]

---

[166] Regan v. Taxation With Representation of Washington, 461 U.S. 540, 545 (1983).

[167] *See* Speiser v. Randall, 357 U.S. 513, 518 (1958).

[168] Richard W. Garnett, *A Quiet Faith? Taxes, Politics, and the Privatization of Religion*, 42 B.C. L. REV. 771, 798 (2001) (quoting JOHANNES B. METZ, THEOLOGY OF THE WORLD 153 (Glen-Doepel trans., 1969)) (internal quotations omitted).

[169] Johnson, *supra* note 55, at 882.

[170] *Id.*

[171] *See* Larry B. Stammer, *Partisanship in the Pulpit Can Be Election-Year Issue; Mingling Political and Religious Messages Can Have a Polarizing Effect. Some Members of the Clergy Find They Are Walking a Fine Line.*, L.A. TIMES, Oct. 16, 2004, at B2. The mixed feelings that churches hold about partisan activity is reflected in an August poll by the PEW Research Center and the PEW Forum on Religious and Public Life. In that poll, sixty-five percent of adults were opposed to congregations endorsing political candidates. Only twenty-five percent approved while ten percent had no opinion. *Id.*

[172] Note that I have borrowed the term "African American" or "Black" Church from the

According to the policy statement of the Presbyterian Church (USA): "It is a limitation and denial of faith not to seek its expression in both a personal and public manner, in such ways as will not only influence but transform the social order. Faith demands engagement in the secular order and involvement in the political realm."[173] As a consequence, the Presbyterian Church does not view its activity in the political realm as divorced from its religious goals. Rather, it believes such involvement to be religiously motivated and compelled.

Black churches in America have an even more well-established tradition of political action, both reformist and radical.[174] This tradition has shown its prominence in the churches' role in the Underground Railroad and various abolitionist movements; by African American clergy seeking political office; by the churches' leading and organizing the Civil Rights Movement; and in the churches' mobilization of voters and provision of fora for political candidates to address members of the African American community.[175] These political activities "stemmed from the liberation tradition of the heritage of black churches," which came into existence during the time of slavery and was fueled by the churches' "own interpretations of Old Testament stories, prophetic pronouncements, and New Testament apocalypse."[176]

"[B]lack churches have a long tradition of involvement in electoral politics"[177] and "[i]t has been a continuous tradition for black churches to let both black and white politicians speak from the pulpit during political campaigns."[178] This tradition of political involvement also emerges in the unique political obligations borne by African American clergy members. Because the clergy is financially independent and does not require financial assistance from outside the church community, there is an expectation that they will use this independence to speak out about pressing political issues, especially when others might shy away from public pronouncement.[179]

History, tradition, and scriptural interpretation compel both the Presbyterian Church (U.S.A.) and the African American Church to in-

---

authors cited and intend to use it in the same sense: to refer to the seven independent, historical, totally African American–controlled denominations that were founded after the Free Africa Society of 1787. These include the African Methodist Episcopal Church; the African Methodist Episcopal Zion Church; the Christian Methodist Episcopal Church; the National Baptist Convention, U.S.A., Incorporated; The National Baptist Convention of America, Unincorporated; the Progressive National Baptist Convention; and the Church of God in Christ, along with a handful of smaller communions. *See generally* C. ERIC LINCOLN & LAWRENCE H. MAMIYA, THE BLACK CHURCH IN THE AFRICAN-AMERICAN EXPERIENCE 202 (1990).

[173] PRESBYTERIAN CHURCH (U.S.A), GOD ALONE IS LORD OF THE CONSCIENCE: A POLICY STATEMENT ADOPTED BY THE 200TH GENERAL ASSEMBLY 48 (1998).

[174] *See generally* LINCOLN, *supra* note 172.

[175] *See* James, *supra* note 13, at 65–66.

[176] LINCOLN, *supra* note 172, at 202.

[177] *Id.* at 215.

[178] *Id.* at 206.

[179] *Id.* at 207.

volve themselves in politics in order to push for certain social goals. The scope of this obligation is a matter of religious faith and interpretation and, for obvious reasons, one best determined by religious and not secular leaders. To assert that the fulfillment of this political obligation may require issue advocacy but not candidate advocacy, or to claim that one is "religious" while the other purely political is nothing but sophistry.

The social obligations of the church compel it to advance certain social agendas. And the nature of religious leadership requires priests and ministers to discern which social goals are religiously compelled and to support those goals. For many ministers, the social values and goals that they view to be of paramount importance to be embodied by certain political parties. Churches that prioritize the preservation of human life at all stages and forms may legitimately feel that their religion requires them to support the party that opposes abortion, and churches who prioritize social equality may legitimately feel that their religion compels them to endorse the party that most nearly supports that aim. To punish churches for taking the obvious step to endorse certain candidates or parties is to limit the churches' abilities to convey a religiously compelled message. A minister would be acting in no less of a religious capacity because he endeavors to end the oppression of racial minorities by endorsing Kennedy, a Democrat, than if he did so by supporting the Civil Rights Act. Indeed, his speech might be more effective if directed toward helping a candidate. Thus, campaign activity can be a religious exercise.

The campaign activity prohibition thus burdens the free exercise of religion in at least two ways. First, and most importantly, it impermissibly conditions the receipt of a benefit—namely tax-exempt status—upon the forfeiture of free exercise of religion. Second, it allows the government to remove religious leaders' control over what activities and beliefs constitute religion.

This burden is magnified by the hazy standard adopted in the current statute. Because ministers are forbidden from backhandedly endorsing candidates through code words such as "pro-choice," some ministers are likely to shy away from discussing social issues that are important to them, even though doing so is perfectly legal under current law. As one California minister explained, some clerics possess an imperfect understanding of the separation of church and state. "Unfortunately, too many preachers let that scare them from preaching about the importance of religious values, which is always a political stance,"[180] but not necessarily a partisan one. Consequently, these ministers limit their discussion of a large sphere of issues because they fear losing their church's tax-exempt status. That "chilling effect" is a burden on the free exercise of religion.

While it seems clear that the prohibition burdens free exercise rights, the question remains whether the burden is substantial. After all, the gov-

[180] Stammer, *supra* note 171.

ernment does not forbid churches from engaging in partisan activity, which would be an obvious First Amendment violation. Rather, the government only conditions the receipt of tax-exempt status on the churches' abstention. The D.C. Circuit raised two related justifications in explaining why the campaign activity prohibition was permissible. First, the court noted that simply decreasing the amount of money available to a church for its religious practices did not rise to the level of a constitutionally cognizable burden.[181] Second, the court explained that the federal government's decision not to subsidize an organization's First Amendment activities does not constitute a violation of the organization's First Amendment rights.[182] I contend that while the D.C. Circuit cited the proper legal standard, it failed to apply it correctly to the facts at hand.

With respect to the D.C. Circuit's first argument, the court failed to grasp a fundamental distinction between the *Hernandez v. Commissioner*[183] and *Jimmy Swaggart Ministries v. Board of Equalization*[184] lines of cases, from which the precedent is derived, and the I.R.C.'s political prohibitions. The assertion in both *Swaggart* and *Hernandez*, that the tax burdens considered therein were not constitutionally problematic, stemmed from the fact that those burdens did nothing more than change the amount of money available for religious activity.[185]

The campaign prohibition functions quite differently. Instead of merely causing an across-the-board decrease in available resources, the campaign-activity prohibition penalizes churches for holding, and acting upon, a specific religious belief. In *Swaggart* and *Hernandez*, the tax concerns did not affect the content of the organization's belief or practice. They affected only the quantity of the activity, not its fundamental character.

In the present case, the campaign prohibition actually warps religious practice and belief. By only withholding tax-exempt status for participating in certain activities, the government has ensured that most churches will conform their religious behavior to the government's preferences. The actual content of the religious message delivered by churches has been

---

[181] Branch Ministries v. Rossotti, 211 F.3d 137, 142 (D.C. Cir. 2000).

[182] *Id.* at 143–44.

[183] 490 U.S. 680, 699 (1989).

[184] 493 U.S. 378, 391 (1990).

[185] *See Swaggart*, 493 U.S. at 391; *Hernandez*, 490 U.S. at 699. In *Swaggart*, the court ruled on the constitutionality of a California law requiring retailers to collect a six percent sales tax for sales of certain goods. The court held that the tax was constitutional since the only burdens it placed upon Jimmy Swaggart Ministries were a reduction of income, resulting from a decrease in consumer demand for goods (which presumably would be more expensive with the tax) and administrative costs. The *Hernandez* case presented a functionally similar scenario. In *Hernandez*, the issue was whether the IRS must allow individuals who practice Scientology to deduct monies paid to the church in return for auditing services. The court ruled that the IRS was correct, and that there was no constitutionally cognizable burden because the refusal to grant a tax deduction merely decreased the amount of money available to spend on auditing. As in *Swaggart*, the tax law in question did not punish the church for certain beliefs it held. It merely reduced the amount of resources available for religious activity in general.

impacted, not just the amount that they can speak. This distinction is evident in *Swaggart*, where the court stated that "the sales and use tax is not a tax on the right to disseminate religious information, ideas, or beliefs *per se*; rather, it is a tax on the privilege of making retail sales on tangible personal property . . . ."[186] Neither *Swaggart* nor *Hernandez* dealt with a tax on the right to disseminate religious information, ideas, or beliefs. The tax-campaign prohibition does. Churches that convey a religious message discouraged by the state are taxed; churches who abstain are not. Consequently, the prohibition imposes a substantial burden on religious belief and cannot be cast aside as simply reducing the amount of resources available for religious activities.[187] For example, now liable for taxes, churches would face new administrative burdens. Some churches undeniably do have substantial income from sources other than donations, and loss of tax-exempt status would cause them significant financial injury. Moreover, there is a clear symbolic value at stake: loss of tax-exempt status suggests government disfavor and, potentially, the calling into question of the church's religious and charitable identity. The tangible nature of these burdens, at least for some churches, probably explains why most ministers adhere to the letter of the law and show real fear of losing their tax-exempt status.

The campaign-activity prohibition can be distinguished from *Swaggart* in another way. Prior to *Swaggart*, the court had held that licensing fees that must be paid by religious organizations before they engage in religious activities were unconstitutional.[188] In *Swaggart*, the court limited its holding in the earlier cases to flat license taxes that served as prior restraints, noting the dissimilarity between such taxes and the taxes considered by the court in *Swaggart*.[189] The campaign-activity prohibition, however, bears more resemblance to the unconstitutional licensing taxes than it does to the sales tax at issue in *Swaggart*. Like the licensing taxes, the campaign prohibition acts like a prior restraint. In order to express certain beliefs, a church must surrender its tax-exempt status. The

---

[186] *Swaggart*, 493 U.S. at 389.

[187] One additional critique is that the burden on churches is superficial; the loss of tax-exempt status may not inflict real financial injury on churches. *See* Ellen P. Aprill, *Churches, Politics, and the Charitable Contribution Deduction*, 42 B.C. L. REV. 843, 844–46 (2001); Feld, *supra* note 123, at 936. Church revenue is derived primarily from individual donors. In 1996, religious congregations received $81.2 billion in total revenue. Aprill, *supra*, at 844. Of that figure, $68 billion came from private donations, ninety-four percent of which came from individuals. *Id.* The tax code does not consider private donations to be income, and as such they are non-taxable. *See* 26 U.S.C. § 170 (2000). Consequently, even if a church were to lose its § 501(c)(3) status, it still would not be required to pay income taxes on donations from congregants. Feld, *supra*, at 936. And because many individuals do not take their deduction for donating to churches anyway, *see* Aprill, *supra*, at 845–46, they would not be impacted by the change in tax status. However, such an argument misses the point.

[188] Follett v. Town of McCormick, 321 U.S. 573 (1944); Murdock v. Pennsylvania, 319 U.S. 105, 108–09 (1943).

[189] *Swaggart*, 493 U.S. at 386–87.

second a church engages in forbidden activity, it no longer qualifies as a tax-exempt entity under § 501(c)(3). Thus, churches are prevented—at least much of the time—from exercising their First Amendment rights by a prior restraint. Admittedly, enforcement issues may make the campaign-activity prohibition a more permeable restraint than a flat licensing tax, but the difference is not as substantial as it may appear. Much as an itinerant minister could preach without the benefit of a license until discovered by the authorities, churches today endorse political candidates and engage in political activity until discovered by the IRS. The mechanics may differ, but the principle is the same. The tax laws serve as a prior restraint on the free exercise of religion, just like the licensing taxes ruled unconstitutional by the Court in *Murdock*[190] and *Follett*.[191]

The D.C. Circuit's misunderstanding of the campaign-finance prohibition is also reflected in its misguided application of *Regan v. Taxation with Representation* in *Branch Ministries v. Rossotti*.[192] In *Regan*, the court heard a challenge to the constitutionality of an I.R.C. provision banning lobbying by tax-exempt organizations. The court ruled that while the government may not deny a benefit to a person for the exercise of a constitutional right, the government is not required to subsidize First Amendment activity.[193] The court then proceeded to uphold the constitutionality of the provision. However, as three judges of the *Regan* court recognized, and the rest of the court later accepted,[194] the availability of alternate means of communication was essential to the constitutionality of § 501(c)(3)'s lobbying restrictions.[195] If no alternate means of communication existed, then "an otherwise eligible organization [would be deprived] of its tax-exempt status and its eligibility to receive tax-deductible contributions for all its activities, whenever one of those activities is 'substantial lobbying.'"[196] And because lobbying is protected by the First Amendment, the lack of an alternate means of communication would ensure that the I.R.C. denied "a significant benefit to organizations choosing to exercise their

---

[190] *Murdock*, 319 U.S. at 108–09.

[191] *Follett*, 321 U.S. at 573.

[192] *See Branch Ministries*, 211 F.32 137, 143, *citing* Regan v. Taxation with Representation, 461 U.S. 540, 552–53 (1983).

[193] *See id.* at 545–46.

[194] *See* FCC v. League of Women Voters, 468 U.S. 364, 400–01 (1984).

[195] *Regan*, 461 U.S. at 552–53 (Blackmun, J., concurring). The Court subsequently confirmed that this was an accurate description of its holding. *See FCC*, 468 U.S. at 400. Also, note that neither of the seminal Supreme Court cases allowing government to impose conditions on federal funding are applicable here. Rust v. Sullivan, 500 U.S. 173, 197–98 (1991), endorses the *FCC* Court's analysis regarding the need for alternate avenues of communication, and South Dakota v. Dole, 483 U.S. 203 (1987), only addresses the federal government's ability to condition the payment of funds to states on compliance with requirements the federal government could not otherwise impose.

[196] *Regan*, 461 U.S. at 552 (Blackmun, J., concurring).

constitutional rights,"[197] a result prohibited by *Speiser v. Randall* in the absence of a compelling government interest.[198]

The crucial distinction to be drawn between *Regan* and the situation confronted here is that the revocation of tax-exempt status is not simply a refusal by the IRS to subsidize church speech, but rather serves as a punitive measure.[199] In *Branch Ministries*,[200] the Church relied on *Regan* to argue that its free exercise rights were substantially burdened because it lacked an alternate means by which to express its opinions about candidates.[201] The D.C. Circuit responded that the church retained an alternate avenue for expression, because it could set up a parallel § 501(c)(4) organization, which could then establish a PAC to communicate about candidates.[202] While the D.C. Circuit's assertion may be formally correct, it misses the fundamental point of Regan's "alternative means" requirement: namely, it fails to require that the institution in question retained a viable alternative to the channel of communication that had been burdened. If no viable alternative exists, then the constitutional problem cannot be remedied unless the burden is removed.

Because of the requirements imposed on the creation and maintenance of these additional organizations, communications from the PAC could not be communications from the church itself.[203] While the requirements imposed on a 501(c)(3)'s ability to create downstream organizations are somewhat lax, they would still preclude the church from making related communications within the church itself and would almost certainly preclude the minister or priest from speaking on behalf of the PAC, whether within or outside the church. They would be, by legal necessity, communications from a different entity, albeit one that maintained some connection with the church. As a consequence, the church would be stripped of its religious voice, an outcome not permitted by *Regan*.[204]

Both the church's free exercise rights and its religious/political message are bound up in the identity of the speaker to a unique degree. Unlike most organizations, churches communicate first and foremost through their ministers' speaking to an assembled congregation. The religious character of the organization and the moral gravitas that defines it derive from this configuration. By requiring churches to communicate all religious messages with partisan connotations through a PAC that cannot even be established by the church, the connection to the church of any message issued will be attenuated; it will certainly no longer be religious. There is,

---

[197] *Id.*

[198] 357 U.S. 513, 528–29 (1958).

[199] *See* Regan v. Taxation with Representation, 461 U.S. 540, 546 (1983); *F.C.C.*, 468 U.S. at 400, for a discussion of the subsidy / penalty distinction.

[200] Branch Ministries v. Rossotti, 211 F.3d 137, 143 (D.C. Cir. 2000).

[201] *Branch Ministries*, 211 F.3d at 143.

[202] *See supra* notes 132–136 and accompanying text.

[203] *See Regan*, 461 U.S. at 552–53 (Blackmun, J., concurring).

[204] *See id.* at 552.

quite simply, no adequate substitute available for that form of communication. Consequently, the burden placed upon churches is very substantial indeed.

The campaign-activity prohibition also substantially burdens religion and religious institutions by encroaching upon the ability of the church to define what is and what is not religious. By co-opting this authority from the church, and defining which church actions are religious and which non-religious, the government may "subtly reshape[ ] religious consciousness itself. In other words, by telling religion what it may say, . . . , and by telling faith where it belongs, government [may] mold[ ] religion's own sense of what it is."[205] The First Amendment "stands as an expression of principle on the part of the Founders of our Constitution that religion is too personal, too sacred, too holy, to permit its 'unhallowed perversion' by a civil magistrate."[206] Thus, when a civil magistrate—here the IRS—gets involved in the unhallowed business of defining and regulating religious content, the church's free exercise rights are, almost by definition, substantially burdened.

2. *Compelling Interest Inquiry*

Because of the lack of legislative history,[207] the compelling interest inquiry is somewhat less clear than it could be. However, while Congress has not specifically articulated what it believes the interest at stake to be, three main arguments are used to explain why the government has a compelling interest in the campaign-activity prohibition. The first argument asserts that the prohibition is required in order to maintain a tax system that can be easily administered without allowing myriad exceptions for different religious groups.[208] The second argument is that the prohibition "reflect[s] Congressional policies that the U.S. Treasury should be neutral in political affairs."[209] The final argument is that the prohibition is necessary to ensure that tax-deductible money is not used for partisan activity.[210] While several of these arguments suggest that Congress may have a compelling interest in tempering churches' *financial* involvement in electoral politics, none of them provides a compelling justification for the prohibition of activities that do not involve the expenditure of tax-exempt monies.

---

[205] Garnett, *supra* note 168, at 796.

[206] Engel v. Vitale, 370 U.S. 421, 431–32 (1962).

[207] *See supra* notes 55–67 and accompanying text.

[208] *See* Hernandez v. Comm'r, 490 U.S. 680, 699–700 (1989).

[209] H.R. REP. No. 100-391, pt. 2, at 1625 (1987), *as reprinted in* 1987 U.S.C.C.A.N. 2313-1205.

[210] *See Hernandez*, 490 U.S. at 699–700.

The first argument is the most easily dismissed.[211] Simply stated, ending the campaign activity prohibition would not create any additional exceptions to the administration of the tax code. Churches would be allowed to engage in additional behavior, but the state would take on no new administrative duties, and no additional exemptions would be created.[212] The tax system would continue to exist exactly as it does now without additional burdens on the government.[213]

The second argument also may be untenable, as it presupposes that removing the political-activity prohibition would force the U.S. Treasury to assume some sort of non-neutral role in U.S. politics. However, this assumption is belied by the tremendous demographic differences among churches. Although it may be convenient to group churches as a block, there is no reason to think that institutions that are so diverse along such a range of dimensions (geographical location, racial demographics, etc.) would show a predictable bent toward any particular political party.[214] Even if they did, this argument would still be flawed, because it supports a prohibition on churches spending tax-exempt funds for partisan activity, not a prohibition on all forms of activity. Allowing churches to participate in putatively partisan activity would not place the Treasury at risk of non-neutrality if it forbade expenditures of tax-exempt monies for the activity.

Even if the above reasons were insufficient justifications for drawing the line at the expenditure of tax-exempt funds, the Establishment Clause would provide an additional reason for drawing this line. Over a half-century ago, the Supreme Court explained that the Establishment Clause means that "[n]either [a state nor the Federal Government] can pass laws which aid one religion, aid all religions, or prefer one religion over another."[215] Though time has seen the Court alter the test it uses to examine Establishment Clause claims, those words remain true today. As the Court made clear in *Lemon v. Kurtzman*,[216] and reaffirmed many times since,[217] the government cannot pass a law that has the primary effect of advanc-

---

[211] *See supra* note 164 and accompanying text.

[212] There is also no reason to think that allowing churches to become involved in partisan activity would involve additional enforcement needs. The IRS already avoids closely monitoring church activity, and it could continue to do so under a more lax standard. And by drawing the line at financial expenditures, IRS enforcement actually becomes a simpler, more objective process with an easily enforceable, bright-line rule.

[213] Nor would this create the need to make more exceptions for other religious organizations at a later date. *See* discussion *supra* note 164.

[214] And even if churches as a whole tended to side more with one party than another, this general bias might not be reflected in the political activity engaged in by churches as a block. For instance, even if more churchgoers were to hold Republican sympathies, it is possible that the most active group of churches would be African American churches that held Democratic sympathies.

[215] Everson v. Bd. of Educ., 330 U.S. 1, 15 (1947).

[216] 403 U.S. 602, 612 (1971) (citing Bd. of Educ. v. Allen, 392 U.S. 236, 243 (1968)).

[217] *See, e.g.*, Zelman v. Simmons-Harris, 536 U.S. 639 (2002); Mitchell v. Helms, 530 U.S. 793 (2000); Agostini v. Felton, 521 U.S. 203 (1997).

ing (or inhibiting) religion. Changing the tax laws to allow religious institutions a government-subsidized financial advantage over non-religious organizations in the political sphere would clearly advance religion and would provide religious organizations with a distinct advantage over their secular counterparts in the political sphere. Consequently, RFRA and the Establishment Clause serve to carve out statutory and constitutional parameters for the government regulation of politico-religious activity by churches. RFRA requires that churches be allowed to communicate religious messages—even those with a political position—without fear of financial penalty, while the Establishment Clause forbids the government from allowing churches to use their tax-exempt funds to engage in partisan activities.

V. PROPOSED SOLUTIONS

A. A New Bright-Line Rule

Carving out a clearly delineated safe harbor for politico-religious activity will alleviate pressing First Amendment concerns much more successfully than either the status quo or the leading alternatives.[218] While the status quo cripples enforcement efforts with interpretive ambiguities and the most discussed alternatives either fail to cure the statute's free exercise infirmities or raise new First Amendment concerns, the safe harbor proposal both palliates the existing free exercise problems and distills enforcement into a simple matter of whether or not money was spent on political activity.

As currently written, the campaign activity prohibition requires the IRS to undertake a highly subjective, fact-intensive inquiry in order to determine what behavior is acceptable and what behavior is not.[219] By instead defining the limit of permissible partisan activity to end at the expenditure of tax-exempt money for partisan purposes, the IRS could create a bright-line rule that would serve two purposes. First, it would provide churches with better guidance so that they may more successfully comply with the law, thus reducing both the chilling effect of the current law and the incidence of unintentional violations. Second, it would allow the IRS to determine more accurately which actions are and are not violations.[220] Consider a few of today's most vexing concerns and how

---

[218] *See infra* Part V.B.

[219] *See supra* notes 89, 125 and accompanying text.

[220] Of course, the IRS would still be required to make determinations involving which actions are partisan and which not partisan, but the IRS would have less cause to avoid enforcing clear violations of the law. Currently, the IRS ignores most violations, presumably because they seem so minor and the revocation of a church's tax-exempt status so disproportionate a punishment. Under the proposed interpretation, several of the most difficult problems would be resolved because the IRS would be addressing only the expenditure of tax-exempt funds. Moreover, many new gray area concerns could be alleviated by the addition of a *de minimis* spending exception. Under such a rule, *de minimis* expenditures such

easily they would be resolved under the new standard. Candidate endorsement from the pulpit would be clearly permissible, as would the distribution of voter guides produced by outside organizations, irrespective of their political bent. Candidate forums and visits also would be acceptable in most situations. Fundraising by churches, donations to campaigns, and partisan advertisements would all be forbidden.[221] And if the IRS did enforce these laws, there would almost certainly be little outcry, given the Establishment Clause concerns that support the IRS's position. In practice, the prohibition would exclude approximately the same range of activities, but churches would no longer be chilled from fully enjoying their free exercise rights, and churches would no longer be rewarded for violating the letter of the law. Enforcement would be scaled back little, if at all, and churches would be free to exercise their religion freely without the specter of financial punishment looming over them.[222]

B. Alternative Proposals

There are three other alternatives that have been discussed: the application of federal election disclosure rules, the adoption of a "substantial part" test, and the Crane-Rangel Amendment's five percent rule.[223] None of these alternatives, however, address adequately both the constitutional and pragmatic policy concerns raised by the campaign-activity prohibition.

Federal election disclosure rules require an organization that speaks out on behalf of a "clearly identified" electoral candidate to disclose its expenditures if the speech is considered to be "express advocacy."[224] A similar rule could be used in the church context, allowing the church to engage in political activities without fear of IRS reprisal as long as the activities did not amount to "express advocacy." However, this shift in standards would hardly solve the problem. "Express advocacy" is still not a bright-line test; thus, enforcement and adherence would remain problematic[225] and would not be permissive enough to comply with RFRA.

---

as the allocation of meeting space or the waiver of small fees would not pose enforcement difficulties.

[221] While it may seem that churches could involve themselves in limited activities without the use of tax-exempt funds, the risk of commingling tax-exempt and non-tax-exempt funds makes a more prophylactic prohibition on churches spending money for partisan purposes a desirable addition to the law.

[222] This is also important because a church's ability to freely exercise its religion would no longer be dependent on its financial status or degree of risk aversion.

[223] Erik J. Ablin, *The Price of Not Rendering to Caesar: Restrictions on Church Participation in Political Campaigns*, 13 NOTRE DAME J.L. ETHICS & PUB. POL'Y 541, 585–86 (1999) (discussing various proposals for removing the restrictions on church campaign activity).

[224] *Id.* at 583.

[225] For a discussion of the vagaries confronted by courts in the application of analogous tests, see Eugene Volokh, *A Common-Law Model for Religious Exemptions*, 46

Candidate endorsement by a minister, for instance, is protected by RFRA, but would still violate an express advocacy test.

Adopting the "substantial part" standard would allow for more church participation in political activity by applying the same § 501(c)(3) standard currently used for lobbying. A church would be allowed to involve itself in electoral activity as long as that activity did not constitute a "substantial part" of its activities.[226] This proposal has two key flaws. First, the "substantial part" inquiry is extremely subjective and invites extensive regulatory oversight, which may be problematic from an Establishment Clause perspective.[227] Second, the standard would allow churches to engage in activity, such as making donations to political campaigns, that is repugnant to the Establishment Clause so long as that activity did not constitute a substantial part of the church's activity.

The Crane-Rangel Amendment—the brainchild of Representatives Philip Crane (R-Ill.) and Charles Rangel (D-N.Y.)—would have amended § 501(c)(3) to allow churches to spend funds on political campaigning so long as their expenditures did not exceed five percent of gross revenues.[228] While the amendment would allow for the full enjoyment of free exercise rights guaranteed by RFRA, it would proceed one step further than it ought, by allowing churches to spend up to five percent of their revenues on electioneering activities. Because churches are tax-exempt organizations, their revenues are tax exempt. Therefore, the Crane-Rangel Amendment would permit churches to spend tax-exempt funds on partisan political activities. As discussed in Part IV, allowing churches, but not other charitable organizations, to make such expenditures probably would violate the Establishment Clause.[229] And while allowing all charitable organizations to make such expenditures would alleviate the Establishment Clause concerns, it would create a regulatory and administrative nightmare for both the IRS and participating agencies[230] and might also spawn entanglement concerns.

UCLA. L. REV. 1465, 1474 n.21 (1999).

[226] Ablin, *supra* note 223, at 584.

[227] Lemon v. Kurtzman, 403 U.S. 602, 612 (1971). Probing regulatory oversight may constitute excessive entanglement in the religious context. *See generally* Agostini v. Felton, 521 U.S. 203, 233 (1997).

[228] Ablin, *supra* note 223, at 585.

[229] *See* Everson v. Bd. of Educ., 330 U.S. 1, 15 (1947); Zelman v. Simmons-Harris, 536 U.S. 639 (2002); Mitchell v. Helms, 530 U.S. 793 (2000); Agostini v. Felton, 521 U.S. 203 (1997).

[230] Churches and other charitable organizations would be forced to keep strict records of all revenues, and the IRS would be required to untangle the financial records of numerous nonprofits, taxing its limited resources.

VI. CONCLUSION

Both pragmatic policy concerns and free exercise concerns support abandoning the current formulation of the campaign activity prohibition. Consider, for instance, the candidate/issue dichotomy discussed above.[231] While churches may comment on issues, they may not comment on candidates. The IRS recognizes the easily blurred distinction between candidate and issue advocacy, and it requires that its agents make a subjective evaluation of the church's religious speech to discern issue commentary from veiled candidate commentary.[232] Given the difficulty of determining whether certain speech constitutes candidate, rather than issue, advocacy and of proving the veracity of the initial, subjective determination, the IRS is understandably loath to enforce the campaign-activity prohibition in borderline cases. The same pattern is repeated in a variety of contexts: candidate forums, the distribution of voter guides by the Christian Coalition, even candidate fundraising.[233] Unless the violation is too visible to ignore and too partisan to debate,[234] the IRS seems unwilling to enforce the prohibition.

This state of affairs may be inevitable given the fundamental nature of the right at stake and the lack of a bright-line IRS rule. The IRS must choose between substantially abandoning enforcement of the rule and becoming involved in a legal quagmire in which it must fight a battle (probably public) against various religious organizations, threatening the free exercise rights of various churches armed with nothing more than its subjective interpretation of the churches' actions. While the IRS may be able to keep most churches in line by using the occasional threat and prosecuting the most egregious offenders, the upward spike in partisan activity seen in the 2004 election suggests that the IRS may be fighting a losing battle.[235] Unchecked by regulatory action, churches are likely to become more brazen in their violation of the campaign activity prohibition. And given that more vigorous enforcement of the current standard is fraught with unappealing concerns, the IRS is unlikely to engage in stricter enforcement of partisan activity unless armed with a bright-line rule upon which it can more fearlessly rely to regulate church activity.

Simply allowing churches to engage in political activity so long as they avoid spending tax-exempt money on those activities solves all of these problems. It ensures ministers and worshipers the free exercise rights

---

[231] *See supra* notes 127–128 and accompanying text.

[232] *See supra* notes 127–128 and accompanying text.

[233] *See* Branch Ministries v. Rossotti, 211 F.3d 137 (D.C. Cir. 2000), *supra* notes 91–92, 112–115 and accompanying text.

[234] One example of this type of violation is the actions taken by the Church at Pierce Creek. *See* Branch Ministries v. Rossotti, 40 F. Supp. 2d 15, 18 (D.D.C. 1999). *See also supra* notes 119–122 and accompanying text.

[235] *See supra* notes 96–99 and accompanying text.

guaranteed them by RFRA, provides a bright-line test that allows churches to adhere to IRS guidelines and for the IRS to enforce those guidelines, and avoids Establishment Clause concerns by prohibiting the use of tax-exempt dollars. Prudence recommends such a course, and RFRA and the Establishment Clause command it.

RECENT DEVELOPMENTS

THE DETENTION OF ENEMY COMBATANTS ACT

> I frankly do not know whether these tools are sufficient to meet the Government's security needs . . . . It is far beyond my competence, or the Court's competence, to determine that. But it is not beyond Congress's.[1]
>
> —Associate Justice Antonin Scalia

> If the law in its current state is found by the President to be insufficient to protect this country from terrorist plots . . . then the President should prevail upon Congress to remedy the problem.[2]
>
> —Judge Henry F. Floyd

In 1971, Congress enacted the Non-Detention Act,[3] a statutory prohibition on the federal government's detention of U.S. citizens "except pursuant to an Act of Congress."[4] At present, the Non-Detention Act sits squarely in the center of the ongoing litigation over executive detention of U.S. citizens accused of being enemy combatants in the war on terrorism. In two recent cases, *Hamdi v. Rumfeld*[5] and *Padilla v. Rumsfeld*,[6] the threshold question before the courts was whether Congress had previously authorized the detentions being contested. While the Supreme Court plurality in *Hamdi* found that Congress's Authorization for Use of Military Force[7]

---

[1] Hamdi v. Rumsfeld, 542 U.S. 507, 578 (2004) (Scalia, J., dissenting).

[2] Padilla v. Hanft, 389 F. Supp. 2d 678, 692 (D.S.C. 2005).

[3] 18 U.S.C. § 4001(a) (2000) [hereinafter the "Non-Detention Act"] (the Second Circuit adopted this short title in Padilla v. Rumsfeld, 352 F.3d 695, 699 (2d Cir. 2003), *rev'd*, 542 U.S. 426 (2004)).

[4] *Id.* 18 U.S.C. § 4001(a) ("[n]o citizen shall be imprisoned or otherwise detained by the United States except pursuant to an Act of Congress.") *Id. See also Developments in the Law—The National Security Interest and Civil Liberties*, 85 HARV. L. REV. 1130, 1317 n.133 (1972).

[5] 542 U.S. 507. Yasser Hamdi was born in Louisiana in 1980. He was captured in Afghanistan in late 2001 and transferred to Guantanamo Bay, Cuba in January 2002. In April 2002, upon learning of Hamdi's American citizenship, authorities transferred him first to a military brig in Norfolk, Virginia, and then to a brig in Charleston, South Carolina. *See id.* at 510–11.

[6] 542 U.S. 426 (2004). On May 8, 2002, Jose Padilla arrived at Chicago's O'Hare International Airport from Pakistan. Federal agents immediately detained him pursuant to a material witness warrant issued by the District Court for the Southern District of New York as part of the continuing investigation into the September 11 attacks. On May 22, Padilla's attorney moved to vacate the warrant. On June 9, while the motion was pending, the President issued an order to the Secretary of Defense designating Padilla an enemy combatant and directing the Secretary to detain him in military custody. That same day, Padilla was taken into military custody and transferred to the U.S. Navy brig in Charleston, South Carolina. *See id.* at 430–32.

[7] Pub. L. No. 107-40, 115 Stat. 224 (2001) [hereinafter "AUMF"].

(AUMF) had indeed authorized detentions of U.S. citizens captured in a *foreign* zone of combat,[8] the question of whether or not Congress has authorized detentions of citizen enemy combatants[9] captured *inside* the United States remains open.[10]

This Recent Development examines the proposed Detention of Enemy Combatants Act[11] (DECA), which would grant the President broad authority to detain any U.S. citizen deemed an "enemy combatant," regardless of the place of capture.[12] First, I briefly examine whether the AUMF already authorizes detentions of citizens captured inside the United States. While the district and circuit courts have divided sharply over this issue, Justice Stevens's dissent in *Padilla*[13] and Justice Scalia's dissent in *Hamdi*[14] provide evidence that a majority of the Supreme Court may hold that the AUMF does not authorize such detentions. Second, I examine the efficacy of DECA's due process provisions, concluding that while imposing several requirements on the executive, its provisions are either ambiguous or silent regarding key issues. As a result, I conclude that DECA is unlikely to achieve its authors' intended goal of ensuring due process for detainees.

Since this Recent Development is concerned with statutory authorization for executive detentions of citizens captured inside the United States,[15]

---

[8] *Hamdi*, 542 U.S. 519 ("[I]t is of no moment that the AUMF does not use specific language of detention. Because detention to prevent a combatant's return to the battlefield is a fundamental incident of waging war, in permitting use of 'necessary and appropriate force,' Congress has clearly and unmistakably authorized detention in the narrow circumstances considered here.").

[9] The term "enemy combatant" has consistently eluded definition. While the *Hamdi* plurality acknowledged that "there is some debate as to the proper scope" of the term, the plurality went on to state that the "permissible bounds of the category will be defined by the lower courts as subsequent cases are presented to them." *Id.* at 516. In *Ex parte Quirin*, the Supreme Court stated that "citizens who associate themselves with the military arm of the enemy government, and with its aid, guidance, and direction enter this country bent on hostile acts are enemy belligerents within the meaning of . . . the law of war." 317 U.S. 1, 37–38 (1942). *See also* David D. Coron & Jenny S. Martinez, *International Decision: Availability of U.S. Courts to Review Decision to Hold U.S. Citizens as Enemy Combatants—Executive Power in War on Terror*, 98 A.J.I.L. 782, 786 (2004) (noting that the *Hamdi* plurality gave no direction to lower courts as to where they should find the "bounds of the category"). For more discussion of the term in the context of international law, see Ryan Goodman & Derek Jinks, *International Law, U.S. War Powers, and the Global War on Terrorism*, 118 HARV. L. REV. 2653, 2654–58 (2005).

[10] In *Padilla*, the Supreme Court found that the lower courts had lacked jurisdiction to hear the case and therefore avoided reaching the merits of Padilla's claim. 542 U.S. at 430 ("We confront two questions: First, did Padilla properly file his habeas petition in the Southern District of New York; and second, did the President possess authority to detain Padilla militarily. We answer the threshold question in the negative and thus do not reach the second question presented.").

[11] H.R. 1076, 109th Cong. (1st Sess. 2005).

[12] *Id.* § 3(a). For a detailed discussion of DECA, *see infra* Part II.

[13] *See* 542 U.S. at 455 (Stevens, J., dissenting).

[14] *See* 542 U.S. at 554–99 (Scalia, J., dissenting).

[15] *See generally* Stephen I. Vladeck, Note, *The Detention Power*, 22 YALE L. & POL'Y REV. 153 (2004) (surveying the history of the President's power to detain U.S. citizens and finding that no statutes today satisfy the requirements of the Non-Detention Act's application to *Padilla* and *Hamdi*).

arguments about possible constitutional sources for the President's authority to detain are not discussed.[16]

I. THE AUMF AND DETENTION OF CITIZENS CAPTURED INSIDE THE UNITED STATES

Throughout the *Padilla* litigation, executive branch attorneys argued that Congress had already authorized detentions of citizen enemy combatants captured inside the United States when it passed the AUMF on September 18, 2001.[17] In response, each of the courts involved in the litigation thus far has come to different conclusions about whether the AUMF satisfies the requirements of the Non-Detention Act. In the first round of litigation, the District Court for the Southern District of New York found that the AUMF had authorized Padilla's detention,[18] while the Second Circuit found that it had not.[19] After the Supreme Court vacated his claims on procedural grounds, Padilla began his second round of litigation in the District Court for the District of South Carolina, which held that his detention was unauthorized,[20] The Fourth Circuit recently reversed that decision[21] and the issue is now ripe for the Supreme Court to decide on the merits.

After briefly examining each of the lower court decisions as to whether the AUMF authorized Padilla's detention, I will analyze possible outcomes in the Supreme Court. Assuming that DECA is not enacted before

---

[16] While executive branch attorneys have argued that the President's Article II war powers encompass the power to detain, courts would likely avoid deciding such a contentious issue if clear statutory authorization for detentions existed. *See, e.g.*, *Hamdi*, 542 U.S. at 516–17 ("The Government maintains that no explicit congressional authorization is required, because the Executive possesses plenary authority to detain pursuant to Article II of the Constitution. We do not reach the question whether Article II provides such authority, however, because we agree . . . that Congress has in fact authorized Hamdi's detention, through the AUMF."). *See also* Padilla v. Rumsfeld, 352 F.3d 695, 722 (2d Cir. 2003) ("Since we conclude that the Non-Detention Act applies to military detentions such as Padilla's, we would need to find specific statutory authorization in order to uphold the detention.").

[17] In addition, attorneys for the executive argued that the Non-Detention Act should be read so as to avoid conflict with President's inherent war powers. *See Padilla*, 233 F. Supp. 2d 564, 596–97 (S.D.N.Y. 2002). On appeal to the Second Circuit, executive branch attorneys also argued that funding of military detention facilities in 10 U.S.C. § 956(5) provided further proof of congressional authorization. *See Padilla*, 352 F.3d at 722; *see also* Curtis A. Bradley & Jack L. Goldsmith, *Congressional Authorization and the War on Terrorism*, 118 HARV. L. REV. 2047, 2117–23 (2005) (arguing that the AUMF's text suggests that the President is authorized to use force anywhere he encounters the enemy covered by the AUMF, including the United States). *But see generally* Vladeck, *supra* note 15 (surveying the history of the President's power to detain U.S. citizens and finding that no statutes today satisfy the requirements of the Non-Detention Act's application to Padilla and Hamdi).

[18] *Padilla*, 233 F. Supp. 2d at 598–99, *rev'd*, 352 F.3d 695 (2d Cir. 2003).

[19] *Padilla*, 352 F.3d at 720. The appellate panel was divided, and, in his dissent Judge Wesley supported the district court's broad reading of the AUMF and found that the June 9 Order fell within the authority granted to the President "to stop al Qaeda from killing or harming Americans here or abroad." *Id.* at 730 (Wesley, J., dissenting).

[20] *See* Padilla v. Hanft, 389 F. Supp. 2d 678, 690–91 (D.S.C. 2005).

[21] *See* Padilla v. Hanft, 423 F.3d 386, 397 (4th Cir. 2005).

the issue reaches the Supreme Court, I conclude that there is evidence that a majority of the court is not inclined to find that the AUMF authorized Padilla's detention. However, following the *Hamdi* decision, the government has amended its claims against Padilla in an effort to bring his detention within the scope of the *Hamdi* plurality's holding.[22] As discussed below, the Fourth Circuit heavily relied on those new arguments in finding that the *Hamdi* plurality's analysis applied and that the AUMF likewise authorized Padilla's detention. Such arguments might be similarly persuasive in the Supreme Court.

At the first stage of Padilla's case, the District Court for the Southern District of New York found that while the Non-Detention Act applied to Padilla's detention, the AUMF had authorized the detention.[23] The court offered scant analysis of the Non-Detention Act beyond finding that its plain language encompassed all detentions of U.S. citizens, including Padilla's detention.[24] The court then focused on the AUMF's broad language authorizing the President to "[u]se all necessary and appropriate force" against organizations and persons responsible for the September 11 attacks "in order to prevent any future acts of international terrorism."[25] The district court reasoned that since the President's order to detain Padilla asserted that he was an unlawful combatant on behalf of al Qaeda, which was responsible for the September 11 attacks, Padilla's detention fell within the authority granted by the AUMF.[26]

On appeal, the Second Circuit reversed that decision. After engaging in an extensive analysis of the Non-Detention Act, the court applied a strong clear statement requirement to the AUMF.[27] The court found evidence in

---

[22] *Compare* Brief for Respondent-Appellant, Padilla v. Rumsfeld, 352 F.3d 695 (2d Cir. 2003) (No. 03-2235), 2003 WL 23622382, at *4 ("While in Afghanistan . . . Padilla met several times with al Qaida officials . . . .") *with* Opening Brief for Appellant, Padilla v. Hanft, 423 F.3d 386 (4th Cir. 2005) (No. 05-6396), 2005 WL 1521249, *23 (describing Padilla as performing guard duties in Afghanistan for the Taliban and asserting that Padilla was "armed with a Kalashnikov assault rifle and ammunition."); *see also* Neil A. Lewis, *Court Gives Bush Right to Detain U.S. Combatant*, N.Y. TIMES, Sept. 10, 2005, at A1 (noting that "the government no longer presents as the main charge against [Padilla] that he intended to set off a 'dirty bomb' . . . . Instead, as his case made its way through the court system for the second time, the government all but eliminated that accusation . . . . Government lawyers argued that the main new reason he should be detained as an enemy combatant was that he fought American forces in Afghanistan alongside Qaeda colleagues.").

[23] *See Padilla*, 233 F. Supp. 2d at 596–99.

[24] *See id.* at 596 (citing *Howe v. Smith*, 452 U.S. 473, 480 n.3 (1981) ("[T]he plain language of § 4001(a) proscribe[es] detention *of any kind* by the United States . . . .") (emphasis in original)).

[25] *Padilla*, 233 F. Supp. 2d at 598–99 (citing AUMF § 2(a)).

[26] *See id.* at 599; *see also Padilla*, 352 F.3d at App. A (containing the text of the President's June 9 Order).

[27] *See Padilla*, 352 F.3d at 718–22. The court applied the same clear statement requirement to the other statutory sources of authorization that the executive's attorneys proffered. *Id.* at 723–24 ("In light of . . . the Non-Detention Act's requirement that Congress specifically authorize detentions of American citizens . . . we decline to impose on section 956(5) loads it cannot bear.").

the legislative history of the Non-Detention Act that Congress intended that only an express authorization of detention would suffice.[28] Accordingly, the court concluded "that precise and specific language authorizing the detention of American citizens is required to override [the Non-Detention Act's] prohibition."[29] As the court noted, the AUMF does not contain any language expressly authorizing detention.[30]

The Second Circuit also declined to infer authorization to detain from the AUMF, grounding its analysis in the Supreme Court's historical preference for clear statements of authorization where a constitutionally protected liberty is implicated.[31] The court cited the Supreme Court's statement in *Ex parte Endo*: "We must assume, when asked to find implied powers in a grant of legislative or executive authority, that the law makers intended to place no greater restraint on the citizen than was clearly and unmistakably indicated by the language they used."[32]

Applying that principle to the AUMF, the Second Circuit found that it lacked the "clear" and "unmistakable" language required to authorize detentions of citizens captured inside the United States.[33] Moreover, the court found that the congressional debates over the AUMF "are at best equivocal as to the President's powers and never mention the issue of detention."[34] Finally, the court found it incongruous that Congress had specified in the AUMF that the War Powers Resolution remained applicable to any military action[35] but failed to specify authorization to detain citizens under the Non-Detention Act.[36] Accordingly, the Second Circuit held that the AUMF had not authorized Padilla's detention in light of the strong clear statement requirement elaborated in Endo and embodied in the Non-Detention Act.

When Padilla's claims eventually reached the Supreme Court, a majority of the Court voted to reverse the case on procedural grounds, and the case was remanded for entry of an order of dismissal without prejudice.[37]

---

[28] "You have got to have an act of Congress to detain, and the act of Congress must authorize detention." *Id.* at 730 (citing 117 CONG. REC. H31555 (daily ed. Sept. 13, 1971) (statement of Rep. Eckhardt)).

[29] *Padilla*, 352 F.3d at 730.

[30] *Id.* at 722 (citing AUMF §2(a)).

[31] *See generally* Cass R. Sunstein, *Minimalism at War*, 2004 SUP. CT. REV. 47 (describing three central components of Supreme Court minimalism including "a requirement of clear congressional authorization for executive action intruding on interests with a claim to constitutional protection.").

[32] *Padilla*, 352 F.3d at 723 (citing *Ex parte Endo*, 323 U.S. 283, 300 (1944)).

[33] Responding to the Fourth Circuit's decision regarding Hamdi's detention, the Second Circuit did allow that "it may be possible to infer a power of detention . . . in the battlefield context where detentions are necessary to carry out the war." *Padilla*, 352 F.3d at 723.

[34] *Id.* at 723 n.31.

[35] *Id.* at 729 (citing AUMF §2(b)).

[36] *Id.* at 723.

[37] *Padilla*, 542 U.S. 426, 430 (2004). Chief Justice Rehnquist wrote the majority opinion, in which Justices O'Connor, Scalia, Kennedy, and Thomas joined. *Id.*

As a result, the majority avoided deciding the merits of the case.[38] In brief, the majority held that the Southern District of New York did not properly have habeas corpus jurisdiction over Padilla because he was confined within the District of South Carolina. The Court also held that Secretary of Defense Rumsfeld was improperly listed as respondent, and that the commander of the facility in which Padilla was being held was the proper respondent.[39] Following the Supreme Court's decision, Padilla's attorneys filed his habeas petition in the District of South Carolina and named the commander of the brig as respondent.

On the same day it remanded Padilla's case, the Supreme Court issued its decision on the merits of Yasser Hamdi's detention in a plurality opinion that provoked three dissents. Following the *Hamdi* decision, the government altered the allegations underlying Padilla's detention in an effort to bring his detention within the *Hamdi* plurality's narrow holding that the AUMF authorized detention of enemy combatants to prevent their return to the battlefield.[40]

Despite the new grounds proffered by the executive, the district court in South Carolina granted Padilla's habeas petition and ordered his release. First, the court distinguished Hamdi's detention, noting the Fourth Circuit's dictum that to compare the two cases was "to compare apples and oranges."[41] The court also noted that the *Hamdi* plurality repeatedly stated that its holding was limited to the particular facts of Hamdi's capture in a foreign zone of combat.[42] Second, the district court found that both *Quirin*[43] and *Milligan*[44] imposed a clear statement requirement that "the detention of a United States citizen by the military is disallowed without explicit Congressional authorization."[45] Third, the court found that the AUMF did not authorize Padilla's detention and that his detention was therefore illegal under the Non-Detention Act.[46] Citing the Second Circuit opinion, the court held that the AUMF was not sufficiently clear to authorize Padilla's detention.[47] Finally, the court repeatedly referenced Justice Scalia's *Hamdi* dissent before ordering that, since Congress had not suspended the writ of habeas corpus, Padilla must be released.[48]

---

[38] *Id.*

[39] *See* Hamdi v. Rumsfeld, 542 U.S. 507, 509, 517–19 (2004).

[40] *See* Opening Brief for Appellant, Padilla v. Hanft, 423 F.3d 386 (4th Cir. 2005) (No. 05-6396), 2005 WL 1521249, at *22–*23 ("Padilla's combat activities in Afghanistan are not materially distinguishable from Hamdi's and so Padilla fits squarely within the enemy-combatant definition that the Supreme Court utilized in *Hamdi*.").

[41] Padilla v. Hanft, 389 F. Supp. 2d 678, 685 (D.S.C. 2005) (citing Hamdi v. Rumsfeld, 337 F.3d 335, 344 (4th Cir. 2003) (Wilkinson, J., concurring)).

[42] *See Padilla*, 2005 U.S. Dist. LEXIS 2921, at *18 n.8.

[43] *Ex parte* Quirin, 317 U.S. 1 (1942).

[44] *Ex parte* Milligan, 71 U.S. 2 (4 Wall.) (1866).

[45] *See Padilla*, 389 F. Supp. 2d at 688.

[46] *See id.* at 689.

[47] *See id.*

[48] *See id.* at 692.

In September 2005, a unanimous panel of the Fourth Circuit reversed that holding and ruled that the AUMF had authorized Padilla's detention.[49] Crucially, the Fourth Circuit framed the issue as whether the President had the authority to detain a citizen who "is closely associated with al Qaeda . . . ; who took up arms on behalf of that enemy and against our country in a foreign combat zone of that war; *and* who thereafter traveled to the United States for the avowed purpose of further prosecuting that war on American soil."[50] Framed in this way, the Fourth Circuit found that Padilla's designation as an enemy combatant was indistinguishable from Hamdi's case. Padilla's detention was "no less necessary than was Hamdi's in order to prevent his return to the battlefield."[51] Therefore, "as the AUMF authorized Hamdi's detention by the President, so also does it authorize Padilla's detention."[52]

The Fourth Circuit rejected Padilla's claim that his capture inside the United States is sufficient to distinguish his detention from Hamdi's. In the court's reading, the *Hamdi* plurality did not make capture in a foreign combat zone a necessary condition for detention under the AUMF.[53] Rather, the court read *Hamdi* as stating that the support of forces hostile to the United States in Afghanistan and engaging in armed conflict were sufficient to trigger the authority to detain under the AUMF.[54] Because Padilla poses "the same threat of returning to the battlefield as Hamdi posed," the court concluded that Padilla's detention is as necessary and appropriate as was Hamdi's.[55]

The court of appeals also rejected all of Padilla's other claims, including Padilla's argument that the AUMF failed to provide a sufficiently clear statement of authorization to detain.[56] Echoing the Second Circuit's analysis, Padilla had argued that only a clear statement would suffice. Responding to Padilla's use of *Endo*, the court rejected the existence of a clear statement requirement. It pointed to *Endo*'s statement that the statute's silence regarding detentions did "not of course mean that any power to detain [was] lacking."[57] Furthermore, the court noted that the *Endo* court had ordered the detainee's release because it found that detention of a concededly loyal citizen bore no relation to the statute's purpose of preventing sabotage.[58] In contrast, the Fourth Circuit understood the *Hamdi*

---

[49] *See* Padilla v. Hanft, 423 F.3d 386, 397 (4th Cir. 2005).

[50] *See id.* at 389 (emphasis added).

[51] *See id.* at 392.

[52] *See id.* at 391.

[53] *See id.* at 393–95 ("Nowhere in its framing of the "narrow question" presented did the plurality even mention the locus of capture."); *see also* Bradley & Goldsmith, *supra* note 17, at 2118–20 (arguing that the AUMF "authorizes the President to use force anywhere he encounters the enemy covered by the AUMF, including the United States").

[54] *See Padilla*, 423 F.3d at 393–95.

[55] *See id.* at 393.

[56] *See id.* at 396.

[57] *See id.* at 395–96 (citing *Ex parte* Endo, 323 U.S. at 301).

[58] *See id.* at 395 n.5 (citing Endo, 323 U.S. at 202).

plurality as holding that detention of citizens like Padilla, "who associated with al Qaeda and the Taliban regime, who took up arms against this Nation . . . *and* who entered the United States for the avowed purpose of further prosecuting that war"[59] clearly furthered the purpose of the AUMF. Accordingly, since the AUMF was sufficiently clear to detain Hamdi, and "nothing in the AUMF permits us to conclude that [it] clearly and unmistakably authorized Hamdi's detention but not Padilla's,"[60] Padilla's detention was authorized.

Despite the Fourth Circuit's broad reading of *Hamdi*, there is some evidence that its reliance on the plurality in that case as settling the matter is misplaced. The plurality opinion in *Hamdi* and the majority opinion dismissing Padilla's original litigation provoked strong dissents. A review of those dissents provides some evidence that a majority of the Court may hold that the AUMF did not authorize Padilla's detention.

Although a majority of the Supreme Court dismissed Padilla's first suit on procedural grounds, Justice Stevens's dissent, in which three other Justices joined,[61] included a footnote affirming the Second Circuit's holding regarding the AUMF. That footnote states: "Consistent with the judgment of the Court of Appeals, I believe that the Non-Detention Act prohibits—and the Authorization for Use of Military Force Joint Resolution . . . does not authorize—the protracted, incommunicado detention of American citizens arrested in the United States."[62] Although the footnote did not offer any further detail, the statement is strong evidence that at least four justices have serious reservations about the AUMF's sufficiency under the Non-Detention Act.

Additionally, while Justice Scalia joined the *Padilla* majority in remanding that case on procedural grounds, his strong dissent in *Hamdi* is equally applicable to Padilla's detention. Justice Scalia found that Hamdi was entitled to release, unless criminal charges were filed or Congress suspended the writ of habeas corpus.[63] Regarding the AUMF, Scalia stated four reasons why he did not think it authorized Hamdi's detention.[64] All four reasons were grounded in clear statement principles, and among them was his

---

[59] *Padilla*, 423 F.3d at 397.

[60] *Id.* at 396.

[61] Joining Justice Stevens were Justices Souter, Ginsburg, and Breyer. Padilla v. Rumsfeld, 542 U.S. 426, 455 (2004) (Stevens, J., dissenting). Notably, Justice Breyer had joined with the *Hamdi v. Rumsfeld* plurality in finding that the AUMF had authorized that detention. 542 U.S. 507, 509 (2004).

[62] *Padilla*, 542 U.S. at 455 (Stevens, J., dissenting) (citations omitted).

[63] *Hamdi*, 542 U.S. at 554–55 (Scalia, J., dissenting). Interestingly, Justice Scalia indicates that nothing short of official suspension of the writ of habeas corpus would suffice to authorize military detention. As he says, if the Suspension Clause "merely guarantees the citizen that he will not be detained unless Congress by ordinary legislation says he can be detained; it guarantees him very little indeed." *Id.* at 575. Accordingly, Justice Scalia might find even DECA, which explicitly authorizes detentions, to be insufficient as it is not a suspension of habeas corpus.

[64] *See id.* at 573 (Scalia, J., dissenting).

view that the AUMF lacked "the clarity necessary to overcome the statutory prescription" of the Non-Detention Act.[65] Importantly, Justice Scalia explicitly stated that his views regarding Hamdi's detention also applied to Padilla's case.[66]

In reviewing Justice Stevens's footnote, some scholars have suggested that the footnote merely rejects "incommunicado" detentions,[67] and that the Court might adopt a *Hamdi*-like compromise that recognized the authority to detain but imposed procedural safeguards in the case of citizens captured on American soil.[68] Indeed, Justice Stevens himself wrote that whether Padilla "is entitled to immediate release is a question that reasonable jurists may answer in different ways,"[69] and his opinion placed heavy emphasis on Padilla's lack of access to counsel.[70] In light of the *Hamdi* plurality opinion and its use in the Fourth Circuit's *Padilla* opinion, it is certainly plausible to suggest that a compromise that authorizes detention under AUMF will be forthcoming.

Indeed, the executive's efforts to portray Padilla's detention and Hamdi's detention as similar make a Supreme Court ruling upholding Padilla's detention more likely. As the Fourth Circuit repeatedly noted, Padilla had allegedly returned to the United States to continue hostilities that he had begun in Afghanistan.[71] Padilla's presence in the United States was a continuation of the hostilities that the AUMF had addressed, and there may be no obvious reason to limit the AUMF's scope to foreign battlefields.[72]

As the jumble of lower court opinions and Supreme Court plurality and dissenting opinions suggests, reasonable jurists may indeed differ as to whether the AUMF authorizes detentions of citizens captured inside the United States. Regardless of how the Supreme Court ultimately decides Padilla's case, substantial uncertainty about the legality of future detentions will remain, especially if another plurality opinion restricted to the "narrow question presented" is issued. This continuing uncertainty highlights the need for Congress to clarify the executive's authority in this area.[73]

---

[65] *Id.*

[66] *See id.* at 577 ("[C]urrently we know of only two [parties to whom his dissent would apply], Hamdi and Jose Padilla.").

[67] *See* Bradley & Goldsmith, *supra* note 17, at 2120 n.324. Professors Bradley and Goldsmith also argue that a clear statement is required only if the President's actions restrict the liberty of non-combatants in the United States. *See* Curtis A. Bradley & Jack L. Goldsmith, *Rejoinder: The War on Terrorism: International Law, Clear Statement Requirements, and Constitutional Design*, 118 HARV. L. REV. 2683, 2693 (2005).

[68] *See* Bradley & Goldsmith, *supra* note 17, at 2121–23 and accompanying notes.

[69] *Padilla*, 542 U.S. at 465 (Stevens, J., dissenting).

[70] *See id.* ("Access to counsel for the purpose of protecting the citizen from official mistakes and mistreatment is the hallmark of due process.").

[71] *See supra* notes 49–60 and accompanying text.

[72] *See* Bradley & Goldsmith, *supra* note 17, at 2117–23 and accompanying notes.

[73] *See* Michael Chertoff, *Law, Loyalty, and Terror*, 9 WKLY. STANDARD, Dec. 1, 2003, 15, 17 (stating the need "to debate a long-term and sustainable architecture for the process of determining when, why, and for how long someone may be detained as an enemy combatant . . . .").

The Detention of Enemy Combatants Act attempts to offer such clarity, and the remainder of this Recent Development focuses on its provisions.

II. THE DETENTION OF ENEMY COMBATANTS ACT (DECA)

DECA clearly and unmistakably authorizes executive detention of any "United States person or resident"[74] as an "enemy combatant"[75] if he "is a member of al Qaeda, or knowingly cooperated with a member of al Qaeda in the planning, authorizing, committing, aiding, or abetting of one or more terrorist acts against the United States."[76] If the bill were intended merely to authorize citizen detentions, that language would be sufficient. However, DECA is also intended to ensure due process for detainees.[77] Accordingly, much of DECA is devoted to imposing procedural requirements on the executive branch's authority to detain. After reviewing DECA's major provisions, I conclude that most of these requirements suffer from too many ambiguities and omissions to effectively guarantee due process for detainees.

First, DECA restricts executive authority by limiting the terms of the authorizing language. The authorization to detain extends only to citizens who are members of, or knowingly cooperate with, al Qaeda. Despite the amorphous nature of the term "al Qaeda," the limitation is still significant. As commonly understood, and DECA itself provides no further definition, al Qaeda encompasses a broad confederacy of affiliated terrorist organizations around the world.[78] Although the term is broad, the requirement of a nexus with al Qaeda in particular means, for example, that detention of a citizen who was a member of an independent terror group would not be authorized.

Second, DECA requires that publicly promulgated standards and criteria govern "the determination that an American citizen or lawful resident is an enemy combatant" as well as the detention itself.[79] DECA places primary responsibility for promulgating these standards with the Secretary of Defense. However, in a blurring of the military and criminal aspects of the war on terror,[80] the Secretary of Defense is directed to consult with the

---

[74] H.R. 1076, 109th Cong. § 3(a) (2005).

[75] *Id.* Regarding the definition of "enemy combatant," DECA's section on congressional findings states that "[e]nemy combatants in the war on terrorism are not defined by simple, readily apparent criteria . . . . And the power to name a citizen as an 'enemy combatant' is therefore extraordinarily broad." *Id.* § 2(8).

[76] *Id.* § 3(a).

[77] *See* 151 CONG. REC. E351 (daily ed. Mar. 3, 2005) (statement of Rep. Schiff).

[78] *See* ROHAN GUNARATNA, INSIDE AL QAEDA: GLOBAL NETWORK OF TERROR 95 (2003); Bradley & Goldsmith, *supra* note 17, at 2109 nn.278–80 and accompanying text (citing PETER L. BERGEN, HOLY WAR, INC. 195–220 (2001)); Rohan Gunaratna, *The Post-Madrid Face of Al Qaeda*, WASH. Q., Summer 2004, at 91, 93 (describing al Qaeda's post–September 11 evolution into "a movement of two dozen groups").

[79] H.R. 1076 § 3(b).

[80] *See generally* Bruce Ackerman, *The Emergency Constitution*, 113 YALE L.J. 1029,

Secretary of State and the Attorney General in creating these standards.[81] Finally, DECA requires that the Secretary of Defense report the standards to the Senate and House Judiciary Committees.[82] Taken together, these requirements purport to subject the executive's authority to detain citizens to significant public comment and review.

Regarding these procedural rules for detention, DECA makes two specifications. First, the rules must "guarantee timely access to judicial review to challenge the basis for a detention."[83] Second, the rules must also "permit the detainee access to counsel."[84] DECA does not address the nature of such review, except for a jurisdictional provision discussed below, and is silent as to what "access to counsel" means. Other sources of law may provide content to these terms, but the use of such sources is itself a potential area of disagreement.[85]

Third, DECA specifies that the President must certify detentions and thereby purports to limit the length of detentions. To effect an authorized detention, the President must certify that United States Armed Forces continue to be engaged in a state of armed conflict with al Qaeda.[86] Additionally, the President must certify that an "investigation with a view toward a prosecution, a prosecution, or a post-trial proceeding in the case of [a detainee] is ongoing" or that detention is warranted "in order to prevent such person or resident from aiding persons attempting to commit terrorist acts against the United States."[87] A presidential certification would be effective for 180 days, after which time the President would be required to make a new certification.[88] As DECA does not specify, it may be assumed that the President could continually recertify a citizen's detention so long as the certification requirements continued to be met.

Fourth, DECA sets certain minimum standards for the conditions of detention. A detainee must be held at "an appropriate location designated by the Secretary of Defense."[89] A detainee must be "treated humanely, without any adverse distinction based on race, color, religion, gender, birth, wealth, or any similar criteria."[90] A detainee must be "afforded adequate food, drink-

---

1032–37 (2004) (describing the inadequacies of the laws of war and the criminal law in addressing the problems terrorism poses).

[81] H.R. 1076 § 3(b).

[82] *Id.*

[83] *Id.* § 4.

[84] *Id.*

[85] *See generally* Laura A. Dickinson, *Using Legal Process To Fight Terrorism: Detentions, Military Commissions, International Tribunals, and the Rule of Law*, 75 S. CAL. L. REV. 1407, 1412–35 (2002) (surveying the arguments that the U.S. Constitution and international law afford procedural rights to detainees).

[86] H.R. 1076 § 5(a)(1).

[87] *Id.*

[88] *Id.* § 5(a)(2).

[89] *Id.* § 5(c)(1).

[90] *Id.* § 5(c)(2).

ing water, shelter, clothing, and medical treatment"[91] and be "sheltered under hygienic conditions."[92] Finally, a detainee must be "allowed the free exercise of religion consistent with the requirements of such detention."[93] Although these provisions may overlap with extant international laws governing treatment of POWs and other military detainees,[94] DECA makes many of those protections explicit.

Fifth, DECA empowers federal courts to review detentions. DECA consolidates such reviews in the U.S. District Court for the District of Columbia, which is given "exclusive jurisdiction to review any detention under this Act to ensure that the requirements of this Act for detaining an accused are satisfied."[95] While some commentators have advocated the creation of a separate federal circuit to hear enemy combatant habeas petitions,[96] DECA's consolidation of jurisdiction would achieve the same efficiencies without the burdens of creating and staffing a new circuit. The provision would also moot arguments within the Supreme Court regarding the proper location for a habeas petition,[97] while creating a "counterarena of centralized . . . expertise within the judiciary."[98]

Finally, DECA requires the President to submit to Congress a "report on the use of the authority provided by this Act."[99] Such reports shall be made at least once every twelve months and "shall specify each individual subject to, or detained pursuant to, the authority provided by this Act."[100] DECA's authors recently succeeded in having a version of this requirement passed as an amendment to the Justice Department Reauthorization Act of 2005.[101]

---

[91] H.R. 1076 § 5(a)(1), 5(c)(3).

[92] *Id.* § 5(c)(4).

[93] *Id.* § 5(c)(5).

[94] *See generally* Goodman & Jinks, *supra* note 9, at 2660–61 (discussing the treatment of detainees mandated under international laws of armed conflict).

[95] H.R. 1076 § 5(b).

[96] *See* Christopher A. Chrisman, *Article III Goes to War: A Case for a Separate Circuit Court for Enemy Combatant Habeas Cases*, 21 J.L. & POL. 31, 93–100 (2005).

[97] *Compare Padilla*, 542 U.S. at 451 (Kennedy, J., concurring) ("In my view, the question of the proper location for a habeas petition is best understood as a question of personal jurisdiction or venue.") *and id.* at 463 (Stevens, J., dissenting) ("the question is one of venue . . . .") *with id.* at 447 (stating that a habeas petitioner must "file the petition in the district of confinement.").

[98] HAROLD KOH, THE NATIONAL SECURITY CONSTITUTION 183 (1990).

[99] H.R. 1076 § 6.

[100] *Id.*

[101] Department of Justice Appropriations Authorization Act, Fiscal Years 2006 through 2009, H.R. 3402, 109th Cong., § 307 (2005) ("No less often than once every 12 months, the Attorney General shall submit to Congress a report on the status of United States persons or residents detained, as of the date of the report, on suspicion of terrorism."). *But see* Center for National Security Studies v. Department of Justice, 331 F.3d 918 (D.C. Cir. 2003) (holding that under statutory exceptions to the Freedom of Information Act, the government is not required to release the names of detainees or their attorneys).

Before reviewing these provisions, it should be noted that DECA's authors have stated that the act is intended to be an interim measure,[102] and Congress has yet to seriously address the issue of citizen detentions. While DECA would satisfy the Non-Detention Act and impose some requirements on executive authority to detain, a more detailed piece of legislation would provide a more thorough framework for detentions of citizens as enemy combatants.[103] The present version of DECA would certainly evolve substantially as it moved through the legislative process. Unfortunately, Congress has not yet addressed the issue and no hearings regarding DECA or citizen detentions have been scheduled in the House.[104] A Senate hearing on citizen detentions was scheduled for October 2002 but was postponed indefinitely.[105] At best then, the current version of DECA represents a rough outline of an authorization that could be improved upon should Congress decide to take up this issue in earnest.

Recognizing those caveats, a brief review of DECA's provisions reveals several serious ambiguities and omissions in the current version of the Act. First, rather than clarifying detainees' status, DECA muddles it by blurring the military and criminal contexts of detention. Second, DECA's reporting requirements are insufficiently detailed and leave too much discretion to the executive to represent effective regulation. Third, DECA's regulation of the conditions of detention fails to address key issues such as the interrogation of detainees. Finally, DECA's provision for judicial review is too vague to ensure meaningful review. Absent more explicit congressional directives regarding the process and conditions of detention, DECA's grant of congressional authorization for executive actions will likely result in less, not more, judicial scrutiny of detentions.

First, DECA adds considerable confusion to detainees' overall status as either military prisoners or criminal defendants.[106] This confusion was evident in the *Hamdi* plurality, which stated that Hamdi was owed procedural safeguards common to the civil and criminal contexts while finding that the AUMF and the laws of war provided authorization for his detention.[107]

---

[102] *See* 151 CONG. REC. E351 (daily ed. Mar. 3, 2005) (statement of Rep. Schiff).

[103] For a discussion of framework statutes in the national security context, see KOH, *supra* note 98, at 69–70.

[104] For a general discussion of reasons why Congress often acquiesces to presidential actions, see *id.* at 123–33. *See USA PATRIOT Act: A Review for the Purpose of Reauthorization: Before the House Comm. on the Judiciary*, 109th Cong. 20 (2005) (statement of Rep. Schiff).

[105] *See The Constitution and the Detention of US Citizens as Enemy Combatants: Hearing Before the Subcomm. on the Constitution, Civil Rights, and Property Rights of the S. Comm. on the Judiciary*, 108th Cong. 1 (2002), *available at* http://judiciary.senate.gov/hearing.cfm?id=471 (last visited Nov. 20, 2005).

[106] *See generally* Bruce Ackerman, *This Is Not A War*, 113 YALE L.J. 1871 (2004).

[107] *Compare* Hamdi v. Rumsfeld, 542 U.S. 507, 533 (2004) ("a citizen-detainee seeking to challenge his status as an enemy combatant must receive notice of the factual basis for his classification, and a fair opportunity to rebut the Government's factual assertions before a neutral decision maker") (quoting Cleveland Bd. of Educ. v. Loudermill, 470 U.S. 532 (1985) *and* Mullane v. Central Hanover Bank & Trust Co., 339 U.S. 306 (1950)) *with Hamdi*, 542 U.S. at 518 ("[T]he capture, detention, and trial of unlawful combatants, by universal

DECA further muddles the question of whether citizen enemy combatants are military prisoners or criminal defendants. Under DECA's scheme, the Secretary of Defense is to promulgate the rules governing detentions of citizens believed to be enemy combatants, but he is to do so in consultation with the Attorney General.[108] The required Presidential certification must assert both that armed conflict with al Qaeda is ongoing and that a prosecution or an investigation "with a view toward a prosecution" is also ongoing.[109] This muddling of the military and criminal contexts has serious implications for detainees' rights. For example, the term "prosecution" has divergent meanings in the military and criminal contexts. In its current form, DECA would plausibly allow the President to certify the "ongoing prosecution" requirement if he planned to try citizen detainees before military commissions.

Second, DECA's reporting requirements are insufficiently detailed and defer too much to executive discretion to represent meaningful regulation. While the Secretary of Defense must publish the criteria for detention and report them to Congress, DECA contains no guidance as to the content of those rules. What methods of gathering evidence would be allowed? What quantum of evidence would be sufficient to trigger detention? Absent these specifications, the promulgation of rules may not have the effect of providing the meaningful notice and review of executive actions that DECA's authors intend.

Indeed, members of the executive branch have previously offered vague descriptions of the process used to determine whether or not a citizen merits detention as an enemy combatant, and DECA does not compel the executive to provide additional details. In a speech to the ABA, then-Counsel to the President Alberto Gonzales offered "to explain the decision-making that led our enemy combatant determinations with respect to U.S. citizens."[110] Regarding Hamdi's capture in Afghanistan, Gonzales stated that a "U.S. military screening team confirmed that Hamdi met the criteria for enemy combatants . . . . In such a situation in a foreign zone of combat, that determination was quite properly made by military personnel on the ground."[111] Gonzales did not offer any guidance as to the criteria that the screening teams use in making their determinations.

Regarding detentions of citizens captured on U.S. soil, Gonzales stated that when it appears that a captured U.S. citizen may be an al Qaeda op-

---

agreement and practice, are important incidents of war.") (quoting *Ex parte* Quirin, 317 U.S. 1). Moreover, as Justice Scalia noted, the plurality analysis adopted a balancing test originally used in a case involving the *"withdrawal of disability benefits!" Hamdi*, 542 U.S. at 574 (Scalia, J., dissenting).

[108] H.R. 1076, 109th Cong. § 3(b) (2005).

[109] *Id.* § 5(a)(1)(A).

[110] Alberto A. Gonzales, Counsel to the President, Speech to the American Bar Association Standing Committee on Law and National Security: Detention Issues in the War on Terrorism 7 (Feb. 24, 2004) (*available at* http://www.abanet.org/natsecurity).

[111] *Id.*

erative, the option to detain him as an enemy combatant is pursued when "it appears that criminal prosecution and detention as a material witness are, on balance, less-than-ideal options as long-term solutions to the situation . . . ."[112] Gonzales did not elaborate as to what was being balanced, what made an option "ideal," or what were the "long-term solutions" being sought. Gonzales then provided an overview of the process by which the Department of Defense, CIA, Justice Department, and White House Counsel communicate their recommendations to the President.[113] First, the Justice Department's Office of Legal Counsel (OLC) ascertains whether a person "has become a member or associated himself with hostile enemy forces."[114] The parameters of the term "associated with" were left undefined. Second, the Department of Defense, CIA, and Justice Department Criminal Division make their recommendations, with the Secretary of Defense transmitting all of the relevant materials to the White House. Finally, the President receives a recommendation from his Counsel and then decides whether to designate a citizen as an enemy combatant.

Gonzales concluded that the procedures in place are "elaborate and careful,"[115] but his description did not address the standards the agencies employ in making their recommendations. A vigorous exchange of materials within the executive is heartening but does not provide any further clarity for citizens as to the basis for their detention. The standards being employed in designating a citizen as an enemy combatant remain mysterious. DECA, as currently structured, may not compel a more full disclosure.

Unless legislation compels more detailed disclosure, executive discretion alone will determine the contours of the procedural rights of detainees. Indeed, other provisions reveal the extent to which DECA explicitly places crucial aspects of detention within executive discretion. The Secretary of Defense is to determine "an appropriate location" for detainees to be held.[116] While a military brig located in the United States would probably be appropriate, would Camp X-Ray at Guantanamo Bay be appropriate? A U.S. naval vessel at sea? A facility operated by a foreign government? Moreover, although DECA states that the rules of detention should allow for timely judicial review of the basis for a detention and permit access to counsel, the meaning of these crucial terms is left unspecified. What constitutes "timely" review? Could a military commission afford "judicial review"?[117] Would access to counsel be unfettered, or could communications with counsel be monitored? Would counsel have access to the evidence underpinning the decision to detain? What if that evidence were

---

[112] *Id.* at 7–8.
[113] *Id.* at 8–9.
[114] *Id.* at 8.
[115] *Id.* at 9.
[116] H.R. 1076 § 5(c)(1).
[117] The *Hamdi* plurality suggested that it could. *See* Hamdi v. Rumsfeld, 542 U.S. 507, 538 (2004).

classified? In its present form, DECA allows the executive to answer such questions of its own accord.

While some items in this parade of horribles may be unlikely to occur, DECA explicitly puts the responsibility for answering these questions in the hands of the executive itself. The rules for detention that the Defense Secretary promulgates should "preserve the Government's ability to detain . . . and protect the confidentiality of that information . . . which, if released, could impede the Government's investigation of terrorism."[118] Although Congress should be wary of interfering in executive decisions regarding ongoing military action, limits on executive discretion cannot be avoided if Congress sincerely desires to create a statutory regime that actually protects citizens from arbitrary detention.[119]

The third major deficiency in DECA's regulatory scheme concerns the conditions of detention. That is, DECA does not clarify what sort of interrogation techniques would be authorized for use in gaining information from citizen detainees. DECA allows for detention to prevent citizens from "aiding persons attempting to commit terrorist acts."[120] Gaining intelligence on terrorist acts being planned would probably be a crucial objective of detentions. DECA itself allows that the rules for detentions promulgated by the Secretary of Defense should themselves "assist in the gathering of vital intelligence."[121] The level of interrogation to which detainees could be subjected is a crucial aspect of due process that DECA again fails to address.

Finally, DECA's provision for judicial review is too vague to ensure meaningful review. Absent a more explicit congressional directive regarding their mandate, courts will be unlikely to scrutinize executive detentions that have received Congress's imprimatur. Courts are normally hesitant to second-guess executive determinations regarding national security and foreign affairs.[122] As the *Hamdi* litigation demonstrated, often courts are loath to investigate the factual underpinnings of a presidential determination that a person is an enemy combatant.[123] As DECA itself notes, "[c]ourts must give broad deference to military judgment concerning the determination of enemy combatant status, POW status, and related questions."[124] Congres-

---

[118] H.R. 1076, 109th Cong. § 4 (2005).

[119] *See* 151 CONG. REC. E351 (daily ed. Mar. 3, 2005) (statement of Rep. Schiff) ("While we must grant broad latitude to our armed forces when it comes to protecting national security, American citizens should not be held indefinitely upon the sole determination of one branch of government . . . .").

[120] H.R. 1076 § 5(a)(1)(B).

[121] *Id.* § 4.

[122] *See generally* WILLIAM REHNQUIST, ALL THE LAWS BUT ONE (1998); *see also* Hamdi v. Rumsfeld, 542 U.S. 507, 579–599 (2004) (Thomas, J., dissenting) ("This detention falls squarely within the Federal Government's war powers, and we lack the expertise and capacity to second-guess that decision.").

[123] *See, e.g.*, Padilla v. Rumsfeld, 233 F. Supp. 2d 564, 609 (S.D.N.Y. 2002) ("The commission of a judge . . . does not run to deciding de novo whether Padilla is associated with al Qaeda and whether he should therefore be detained as an enemy combatant.").

[124] H.R. 1076, 109th Cong. § 2(10) (2005).

sional authorization of executive actions places executive actions and judgments virtually beyond review. As Justice Jackson famously stated, "When the President acts pursuant to an express or implied authorization of Congress, his authority is at its maximum . . . . In these circumstances, and in these only, may he be said (for what it may be worth) to personify the federal sovereignty."[125] Taken alone, congressional authorization would almost certainly have the effect of discouraging courts from undertaking searching reviews of detentions.

However, Congress can authorize detentions and still provide for more searching judicial review. A statutory scheme that contained more detail about whether, and under what conditions, a citizen could be detained would give courts a firm footing for scrutinizing executive action. DECA would clearly benefit from inclusion of express language regarding evidentiary standards, burdens of proof, and restrictions on interrogation techniques. While a blanket authorization sets presidential powers at their maximum, meaningful proscriptions would set the President's powers "at their lowest ebb"[126] should he take actions incompatible with those proscriptions.

By its own terms, DECA is a retroactive authorization of detentions that are already occurring.[127] Some may quarrel with DECA's legalization of detentions as being another example of Congress's rush to fall in line with the President during a crisis and to ratify clearly unconstitutional conduct.[128] Others have argued that legislative ratification of executive actions engages the polity and provides a mechanism for accountability once the crisis has abated.[129] Without confronting either of those two views directly, this Recent Development has attempted to show that Congress can profitably use the opportunity to authorize detentions to make a meaningful contribution to the formation of national policy in this crucial area. Despite the flaws of DECA in its present form, it is intended to protect citizens from arbitrary detention and guarantee them due process if they are detained. Far from being another example of Congress kowtowing during a crisis, a robust version of DECA could achieve those goals and inaugurate the balanced institutional participation of all three branches of government in this area of intersection between civil liberties and national security.[130]

—*James Weingarten**

---

[125] Youngstown Sheet & Tube Co. v. Sawyer, 343 U.S. 579, 635–36 (1952) (Jackson, J., concurring).

[126] *Id.* at 637.

[127] *See* H.R. 1076 § 2(7) ("United States citizens and residents *have been detained as enemy combatants* in the struggle with al Qaeda.") (emphasis added).

[128] *See* Bruce Ackerman, *The Emergency Constitution*, 113 YALE L.J. 1029, 1029–45 (2004).

[129] *See* Samual Issacharoff & Richard H. Pildes, *Emergency Contexts Without Emergency Powers: The United States' Constitutional Approach to Rights During Wartime*, 2 INT'L J. CONST. L. 296 (2004).

[130] *See* KOH, *supra* note 98.

* B.A., Yale University, 2001; J.D. Candidate, Harvard Law School, Class of 2007.

FLORIDA'S PROTECTION OF PERSONS BILL

Sixty-four years ago, Professor Prosser summarized the common law rule of self-defense applicable to both civil and criminal claims:

> [S]ince the law has always placed a higher value upon human safety than upon mere rights in property, it is the accepted rule that there is no privilege to use any force calculated to cause death or serious bodily injury to repel the threat to land or chattels, unless there is also such a threat to the defendant's personal safety as to justify self-defense.[1]

Earlier this year, the Florida House and Senate departed from this universally accepted principle. Florida's Protection of Persons Bill allows a person outside of his home to stand his ground in the face of an attack and inside of his home or vehicle against an intruder, even if there is no threat of harm.[2]

The bill was conceived of by former National Rifle Association ("NRA") President Marion P. Hammer. Florida Senator Durrell Peaden (R-Crestview) sponsored the bill in the Senate where, with vigorous backing by the NRA,[3] it passed unanimously.[4] To expedite the vote in the Florida House of Representatives, Rep. Dennis Baxley (R-Ocala) sponsored a bill identical to the Senate version.[5] After several Democrats offered two ill-fated amendments in an attempt to limit the bill,[6] the House passed the bill in its original form by an overwhelming vote margin of 94 to 20.[7] Gover-

---

[1] WILLIAM PROSSER & W. PAGE KEETON, PROSSER AND KEETON ON THE LAW OF TORTS 132–33 (5th ed. 1984); *see also* Stuart Green, *Castles and Carjackers: Proportionality and the Use of Deadly Force in Defense of Dwellings and Vehicles*, 1999 U. ILL. L. REV. 1, 33 (referring to "the universally adopted common law rule that deadly force may not be used in defense of personal property"). *See generally* 40 AM. JUR. 2D *Homicide* § 177 ("In general, it may be said that the law countenances the taking of human life in connection with the defense of property only where an element of danger to the person of the slayer is present or where the slaying is necessary to prevent the commission of a felony."); People v. Riddle, 649 N.W.2d 30, 46 (Mich. 2002) ("[T]he cardinal rule, applicable to *all* claims of self-defense, is that the killing of another person is justifiable homicide if . . . the defendant honestly and reasonably believes that he is in imminent danger of death or great bodily harm."); RESTATEMENT (SECOND) OF TORTS § 85 (1965).

[2] H.B. 249, 107th Leg. (Fla. 2005).

[3] *See* Manuel Roig-Franzia, *Florida Gun Law to Expand Leeway for Self-Defense*, WASH. POST, Apr. 26, 2005, at A1.

[4] S.B. 436, 107th Leg. (Fla. 2005).

[5] H.B. 249.

[6] Rep. Arthenia Joyner (D-Tampa) proposed A 848887 and Rep. Dan Gelber (D-Miami Beach) proposed SA 839703. Both were filed on March 30, 2005. On March 31, 2005, the Senate bill was submitted as a substitute.

[7] H.B. 249. Although the bill passed in the Florida House and Senate with bipartisan support, Rep. Gelber stated that many lawmakers supported the bill out of fear of the NRA. *See* Roig-Franzia, *supra* note 3, at A1. Giving credence to this perception is NRA executive vice president Wayne LaPierre, who told *Time Magazine*, "[p]oliticians are putting their career [sic] in jeopardy if they oppose this type of bill." Michelle Cottle, *Shoot First, Regret Legislation Later*, TIME, May 9, 2005, at 80.

nor Jeb Bush promptly signed the bill, which became effective on October 1, 2005.[8]

The first significant change that the bill makes is that it eliminates the duty to retreat before using deadly force if the person acting in self-defense is in a place where he has a right to be.[9] Previously, a person outside of his home or workplace had a duty to use every reasonable means to avoid danger, including retreat, prior to using deadly force.[10]

Although jurisdictions that adhere to the common law duty to retreat recognize the costs of requiring a person assailed to "seek dishonor in flight,"[11] the supreme value of life serves as the justification for this duty.[12] However, since the late nineteenth century, the duty to retreat has eroded as most jurisdictions have begun to view it as an unreasonable burden on societal notions of courage and dignity.[13] Florida joins these jurisdictions

---

[8] Governor Bush said he supported the measure because, "to have to retreat and put yourself in a very precarious position defies common sense." *See* Abby Goodnough, *Florida Expands Right to Use Deadly Force in Self-Defense*, N.Y. TIMES, Apr. 27, 2005, at A1. Contrary to the Governor's depiction of existing state law, Florida does not impose a duty to retreat if fleeing would jeopardize a person's safety. *See* Danford v. State, 43 So. 593, 595 (Fla. 1907) ("[I]t is the duty of a party to avoid a difficulty which he has reason to believe is imminent, *if he may do so without apparently exposing himself to death or great bodily harm.*") (emphasis added).

[9] H.B. 249, creating FLA. STAT. ANN. § 776.013(3) (West 2005), which states:

> [A] person, not engaged in an unlawful activity, who is attacked in any other place where he has a right to be has no duty to retreat and has the right to stand his or her ground and meet force with force, including deadly force if he or she reasonably believes it is necessary to do so, to prevent death or great bodily harm to himself, herself, or another or to prevent the commission of a forcible felony.

[10] State v. James, 867 So. 2d 414 (Fla. Dist. Ct. App. 2003).

[11] Joseph H. Beale, Jr., *Retreat from Murderous Assault*, 16 HARV. L. REV. 567, 577 (1903). For a discussion of the common law duty to retreat, see Regina v. Smith (1837), 173 Eng. Rep. 441 (K.B.); Regina v. Bull (1839), 173 Eng. Rep. 723 (K.B.).

[12] *See, e.g.*, State v. Shaw, 441 A.2d 561, 565 (Conn. 1981) (declaring that the duty to retreat is premised on the "recognition of . . . the great value of human life"); People v. Canales, 624 N.W.2d 918, 919 (Mich. 2001) (Corrigan, C.J., dissenting) (asserting that the sanctity of human life is the primary principle underlying the duty to retreat rule: "[h]uman life is not to be lightly disregarded, and the law will not permit it to be destroyed unless upon urgent occasion") (quoting Pond v. People, 8 Mich. 150, 173 (1860)).

[13] This trend emerged in the unsettled territories of the late nineteenth century, and decisions that removed the duty to retreat bear the imprint of the code of the West. The Ohio Supreme Court is credited with creating what has become known as the "true man" rule in *Erwin v. State*, 29 Ohio St. 186 (1876):

> Does the law hold a man who is violently and feloniously assaulted responsible for having brought such necessity upon himself, on the sole ground that he failed to fly from his assailant when he might have safely done so? [A] true man, who is without fault, is not obliged to fly from an assailant, who by violence or surprise, maliciously seeks to take his life or do him enormous bodily harm.

Id. at 199–200. For a discussion of other jurisdictions, see WAYNE R. LAFAVE & AUSTIN W. SCOTT, JR., CRIMINAL LAW § 5.7(f) (2d ed. 1986); *see also* 40 AM. JUR. 2D HOMICIDE § 164.

and now allows a person to stand his ground when attacked and to meet force with force, even if an avenue of safe retreat is within reach.[14]

Second, the bill claims to codify the castle doctrine,[15] a common law privilege that allows a person attacked within his dwelling to stand his ground.[16] This attempt to codify the castle doctrine is significant because it not only extends the conception of one's castle to include vehicles[17] but also eliminates the requirement of necessity. The bill accomplishes the latter change through Florida Statute § 776.013, which sets forth a presumption that removes the home or vehicle occupant's burden of proving that he feared for his safety.[18] Now, a person who uses deadly force against an intruder is presumed to have a reasonable fear of death or bodily injury.[19] According to the Senate Committee Report, this presumption is irrebuttable.[20] Therefore, a court will not entertain arguments showing the nonexistence of the presumed fact, even in the face of overwhelming evidence.[21] Rather, a court will direct a jury that if they find the basic fact, that the victim was unlawfully in the actor's dwelling or vehicle, to be proven, then they *must* find the presumed fact that the actor had a reasonable fear of imminent death or bodily injury. This finding in turn justifies the use of deadly force, regardless of the circumstances.[22]

The decision of the Florida Court of Appeals in *Quaggin v. State*[23] highlights the effect of this presumption. In *Quaggin*, the defendant lived

---

[14] H.B. 249, amending FLA. STAT. § 776.012(2) (2005). The bill also removes the duty to retreat when using deadly force in defense of others. *See* H.B. 249 (codified as amended at FLA. STAT. § 776.031(3) (2005)); *see also* Anthony J. Sebok, *Florida's New "Stand Your Ground" Law: Why It's More Extreme Than Other States' Self-Defense Measures, and How It Got that Way*, FINDLAW, May 2, 2005, http://writ.news.findlaw.com/sebok/20050502.html.

[15] Steve Bousquet, *Bill Would Relax Rules on Self-Defense*, ST. PETERSBURG TIMES, Feb. 29, 2005, at B1.

[16] *See, e.g.*, Regina v. Ford, *cited in* Raven's Case (1685), 84 Eng. Rep. 1078 (K.B.); People v. Tomlins, 107 N.E. 496, 497 (N.Y. 1914) (Cardozo, J.) (quoting 1 SIR MATTHEW HALE, PLEAS OF THE CROWN, 486). For a description of the castle doctrine as it has existed in Florida, see, for example, Danford v. State, 43 So. 593, 598 (Fla. 1907) (holding that a person attacked in his own home or premises may stand his ground and use such force as may appear reasonably necessary to defend himself or another from great bodily harm).

[17] H.B. 249 (codified as FLA. STAT. § 776.013(5)(c) (2005)).

[18] H.B. 249 (codified as FLA. STAT. § 776.013(1) (2005)). The presumption does not apply if the person who uses defensive force is engaged in an unlawful activity or if the person against whom the defensive force is used has a legal right to be in the dwelling, is a child or grandchild of the person using defensive force, or is a law enforcement officer. *See id.* (codified as FLA. STAT. § 776.013(2) (2005)).

[19] *Id.*

[20] Fla. S. Rep. No. 107-436, 6pt. III, at 6 (2005) (Judiciary Rep.) ("Legal presumptions are typically rebuttable. The presumptions created by the committee substitute, however, appear to be conclusive."). *Accord* Fla. H.R. Rep. No. 107-249 (2005) pt. B, at 4 (Judiciary Rep.) ("[A] person is presumed, rather than having the burden to prove, to have a reasonable fear.").

[21] FLEMING JAMES, JR., & GEOFFREY C. HAZARD, JR., CIVIL PROCEDURE § 7.9, at 248 (1st ed. 1965).

[22] *Id.*

[23] 752 So. 2d 19 (Fla. Dist. Ct. App. 2000).

in a structure surrounded by "piles of junk" and several abandoned trailers.[24] Two children found several comic books and Pez dispensers in one of these trailers and began searching for the owner so they could ask permission to take what they had found.[25] They found another structure they thought was abandoned and entered through an unlocked sliding glass door.[26] The owner of the dwelling jumped up and asked what they were doing there.[27] Before the children could answer, the owner fired at close range and killed one of the boys.[28] The Florida Court of Appeals held that the defendant had to believe that a forcible felony was being committed in order to claim self-defense.[29] The court further held that the necessity of using deadly force in response must be reasonable, i.e., the appearance of danger must have been so real that a reasonably prudent person under the same circumstances would have believed that the danger could be avoided only through the use of deadly force.[30]

The new bill eliminates these requirements. The defendant's use of force is now justified because of a conclusive presumption that the children posed a threat that was sufficient to create a reasonable fear of death or imminent bodily harm. This could be true even if the children had not forcibly entered the dwelling, as was the case here. Florida Statute § 776.013(B), a new section created by the bill, extends the presumption to a "person who uses defensive force [who] knew or *had reason to believe* that an unlawful and forcible entry or unlawful and forcible act was occurring or had occurred."[31] This drastically changes Florida law, which previously required a reasonable belief and factual grounds that the use of deadly force was necessary.[32] In *Quaggin*, for example, a previous burglary in the defendant's home or even a mistaken belief that he locked the sliding glass door would now give the defendant reason to believe that these intruders had unlawfully and forcibly entered his dwelling.

The third significant change is Florida Statute § 776.032, which grants the person using self-defense immunity from criminal prosecution and civil action.[33] In addition, § 776.032(2) states that a person who uses deadly force cannot be arrested unless there exists probable cause that the force

---

[24] *Id.* at 20.
[25] *Id.*
[26] *Id.*
[27] *Id.*
[28] *Id.*
[29] *Quaggin*, 752 So. 2d at 23.
[30] *Id.*
[31] FLA. STAT. § 776.013(B) (2005) (emphasis added).
[32] *See* Falco v. State, 407 So. 2d 203, 209 (Fla. 1981); *see also* 16 FLA. JUR. 2D *Criminal Law* § 1077 (1979). *See generally* 40 AM. JUR. 2D *Homicide* § 155.
[33] H.B. 249, 107th Leg. (Fla. 2005) (codified as FLA. STAT. § 776.032 (2005)); *cf.* RESTATEMENT (SECOND) OF TORTS § 65 (1985), which would allow the assailant to sue the person who acted in self-defense if the force used by that person were not a proportional response to the threat posed by the assailant.

used was unlawful.[34] Therefore, the parents of the slain child in *Quaggin* would have no recourse for their loss. Further, § 776.032 would require that the parents pay "reasonable attorney's fees, court costs, compensation for loss of income, and all expenses incurred by the defendant . . . if the court finds that the defendant is immune from prosecution."[35]

Rep. Baxley, the bill's sponsor in the House, and Marion Hammer, the bill's conceiver, have stated that the bill was not introduced in response to a specific case or incident but rather was an attempt to counterbalance the protection courts give to the rights of criminals vis-à-vis the rights of their victims.[36] They argue that the Protection of Persons Bill forecloses the possibility of being prosecuted for protecting the home and family.[37] Cited as support is the oft-quoted observation by Justice Holmes that "detached reflection cannot be demanded in the presence of an uplifted knife."[38] The presumption and immunity clauses of the bill protect the person who has a split second to defend his life from a prosecutor who has months to engage in detached reflection on how the defendant responded and how he should have responded.[39]

Opponents of the bill, mainly prosecutors and law enforcement officers, believe that this measure is a solution to a problem that does not exist.[40] Hammer claims the bill allows a person in her home to use deadly force against an intruder without having to ask, "Are you here to rape me and kill me or are you just here to beat me and steal my jewelry?"[41] However, in Florida the justifiable use of deadly force in self-defense has never turned on the criminal's intent, but on whether the use of deadly force was reasonable.[42] Even under the old law, an occupant would have been justified in using deadly force if he believed that an intruder posed grave danger and a reasonable person in the same circumstances would have held the same belief.[43] This was true regardless of whether the intruder was a jewel-thief or a rapist. Lawmakers and lobbyists in support of the bill did not cite any case in which this standard led to the prosecution of someone who used deadly force to protect his home. Indeed, a Florida prosecutor recently dropped charges against a homeowner who did specifically that, claiming

---

[34] H.B. 249.

[35] FLA. STAT. § 776.032(3) (2005).

[36] *See* Marion P. Hammer, *At Last, Balance Shifts Away from Criminals*, ATLANTA JOURNAL-CONSTITUTION, May 2, 2005, at 11A; *see also Talk of the Nation* (National Public Radio broadcast May 2, 2005).

[37] *See* Hammer, *supra* note 36; *see also Talk of the Nation*, *supra* note 36.

[38] Brown v. United States, 256 U.S. 335, 343 (1921).

[39] *See Talk of the Nation*, *supra* note 36; *see also* Hammer, *supra* note 36.

[40] *See* Bill Cotterell, *House Passes Self-Defense Bill*, TALLAHASSEE DEMOCRAT, Apr. 6, 2005, at A4.

[41] *See* Paul Flemming, *Bill Would Broaden Deadly Force Rights*, NEWS-PRESS (FORT MYERS, FLA.), Jan. 28, 2005, at 3B.

[42] *See* Quaggin v. State, 752 So. 2d 19, 23 (Fla. Dist. Ct. App. 2000).

[43] *Id.*

that the homeowner "clearly had the lawful right to protect his property, his wife and himself from an intruder."[44]

While removing the duty to retreat brings Florida in line with the majority rule, the conclusive presumption sets Florida apart from all other states because it contravenes an ancient and universally adopted principle that restricts the use of deadly force to an actual or threatened harm to persons. It does so not by eliminating the requirement of reasonable apprehension but through a conclusive presumption that automatically establishes it, regardless of whether it actually existed. On this point, Professor Fleming James notes:

> The conclusive presumption is not really a procedural device at all. Rather it is a process of concealing by fiction a change in the substantive law. When the law conclusively presumes the presence of B from A, this means that the substantive law no longer requires the existence of B in cases where A is present, although it hesitates as yet to say so forthrightly.[45]

Because the children in the *Quaggin* example were guilty of no crime other than trespassing, a misdemeanor,[46] a conclusive presumption that extends the justifiable use of deadly force to these circumstances conceals a change in the elements of self-defense that now allows for deadly force to be used in protection of property. As Professor Anthony Sebok observes, a conclusive presumption that only requires the unlawful entry of the intruder justifies the use of deadly force against an unarmed and helpless intruder, even if such force is used solely as an act of revenge.[47]

An introductory paragraph of the bill cites the castle doctrine as the authority for this change.[48] However, the castle doctrine is the descendant of two common law privileges, self-defense in the home and the defense of habitation, neither of which support the effect of the presumption.

The right to defend oneself in the home has always required that the occupant perceive a threat sufficient to warrant deadly force.[49] When the general privilege of self-defense was established in English common law, it was limited by the requirement of necessity.[50] Primarily, this doctrine required that the person using self-defense not be the aggressor, that the apprehension of harm be reasonable, and that no more force than is nec-

---

[44] *See* Flemming, *supra* note 41.

[45] *See* JAMES & HAZARD, *supra* note 21, at 248.

[46] FLA. STAT. § 810.08(2)(b) (2005) (classifying trespassing in a structure where there is another person as a first-degree misdemeanor).

[47] *See* Sebok, *supra* note 14 (claiming that the presumption would protect a teacher who finds a student defacing the interior of his car and kills him in response).

[48] H.B. 249, 107th Leg. (Fla. 2005) (describing the castle doctrine as "a common-law doctrine of ancient origins which declares that a person's home is his or her castle . . .").

[49] *See infra* notes 50–55 and accompanying text.

[50] *See* Beale, *supra* note 11, at 569.

essary be applied.[51] In addition, the victim of an unprovoked attack had a duty to "retreat to the wall" before using deadly force against the aggressor,[52] provided that the victim can retreat without jeopardizing his own safety.[53] It was believed that a person in his home has already retreated to the wall.[54] Therefore, a duty that would require the occupant to flee the safety of his dwelling goes beyond the traditional boundaries of necessity.[55] While this rule allowed a person to stand his ground in his home, he remained subject to the other requirements of necessity. A conclusive presumption that allows the occupant to circumvent each of these requirements demonstrates that this bill is not supported by the privilege of self-defense in the home.

The second common law privilege that comprises the modern castle doctrine, the defense of habitation, bears a closer relation to the changes made by the presumption. As old as the English feudal maxim, "A man's house is his castle,"[56] the defense of habitation seeks to protect the sanctity of the home and its occupants from the danger and indignity of having to be a "fugitive from [their] own home."[57] While the two defenses overlap in many instances, the defense of habitation is unique in that it does not require any threat of death or physical injury, only the unlawful entry of an intruder who intends to commit a felony.[58] Florida's new presumption closely tracks the defense of habitation because it also does not require

---

[51] *See* 4 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 144–46 (Wayne Morrison ed., Cavendish Publishing 2001) (1822); *see also* Gainer v. State, 391 A.2d 856, 862–63 (Md. Ct. Spec. App. 1978) (acknowledging that "[t]he castle doctrine is for defensive and not offensive purposes and does not confer a license to kill or to inflict grievous bodily harm merely because the assault takes place within the defendant's home").

[52] *See* SIR MATTHEW HALE, SUMMARY OF THE PLEAS OF THE CROWN 41–42 (P. R. Glazebrook ed., Professional Books Limited 1982) (1678); *see also* Beard v. United States, 158 U.S. 550, 560 (1895).

[53] *See supra* note 8.

[54] *See* Weiand v. State, 732 So. 2d 1044, 1049–50 (Fla. 1999).

[55] Lord Hale vividly described this privilege when he asserted that a man "assailed in his own house . . . 'need not fly as far as he can . . . for he hath the protection of his house to excuse him from flying, for that would be to give up the possession of his house to his adversary by his flight.'" People v. Tomlins, 107 N.E. 496, 397 (N.Y. 1914) (quoting HALE, PLEAS OF THE CROWN, *supra* note 16, at 486); *see also* United States v. Peterson, 483 F.2d 1222, 1235–36 (D.C. Cir. 1973) (acknowledging that the duty to retreat "was never intended to enhance the risk to the innocent").

[56] *See, e.g.*, Semayne's Case, (1604) 77 Eng. Rep. 194, 195 (K.B.) ("[T]he house of everyone is to him as his Castle and Fortress, as well as for defence against injury and violence, as for his repose.").

[57] People v. Tomlins, 107 N.E. 496, 497 (N.Y. 1914). For more on the history of the defense of habitation doctrine, see BLACKSTONE, *supra* note 51, at 173–79; 1 EDWARD COKE, THE SELECTED WRITINGS AND SPEECHES OF SIR EDWARD COKE 135–41 (Steve Sheppard ed., Liberty Fund 2003) (1644); 2 SIR FREDERICK POLLOCK & FREDERIC WILLIAM MAITLAND, HISTORY OF ENGLISH LAW BEFORE THE TIME OF EDWARD I 478–84 (2d ed. 1959) (1895); 3 SIR JAMES FITZJAMES STEPHEN, A HISTORY OF THE CRIMINAL LAW OF ENGLAND 13–15 (1883).

[58] *See* BLACKSTONE, *supra* note 51, at 175–79; *see also* Green, *supra* note 1, at 16.

reasonable apprehension before deadly force is used against an unlawful intruder.

To say that the presumption codifies this ancient defense, however, cuts too far. In his Appeals on Homicide, dated circa 1290, Britton noted that the defense of habitation cannot be claimed by an occupant who uses deadly force with felonious intent.[59] Moreover, the occupant has the burden of proving that the use of deadly force was necessary to protect his family and habitation.[60] Blackstone appears to provide a more lenient rule by bypassing any discussion of necessity and proclaiming that "the law of England has so particular and tender a regard to the immunity of a man's house, that it styles it his castle, and will never suffer it to be violated with impunity . . . ."[61] However, he quickly reins in the effect of this forceful language by stating that deadly force is justified only if the "breaking and entry [of an unlawful intruder is] . . . with a felonious intent, otherwise it is only a trespass."[62] By limiting the defense of habitation to situations in which an intruder has "intent to commit a robbery, a murder, a rape, or any other felony,"[63] Blackstone, like Britton five centuries before him, acknowledges that the defense of habitation requires reasonable apprehension and is subject to the requirement of necessity.

This is where the presumption diverges from the common law privilege. Once the defendant establishes that an unlawful and forcible entry occurred, or he had reason to believe it had, evidence of felonious intent by the defendant or lack of necessity becomes *per se* inadmissible.[64] Accordingly, a presumption that eliminates the requirement of reasonable apprehension departs from the defense of habitation.

Florida is not the first state to reinvent the defense of habitation in broad terms.[65] Indeed, the defense has enjoyed a recent revival as lawmakers address the public's fear of crime with measures that make use of the rhetorical force of terms such as "castles" and "fortresses."[66] Proponents of the Protection of Persons Bill and other similar bills point out that Black-

---

[59] 1 BRITTON, F. M. NICHOLS'S TRANSLATION OF BRITTON 113 (Wm. W. Gaunt & Sons, Inc. 1983) (1530ca). According to one authoritative work, Britton's treatise on English law purports to be a direct codification and enactment of the law by Edward I and was a practical book for lawyers in the royal courts. D. WALKER, THE OXFORD COMPANION TO LAW 154 (1980). Britton's predecessor, Bracton, also recognized the requirement of necessity within the context of defending habitation. *See* POLLOCK & MAITLAND, *supra* note 57, at 478 ("Bracton in his text would allow a man to slay a housebreaker, if to do so was a necessary act of self-defense.").

[60] 1 BRITTON, *supra* note 59, at 113.

[61] BLACKSTONE, *supra* note 51, at 178.

[62] *Id.*

[63] *Id.*

[64] *See supra* notes 18–20 and accompanying text.

[65] *See, e.g.*, LA. REV. STAT. ANN. § 14:20(4) (West Supp. 1998). For an in-depth analysis of the statute, see Green, *supra* note 1, at 16.

[66] *See* Green, *supra* note 1, at 16 (commenting on the adoption of permissive versions of the defense of habitation in politically popular laws with names like "Make My Day," "Shoot the Burglar," and "Shoot the Carjacker").

stone's defense of habitation can be claimed without having to prove that the use of deadly force was necessary.[67] Although Blackstone requires that the intruder must have the intent to commit a felony, one can imagine the case of an unarmed and unimposing robber. Because Blackstone does not explicitly limit the use of deadly force in this circumstance, it is argued that there is precedent for allowing deadly force to be used in the home free from the requirement of necessity.

This reading is problematic for several reasons. To assume that Blackstone's silence on this unlikely scenario serves as an implicit endorsement goes against other common law authorities, such as Britton, who specifically addressed this point. In addition, Blackstone characterizes an intrusion committed by someone who lacks the intent to commit one of the predicate felonies as "only a trespass," for which deadly force would not be justified.[68] This establishes that a violation of property rights, by itself, is insufficient to warrant the use of deadly force. Moreover, this reading is at odds with Blackstone's writing on an act that satisfies the requirements for felonious murder, "the highest crime against the law of nature that man is capable of committing."[69] If Blackstone had recognized an exception to the gravest of crimes, he would have done so explicitly.

This reading also ignores the justifications on which the defense of habitation rests. Common law recognition of the sanctity of the home was not rooted in the property rights of the owner but in the protection a home provides.[70] The right to use deadly force against unlawful intruders is a reflection of an era when occupants "were compelled to protect themselves in their habitations by converting them into holds of defence."[71] Once flushed from their castle, the occupants were vulnerable to dangers in the wild, which were at least as serious as the threat of harm from the intruder.[72] In addition, the occupants who fled stood to lose possession of their dwelling to the intruder.[73] The defense of habitation therefore allowed the occupants to use deadly force to protect themselves and their families from the dangers of the outside world, even if the intruder himself were not instilling a fear of death or bodily harm. Today, dispossession by conquest is no longer a legitimate fear, and the modern conveniences of a tamed coun-

---

[67] *See* Hammer, *supra* note 36, at 11A.

[68] *See* BLACKSTONE, *supra* note 51, at 178.

[69] *See id.* at 140.

[70] *See* State v. Carothers, 594 N.W.2d 897, 900 (Minn. 1999) (noting that the defense of habitation has its roots in the "common law recognition of the home's importance," and citing the established belief that "the house has a peculiar immunity [in] that it is sacred for the protection of [a person's] family.") (quoting State v. Touri, 112 N.W. 422, 424 (Minn. 1907)); *see also* State v. Hare, 575 N.W.2d 828, 832 (Minn. 1998) (noting that, at common law, "defense of the home [was] considered equivalent to defense of life itself") (internal citation omitted).

[71] *See* 1 J. BISHOP, NEW COMMENTARIES ON THE CRIMINAL LAW 517 (8th ed. 1892).

[72] *See, e.g.*, Beale, *supra* note 11, at 580 (explaining that to require retreat would leave resident exposed to dangers against which the home was supposed to protect).

[73] *See* HALE, PLEAS OF THE CROWN, *supra* note 16, at 486.

tryside greatly reduce the threat to occupants who are forced to flee their home. Furthermore, increased population density and technology that allows for instantaneous communication, both of which allow for a quick law enforcement response, further diminish the relevance of the original justifications.[74]

This is not to say the defense of habitation is wholly irrelevant today—modern advances reduce the degree of danger to the person forced to flee his home, but do not eliminate it. The indignity suffered by fleeing from one's home is no less for the owner of a condo than it was for the lord of a castle. In an attempt to balance the desire to protect occupants from these dangers and indignities with the interest in preserving the sanctity of life, most jurisdictions recognize a limited version of the defense of habitation. The majority rule, to which Florida adhered prior to the enactment of the Protection of Persons Bill,[75] allows an occupant to use deadly force only to prevent the commission of a felony within the home.[76] A significant number of jurisdictions have further limited the privilege to the prevention of "forcible" or "deadly" felonies.[77] These measures recognize the right of the occupant to be free from unlawful intrusions but protect the sanctity of life by authorizing the use of deadly force only in situations where the intrusion is followed by a crime of a serious and dangerous nature.

Florida's Protection of Persons Bill blazes a trail in the opposite direction by expanding the rights of persons in their homes and vehicles beyond the limits of the common law privilege. The absolute right created by the bill surpasses even the expansive privileges granted by the jurisdictions that have recently reintroduced the defense of habitation.[78]

---

[74] The reasoning that supported the defense is also poorly suited for modern society. As Blackstone notes, crimes against habitations left the occupant with a "natural right of killing the aggressor" because these invasions interfered with "that right of habitation, which every individual might acquire, even in a state of nature." BLACKSTONE, *supra* note 51, at 175. In the state of nature, argued Blackstone, an invasion of the habitation was "sure to be punished with death" by a stronger inhabitant. *Id.* "Hence, the civil laws of England took their cue from the natural law and punished the violator of the home if its inhabitant was too feeble to do so." Thomas Katheder, *Case Note: Criminal Law—Lover and Other Strangers: Or, When Is a House a Castle?*, 11 FLA. ST. U. L. REV. 465, 484 (1983). This justification has little relevance to a society that no longer fashions its laws based on what would take place in a world without criminal and civil laws.

[75] *See* FLA. STAT. § 782.02 (2005).

[76] *Id.*

[77] *See, e.g.*, Law v. Maryland, 318 A.2d 859, 867 (Md. Ct. Spec. App. 1974) (stating that the felonies which justify the use of deadly force "are such and only such as are committed by forcible means, violence, and surprise such as murder, robbery, burglary, rape or arson") (quoting 1 WHARTON'S CRIMINAL LAW AND PROCEDURE § 206, at 454 (Anderson ed. 1957)); State v. McIntyre, 477 P.2d 529, 534–35 (Ariz. 1970) (holding that a felony must reasonably create fear of great bodily injury); People v. Ceballos, 526 P.2d 241, 245 (Cal. 1974) (holding that a felony must be forcible or atrocious).

[78] *Cf.* LA. REV. STAT. ANN. § 14:20(4) (2005) (requiring that the occupant reasonably believe that the use of deadly force is necessary to prevent the entry or to compel the intruder to leave the premises or motor vehicle); DEL. CODE ANN. tit. 11, § 469 (2005) (requiring that the encounter be unexpected and sudden, that the occupant reasonably believe that the intruder would inflict personal injury upon him or others in the dwelling, or that

These unprecedented changes carry practical implications that will affect both occupant and intruder. By creating a right that is predicated not on necessity but on the respective property interests of occupant and intruder, the bill resurrects an area of confusion regarding the use of deadly force against cohabitants. Because the presumption is rooted in the defense of habitation, which requires an unlawful entry, newly created Florida Statute §776.013(2)(a) states that the presumption does not apply if the person against whom deadly force is used has a legal right to be in the dwelling.[79]

Therefore, the Florida Supreme Court's rulings on the use of deadly force against cohabitants still govern in this circumstance. In the most recent ruling on the issue, *Weiand v. State*,[80] the court reversed an earlier decision, *State v. Bobbitt*,[81] which, like the Protections of Persons Bill, created different rights for occupants who are attacked by cohabitants and those attacked by intruders. In *Bobbitt*, the court held that an occupant had a duty to retreat from his home when both he and the attacker have equal rights to be there.[82] In his *Bobbitt* dissent, Justice Overton illustrated the "illogical distinction" made by the majority.[83] If a son attacks his mother while living at home, the mother has a duty to retreat before she can use deadly force, but, if the son is not residing in the home, the mother has no duty to retreat before such force is used.[84]

Seventeen years later, the court in *Weiand* took judicial notice of the confusion created as a result of this illogical distinction and in particular how it affected victims of domestic violence.[85] Specifically, the court cited studies revealing "the threat of separation is usually the trigger for violence."[86] Recognizing that its decision in Bobbitt was "grounded upon the sanctity of [the aggressor's] property and possessory rights, rather than the sanctity of human life,"[87] the court in *Weiand* imposed on a person attacked by a cohabitant a limited duty to retreat within the residence to the extent reasonably possible but no duty to flee the residence.[88]

Similar to the illogical distinction created by the court in *Bobbitt*, the conclusive presumption seeks to protect people attacked in their homes and

---

the occupant demand that the intruder disarm or surrender); IND. CODE § 35-41-3-2(b) (2005) (requiring that the occupant reasonably believe that deadly force is necessary to prevent or terminate the intruder's unlawful entry into or attack on the occupant's dwelling or curtilage).

[79] *See* H.B. 249, 107th Leg. (Fla. 2005).

[80] 732 So. 2d 1044, 1052 (Fla. 1999).

[81] 415 So. 2d 724 (Fla. 1982).

[82] *Id.* at 726.

[83] *Id.* at 728.

[84] *Id.*

[85] *See Weiand*, 732 So. 2d at 1052.

[86] *Id.* at 1053 (quoting Maryanne E. Dampmann, *The Legal Victimization of Battered Women*, 15 WOMEN'S RTS. L. REP. 101, 112–13 (1993)).

[87] Weiand v. State, 732 So. 2d 1044, 1052 (Fla. 1999).

[88] *Id.* at 1056.

vehicles but does not extend this protection to occupants who defend themselves against cohabitants. Similar to the mother in Justice Overton's example, the mother who uses deadly force against her son must only prove that the son was no longer residing in the home and that she had reason to believe that the son had unlawfully and forcibly entered in order to establish that she held a reasonable fear of death or imminent bodily harm. If the son has yet to move out, the mother's duty changes drastically. The absolute right afforded her by the bill vanishes; instead, she has a duty to retreat to another room in the house before using deadly force, which can only be used in the face of an actual, not presumed, fear of death or bodily harm. This illogical distinction becomes even more pronounced when contrasted with the right to stand one's ground in public. Residents and visitors of the state have the right to stand their ground on a public street,[89] but a victim of domestic violence must retreat to another room if attacked within the sanctity of her own home.[90]

Leaders in Florida's law enforcement community fear that citizens may mistakenly assume that they have total immunity in situations in which the bill does not apply.[91] Six million gun owners in Florida have been given a right to take life even in circumstances where law enforcement officers would not be justified in doing so.[92] Extensive local and national news coverage,[93] as well as the unflagging efforts of the NRA,[94] has informed them of this right, yet nowhere has it been mentioned in the media that the presumption that gives them this right does not apply to the most common uses of self-defense in the home. The illogical distinction, combined with a misinformed public, recreates the confusion that plagued the state before the Florida Supreme Court's ruling in *Weiand* and unfortunately stands poised to claim victims on both sides of the gun.

There are also costs associated with this newly created right in situations where the occupant fully understands the presumption and its limits. While the presumption seeks to protect the law-abiding gun owner who is forced to make a split-second decision, granting him an absolute right to protect his home or vehicle removes whatever cautionary effect the law may have previously had. If a person has reason to believe someone has unlawfully and forcibly entered his home, the bill now allows him to act with impunity in situations where the exercise of discretion often means the difference between life and death.

---

[89] *See supra* notes 9–14 and accompanying text.

[90] *See Weiand* at 1056–57.

[91] The Miami and St. Petersburg police chiefs, as well as the Broward County Sheriff, publicly opposed the bill. *See* Goodnough, *supra* note 8.

[92] *See* Editorial, *The Shoot First State*, WASH. POST, May 1, 2005, at B6. A law enforcement officer cannot use deadly force in a non-deadly situation. *See also* Mercado v. City of Orlando, 407 F.3d 1152, 1156 (11th Cir. 2005).

[93] For local coverage, see *supra* notes 15, 40, 41 and accompanying text. For national coverage, see *supra* notes 3, 7, 8, 14, 36 and accompanying text.

[94] *See* Manuel Roig-Franzia, *supra* note 3, at A1.

CONCLUSION

Although most of the debate surrounding the Protection of Persons Bill, both in the legislature and in the media, has concerned the right to stand one's ground in public, it is the conclusive presumption that marks a radical departure from long-held conceptions of proportionality and necessity.

The fact that the right to stand one's ground has drawn significant criticism in local and national news is surprising given that this right has already been codified in most states.[95] Without reference to the long history of this right in other states, opponents have claimed that the measure legalizes dueling and "will possibly turn the state into the O.K. Corral."[96] These opponents are primarily concerned with unforeseen applications of the measure, such as the use of the privilege of non-retreat between two gang members who both have the legal right to be on a crowded Miami street.[97] These concerns fail to take into consideration the many limits placed on a citizen's ability to stand his ground. For example, the bill states that the privilege is not available to those who are engaged in a criminal activity.[98] In addition, the right to stand one's ground is only available to persons who are without fault in the altercation.[99] Further limiting the privilege of non-retreat is the requirement of proportionality. As Rep. Baxley noted, "[y]ou can only do what somebody does to you."[100] These restrictions on the right to stand one's ground in other states have prevented it from applying to the types of shootouts mentioned by opponents. Similarly, these restrictions will most likely reduce the potential for abuse and misapplication in Florida.

The conclusive presumption, however, is a profound departure from common law. The right to take the life of another is the most powerful right that a state can recognize. Sanford Kadish's theory on the right to resist unlawful aggression describes the standard that must be satisfied in order to exercise this right.[101] When a society emerges from a state of nature, the state recognizes a person's fundamental freedom to preserve himself against aggression and agrees to protect that right from violation by others.[102] However, when a citizen is confronted with an aggressor that is capable of

---

[95] *See supra* notes 13 & 93 and accompanying text.

[96] The first concern was raised by Rep. Gelber, and the second was raised by Rep. Irv Slosberg (D-Boca Raton). *See* Cotterell, *supra* note 40.

[97] *See* Bousquet, *supra* note 15.

[98] H.B. 249, 107th Leg. (Fla. 2005).

[99] *See* Beard v. United States, 158 U.S. 550, 564 (1895). *See generally* 40 AM. JUR. 2D *Homicide* § 164.

[100] Cotterell, *supra* note 40; *see also* State v. Renner, 912 S.W.2d 701, 704 (Tenn. 1995) ("[T]he 'true man' rule implies no license for the initiation of a confrontation or an unreasonable escalation of a confrontation in progress.").

[101] *See* Sanford H. Kadish, *Respect for Life and Regard for Rights in the Criminal Law*, 64 CAL. L. REV. 871, 884–85 (1976).

[102] *Id.*

attacking before the state can intervene, the right to resist unlawful aggression remains with the citizen.[103] Therefore, only a threat to life or limb, and not property, allows a citizen to exercise the right that the State is unable to protect. This conception of necessity is a thread that has carried through Anglo-American self-defense jurisprudence ever since the Norman Conquest.[104] While other measures have caused the thread to fray in recent years, the Protection of Persons Bill cuts it in two. The bill's conclusive presumption states that reasonable apprehension sufficient to warrant deadly force per se exists if an intruder unlawfully and forcibly enters a dwelling or vehicle.[105] In doing so, it transgresses the universally adopted common law rule that deadly force may not be used in defense of personal property.[106] Over two centuries ago, Blackstone observed that "many learned and scrupulous men have questioned the propriety, if not lawfulness, of inflicting capital punishment for simple theft."[107] Converting the lesser crime of trespassing into a capital offense raises the same concern.

A triumphant NRA has vowed to push the passage of similar legislation in every statehouse in the country.[108] If they follow a plan of attack in these states similar to the one used in Florida, they will exert tremendous pressure on lawmakers to pass the bill as quickly as possible. In Florida, this strategy allowed the bill to stand unopposed by gun control groups who were not able to mobilize in time.[109] In addition, it brought the bill in both houses to a vote before the subcommittees analyzing the measure could fully appreciate the effects of the amendments. As the NRA brings similar legislation to other states, legislators should carefully examine the justifications and implications for a bill that violates fundamental principles of self-defense.

—*Daniel Michael*[*]

---

[103] *Id.*

[104] *See* Beale, *supra* note 11, at 567–68.

[105] *See supra* note 64 and accompanying text.

[106] *See* Green, *supra* note 1, at 33.

[107] BLACKSTONE, *supra* note 51, at 187.

[108] This bill is the "first step of a multi-state strategy," declared NRA Executive Vice President Wayne LaPierre. Roig-Franzia, *supra* note 3, at A1. "We start with the red and move to the blue," commented LaPierre. Cottle, *supra* note 7, at 80. LaPierre hopes to capitalize on a political climate dominated by conservative opponents of gun control at the state and national levels. Roig-Franzia, *supra* note 3, at A1.

[109] "I'm in absolute shock," said Sarah Brady, chair of the Brady Center to Prevent Gun Violence. "If I had known about it, I would have been down there." Roig-Franzia, *supra* note 3, at A1.

[*] B.A., New York University, 1998; J.D. Candidate, Harvard Law School, Class of 2007.

MEDICAL MALPRACTICE NON-ECONOMIC DAMAGES CAPS

The medical malpractice debate has been presented to Americans through contrasting images of children crippled by physician error and mothers forced to deliver their babies in bathrooms because of physician flight. While the issues involved in what has been called the medical malpractice crisis are much more nuanced than these images convey, many agree that drastic increases in the cost of medical malpractice insurance have tested the wherewithal of parties on both sides of the stethoscope. In order to temper what some view as a concerning exodus of medical care providers from the industry,[1] about half of the states have passed medical malpractice tort reform legislation.[2] Recently, federal malpractice reform efforts have emerged as well.[3]

Federal and state medical malpractice reform bills vary greatly, but they widely share one common element—a $250,000 cap on non-economic damages. The question addressed by this Recent Development is, why is $250,000 the going rate? The impact of the $250,000 non-economic damages cap has been documented, but the origin of the cap has largely remained a mystery. As Congress considers adopting new legislation, it must cut through the common rhetoric of the medical malpractice tort reform debate to find an equitable solution. An examination of the impact of the $250,000 cap and an understanding of its genesis should inform Congress's efforts.

I. FEDERAL REFORM

For the past few years, Congress has unsuccessfully attempted to address growing concerns surrounding medical malpractice litigation, insurance rates, and the retention of physicians.[4] In March 2005, Senate Majority Leader Bill Frist proclaimed that he would seek medical malprac-

---

[1] Some argue that no support exists for the alleged mass exodus of doctors from the medical profession. The Government Accountability Office recently reported that "[t]he number of physicians in the United States increased about 26 percent from 1991 to 2001, twice as much as the nation's population." U.S. GOV. ACCOUNTABILITY OFFICE, PHYSICIAN WORKFORCE 2 (2003), *available at* http://www.gao.gov/new.items/d04124.pdf. *See also* Thomas O. McGarity et al., *The Truth about Torts: An Insurance Crisis, Not A Lawsuit Crisis* 13 (2005) (A Center for Progressive Reform White Paper).

[2] *Cf. Assembly Third Reading on AB 1380, California Assembly* (May 24, 1999) at 7; CONGRESSIONAL BUDGET OFFICE, THE CONGRESS OF THE UNITED STATES, THE EFFECTS OF TORT REFORM: EVIDENCE FROM THE STATES 6 (2004), *available at* http://www.cbo.gov/ftpdocs/55xx/doc5549/Report.pdf.

[3] *See, e.g.*, S. 354, 109th Cong. (2005); H.R. 534, 109th Cong. (2005).

[4] The HEALTH Act of 2004, modeled after MICRA, limits non-economic damages. H.R. 4280, 108th Cong. (2004). Although it passed in the House of Representatives, it failed to become law because its counterpart, "The Patients First Act of 2003," failed to pass in the Senate. William Gunnar, *Is There an Acceptable Answer to Rising Medical Malpractice Premiums?*, 13 ANN. HEALTH L. 465, 492 (2004).

tice tort reform with a cap on non-economic damages, but he highlighted that the biggest challenge is "getting people to the table" to discuss such federal reform.[5]

In 2005, members of the House and Senate followed state medical malpractice tort reform efforts by proposing House Bill 534 and Senate Bill 354 to address the ever-increasing insurance premiums confronting physicians.[6] These bills seek to shorten the statute of limitations for filing a claim, place limitations on access to punitive damages, repeal the collateral source rule, restructure settlement by requiring periodic payment for future damages, cap contingency fees, abolish joint liability, and cap non-economic damages at $250,000.[7] In addition, House Bill 534 limits the liability of manufacturers, distributors, suppliers, and providers of medical products that comply with FDA standards,[8] while Senate Bill 354 prohibits a health care provider from being named as a party in product liability or class action lawsuits pertaining to any FDA-approved drugs or devices he prescribes.[9]

One recent federal reform effort that has already been passed by the House is House Bill 5, the "Help Efficient, Accessible, Low-cost, Timely Healthcare (HEALTH) Act of 2005,"[10] a replica of House Bill 534. Representative Phil Gingrey (R-Ga.) introduced the bill on July 21, 2005, and only a week later the House of Representatives approved the measure, which would cap non-economic damages in medical malpractice cases to $250,000, restrict contingency fees, require periodic payment of future medical damages, and set a statute of limitations relating to discovery and manifestation of injury.[11] While similar measures have previously passed in the House of Representatives, the true test has been and will continue to be the Senate.[12] House Bill 5's Senate counterpart, Senate Bill 4, the "Healthy America Act of 2005," is now pending in the Senate Committee on Finance.[13] Introduced by Senator Frist, the act closely mirrors the reform scheme of House Bill 5.[14]

---

[5] *Frist Signals Areas of Concession*, CAPP News (Californians Allied for Patient Protection, Sacramento, Cal.), Mar. 14, 2005, at 1, *available at* http://www.micra.org/CAPPnews031405.pdf [hereinafter *Areas of Concession*].

[6] H.R. 534; S. 354. Senate Bill 354 is sponsored by Sen. John Ensign (R-Nev.), and House Bill 534 is sponsored by former Rep. Christopher Cox (R-Cal.).

[7] H.R. 534; S. 354.

[8] H.R. 534 § 7.

[9] S. 354 § 8.

[10] H.R. 5, 109th Cong. (2005) (sponsored by Representative Phil Gingrey).

[11] *Id.*; *Federal Update*, CAPP NEWS (Californians Allied for Patient Protection, Sacramento, Cal.), Aug. 8, 2005, at 2, *available at* http://www.micra.org/ augustnews.pdf [hereinafter *Federal Update*].

[12] Past HEALTH Acts have failed in the Senate despite passage in the House of Representatives. *See, e.g.*, HEALTH Act of 2004, H.R. 4280, 108th Congress (2004); HEALTH ACT of 2002, S. 2793, 107th Cong. (2002).

[13] S. 4, 109th Cong. (2005).

[14] *Federal Update*, *supra* note 11, at 2.

Past medical malpractice tort reform bills have included a wide range of damage-related provisions, from elimination of the collateral source rule[15] to caps on all damages awarded. It is important to understand where non-economic damages fit within the palette of damage options in order to understand the impact of caps on this type of award. Economic damages include but are not limited to costs already incurred or likely to be incurred by the claimants, such as medical care expenses, past wage loss, and estimated losses in future earnings.[16] Non-economic damages include past and future pain, suffering, emotional distress, mental anguish, disfigurement, physical impairment, loss of consortium, loss of companionship, loss of parental guidance, loss of enjoyment of life, humiliation, embarrassment, inconvenience, and loss of society.[17] Though non-economic damages may be regarded as lacking a standard, they serve as a measure of juries' "conscience" by allowing juries to select a monetary figure they consider to be fair.[18]

While lumping non-economic damages into the same pot as punitive damages is tempting to many, punitive damages remain a separate and distinct form of payment. Punitive damages punish and deter wrongdoers. They also prompt juries to focus on the defendant's wealth and the reprehensibility of the conduct rather than on fairly compensating for a wrong, as is the focus of non-economic damages.[19] Punitive damages, however, remain on the shelf for most medical malpractice cases in California, awarded only when a claimant can prove by clear and convincing evidence that the health care provider was guilty of oppression, fraud, or malice.[20]

---

[15] *See* H.R. 4280 § 6. The collateral source rule is a common law rule that prohibits defendants from introducing evidence at trial to show that the plaintiff has received injury compensation from other sources. *See also* Brandon Van Grack, *The Medical Malpractice Liability Limitation Bill*, 42 HARV. J. ON LEGIS. 299, 305 (2005) (discussing the elimination of the collateral source rule in the HEALTH Act of 2004).

[16] NICHOLAS PACE ET AL., CAPPING NON-ECONOMIC AWARDS IN MEDICAL MALPRACTICE TRIALS 7 (RAND Institute for Civil Justice 2004).

[17] *See id.*

[18] OFFICE OF DISABILITY, AGING AND LONG-TERM CARE POLICY, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, CONFRONTING THE NEW HEALTH CARE CRISIS: IMPROVING HEALTH CARE QUALITY AND LOWERING COSTS BY FIXING OUR MEDICAL LIABILITY SYSTEM—THE INCREASINGLY UNPREDICTABLE, COSTLY, AND SLOW LITIGATION SYSTEM IS RESPONSIBLE 8 (2002) [hereinafter CONFRONTING THE NEW HEALTH CARE CRISIS]. ("The perceived problem of pain and suffering awards is not simply the amount of money expended, but also the erratic nature of the process by which the size of awards is determined. Juries are simply told to apply their 'conscience' in selecting a monetary figure they consider to be fair.").

[19] Joseph C. Chambers, *In re Exxon Valdez: Application of Due Process Constraints on Punitive Damages Awards*, 20 ALASKA L. REV. 195, 200 (2003).

[20] *Medical Injury Compensation Reform Act of 1975: Hearing on A.B. 1380 Before the Cal. Assemb. Comm. on the Judiciary*, 1999-2000 Leg. 10 (May 25, 1999), *available at* http://www.leginfo.ca.gov/pub/99-00/bill/asm/ab_1351-1400/ab_1380_cfa_19990525_133157_asm_comm.html.

II. THE ORIGIN OF THE $250,000 CAP

Congressional efforts to reform medical malpractice tort litigation are not unique. Past bills and those currently pending in the House and Senate resemble several state reform efforts, including California's Medical Injury Compensation Reform Act of 1975 (MICRA), which set the precedent for medical malpractice tort reform.[21]

In that year, two of California's largest insurers, Travelers and Argonaut Insurance, proffered rate hikes as steep as 342%[22] and 350%,[23] respectively, for medical malpractice insurance coverage. Nearly half of the doctors in Northern California failed to show up to work on May 1, 1975, and the empty clinics and hospitals were the result of Argonaut Insurance's refusal to provide group coverage to Northern Californian doctors.[24] That refusal effectively quadrupled doctors' insurance costs by forcing them to buy individual policies.[25] Meanwhile, in Southern California, Traveler's Indemnity Company warned Los Angeles physicians of

---

[21] The California Civil Code states:

> In any action for injury against a health care provider based on professional negligence, the injured plaintiff shall be entitled to recover non-economic losses to compensate for pain, suffering, inconvenience, physical impairment, disfigurement and other non-pecuniary damage. (b) In no action shall the amount of damages for non-economic losses exceed two hundred fifty thousand dollars ($250,000).

CAL. CIV. CODE § 3333.2(a) (West 1975). The cap applies in cases of injury or death, and it allows only one $250,000 recovery in a wrongful death case. *See* Yates v. Pollock, 194 Cal. App. 3d 195, 200 (Cal. Ct. App. 1987). Authority exists, however, for allowing separate caps for the patient and for a spouse claiming loss of consortium. Atkins v. Strayhorn, 223 Cal. App. 3d 1380, 1394 (Cal. Ct. App. 1990). The cap on non-economic damages has been upheld as constitutional. Fein v. Permanente Med. Group, 695 P.2d 665, 680 (Cal. 1985).

[22] *Medical Injury Compensation Reform Act of 1975: Hearing on A.B. 1 as Amended in the Second Extraordinary Sess. Before the Cal. S. Comm. on Insurance and Financial Institutions.*, 1975-1976 Leg. 1 (Aug. 6, 1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files) [hereinafter *Aug. 6 Hearing on A.B. 1 Before the Cal. S. Comm. on Insurance and Financial Institutions*] ("Effective November 1, 1975, approximately 5,600 Northern California physicians insured by Travelers Insurance Company will face a 342% rate increase.").

[23] Barry Keene, *California's Medical Malpractice Crisis*, *in* A LEGISLATOR'S GUIDE TO THE MEDICAL MALPRACTICE ISSUE 29 (David G. Warren & Richard Merritt eds., 1976).

[24] Lacey Fosburgh, *Doctors Limit Care in Protest on Coast*, N.Y. TIMES, May 2, 1975, at A1 [hereinafter *Doctors Limit Care*]. *See also* Lacey Fosburgh, *Physician Strike May Be Widened*, N.Y. TIMES, May 26, 1975, at A56; ALBERT J. LIPSON, MEDICAL MALPRACTICE: THE RESPONSE OF PHYSICIANS TO PREMIUM INCREASES IN CALIFORNIA 1 (RAND Corporation 1976); PACE ET AL., *supra* note 17, at 5; *cf.* Letter from Elmer Low, President, California Trial Lawyers Association, to John Miller, Chairman, Judiciary Committee (June 16, 1975) (on file with Cal. Comm. on the Judiciary A.B. 1xx 1975 files) ("The former president of Argonaut Insurance Company has admitted that a disastrous loss in the market was a major factor in that company's decision to impose a 382% increase.").

[25] *Doctors Limit Care*, *supra* note 24, at A1.

a proposed five-fold increase in insurance rates.[26] During that same period, CNA Insurance announced a 190% increase.[27]

California lawmakers faced a time crunch in their quest to address the insurance premium crisis. In an effort to provide short- and long-term solutions, Governor Jerry Brown called a second special session, also known as a second extraordinary session, of the legislature on Saturday, May 17, 1975. Governor Brown explained, "The inability of doctors to obtain insurance at reasonable rates is endangering the health of the people of this state and threatens the closing of many hospitals. The longer-term consequences of such closings could seriously limit the healthcare provided to hundreds of thousands of our citizens."[28]

During the special session, Assembly Health and Human Services Committee Chairman Barry Keene proposed Assembly Bill 1xx[29] to contain sharply rising premiums. Assembly Bill 1xx became the series of statutes known as MICRA that emerged from the special session to solve the insurance premium crisis.[30] MICRA permitted future damages awards in excess of $50,000 to be paid in periodic payments, limited attorney fees in medical malpractice litigation, eliminated the collateral source rule, and imposed a non-indexed[31] $250,000 cap on non-economic damages.[32] With such a cap, the court automatically limits jury awards of non-economic damages that exceed $250,000 without informing the jury of the cap prior to its imposition. Despite efforts to reform the non-indexed cap on non-economic damages, the $250,000 cap has existed in California since the 1975 enactment of MICRA.[33]

In a 2005 interview, Barry Keene, MICRA's author, recounted the 1975 legislative special session from which MICRA was born.[34] He noted

---

[26] *Hearing on AB 1380 Before the Assembly Committee on the Judiciary*, *supra* note 20, at 6.

[27] *Id.*

[28] CALIFORNIANS ALLIED FOR PATIENT PROTECTION, MICRA: CALIFORNIA'S LANDMARK HEALTHCARE LIABILITY LAW: A NATIONAL MODEL FOR ASSURING ACCESS TO AFFORDABLE HEALTHCARE 4 (2005), *available at* http://www.micra.org/MICRAHandbook.pdf [hereinafter LANDMARK HEALTHCARE LIABILITY LAW].

[29] AB 926 was the progenitor of AB 1xx. Keene, *supra* note 23, at 30.

[30] MICRA had strong bipartisan support when it was passed. The legislation passed the 80-member Assembly by a vote of 67-8 and passed the 40-member Senate by a vote of 34-4. Landmark Healthcare Liability Law, *supra* note 28, at 4.

[31] Indexing a monetary cap would allow the adjustment of the monetary amount so that it reflects changes in the cost of living consistent with the Consumer Price Index.

[32] Gunnar, *supra* note 4, at 484.

[33] *Cf. Medical Injury Compensation Reform Act of 1975: Hearing on A.B. 1380 Before the Cal. S. Comm. on the Judiciary*, 1999-2000 Leg. 2 (July 13, 1999), *available at* http://www.leginfo.ca.gov/pub/99-00/bill/asm/ab_1351-1400/ab_1380_cfa_19990714_123743_sen_comm.html [hereinafter *July 13 Hearing on A.B. 1380 Before the Cal. S. Comm. on the Judiciary*] (demonstrating that if AB 1380 had passed in the California Assembly, it would have adjusted the cap each year to reflect the cumulative percentage change in the Consumer Price Index).

[34] *Perspectives: An Interview With MICRA Author Barry Keene*, CAPP NEWS (Californians Allied for Patient Protection, Sacramento, Cal.), Apr. 26, 2005, at 1, *available at*

some alarming circumstances, including the fact that some physicians had chosen to "go bare" instead of paying higher premiums.[35] As tension grew in 1975 over what would be the legislative solution, Keene hoped that his grab bag of various malpractice reform mechanisms would have a reasonable chance of avoiding sabotage by any single interest group.[36] During the legislative battle, Keene asserted, "The malpractice crisis can only be solved by comprehensive reforms requiring equal sacrifices on the part of the insurance companies, the doctors, the hospitals, and the attorneys."[37]

Both a strength and a weakness of Keene's bill was its inclusion of a wide array of measures to remedy the insurance premium crisis. But it was, and still is, unclear which measures were necessary to freeze the rise in insurance premiums. Leaders of the tort reform movement champion MICRA as the model for medical malpractice reform, crediting it as the stabilizing force in the California malpractice insurance market.[38] The popularity of the cap is reflected in the fact that it has been embedded in the federal government's compensation fund for victims of September 11[39] and served as the staple amount for compensation payable under insurance for military personnel killed in action and public safety officers killed on duty.[40] However, detractors point to other legislative measures that could just as likely have slowed rate increases in California[41] and lament

---

http://www.micra.org/CAPPNews42605.pdf [hereinafter *Perspectives 1*] (the first installment of an interview run over three issues). On December 12, 1975, the cap went into effect and has remained unchanged. *See July 13 Hearing on A.B. 1380 Before the Cal. S. Comm. on the Judiciary*, *supra* note 33, at 1.

[35] *Perspectives*, *supra* note 34, at 3.

[36] *Id.* at 3; *see* JULIANNE D'ANGELO FELLMETH & THOMAS A. PAPAGEORGE, INITIAL REPORT MEDICAL BOARD OF CALIFORNIA ENFORCEMENT PROGRAM MONITOR 20 (Center for Public Interest Law University of San Diego School of Law 2004); Keene, *supra* note 23, at 30.

[37] Press Release, Office of Assemblyman Barry Keene (May 19, 1975), at 5.

[38] A 2003 Department of Health and Human Services report endorsed the success of MICRA, illustrating that since 1975, premiums in California have risen 167% while premiums in the rest of the country have increased 505%. OFFICE OF DISABILITY, AGING AND LONG-TERM CARE POLICY, U.S. DEP'T OF HEALTH AND HUMAN SERVICES, ADDRESSING THE NEW HEALTH CARE CRISIS: REFORMING THE MEDICAL LITIGATION SYSTEM TO IMPROVE THE QUALITY OF HEALTH CARE, at 24 (2003), *available at* http://aspe.hhs.gov/daltcp/reports/medliab.htm.

[39] *See* Sept. 11th Victim Compensation Fund of 2001, 67 Fed. Reg. 11, 239 (Mar. 13, 2002).

[40] CALIFORNIANS ALLIED FOR PATIENT PROTECTION, THE CALIFORNIA STORY: MICRA: A SUCCESSFUL MODEL FOR AFFORDABLE AND ACCESSIBLE HEALTH CARE 6 (2002), *available at* http://www.micra.org/CAStory.pdf.

[41] Twelve years after MICRA's passage, California premiums reached an all-time high, 190% higher than when the statutes were enacted. In another attempt to contain the premium rate hikes, California passed the Insurance Rate Reduction and Reform Act, known as Proposition 103. Among other things, Proposition 103 mandated 20% insurance rate reductions and granted the insurance commissioner the ability to reject or alter rate increase requests. Since the passage of Proposition 103 in 1988, malpractice premium rates in California have fallen below national levels. Gunnar, *supra* note 4, at 489–91. *See also The Medical Insurance Crisis: A Review of the Situation in Pennsylvania: Hearing Before*

that MICRA represents a marginalizing ambush of groups who feel the strongest negative impacts of non-indexed non-economic damages caps.[42] Without definitive support for its effectiveness and positive impact, MICRA remains open to criticism on substantive grounds.

III. WHOM DOES MICRA TRULY AFFECT?

MICRA's non-indexed non-economic damages cap has proven detrimental to some Californians injured by medical malpractice. Proponents of non-economic damages caps pay little attention to how caps affect access to the civil justice system or to how they may harm discrete groups of people.[43] Because of such caps, claimants who experience low economic loss but suffer high non-economic loss have claims that are no longer attractive to many lawyers, who will not incur the costs of litigation without the prospect of profiting from damage awards.[44] Moreover, the impact on select groups, namely minorities and women, has been gravely disproportionate.[45]

Lucinda Finley of the State University of New York at Buffalo examined how juries in medical malpractice tort suits allocate their awards between economic and non-economic damages.[46] She found that caps on non-economic damages can discriminate against women because women face unique injuries such as pregnancy loss, sexual or reproductive harm, and sexual assault that can lead to impaired fertility or sexual functioning, miscarriage, incontinence, trauma associated with sexual relationships, and scarring or disfigurement in sensitive, intimate areas.[47] Non-economic damages are the principle means by which juries can demonstrate their sense of the gravity of, and compensate women for, these unique harms.[48] These injuries and their consequences, however, do not significantly impact a woman's wage-earning capability, which is the ranking consideration in the tabulation of economic damages. Consequently, many damage awards to female claimants are based mostly on non-economic loss,

---

the Subcomm. On Oversight and Investigations of the H. Comm. On Energy and Commerce, 109th Cong. 130 (2003) (testimony of Harvey Rosenfield, President, Foundation for Consumer Taxpayer Rights). *But see* WILLIAM HAMM ET AL., MICRA, NOT PROPOSITION 103, ACCOUNTS FOR THE RELATIVELY LOW-GROWTH IN MEDICAL MALPRACTICE INSURANCE COSTS IN CALIFORNIA 2 (2005), *available at* http://www.micra.org/Proposition103Report.pdf.

[42] The MICRA cap would have to be raised to $916,025 to keep up with inflation since 1976. CONSUMER ATTORNEYS OF CALIFORNIA, CAOC INFORMATION BRIEF: "BEHIND THOSE MEDICAL MALPRACTICE RATES" 1 (2005), http://www.caoc.com/CA/index.cfm?event=showPage&pg=MICRABrief02-22-05.

[43] Lucinda Finley, *The Hidden Victims of Tort Reform: Women, Children, and the Elderly*, 53 EMORY L.J. 1263, 1264–65 (2004).

[44] *See id.* at 1265; *see also* PACE ET AL., *supra* note 16, at xxi.

[45] *See* CONSUMER ATTORNEYS OF CALIFORNIA, *supra* note 42; Finley, *supra* note 43, at 1266.

[46] Finley, *supra* note 43, at 1266.

[47] *Id.*

[48] *Id.* at 1266, 1281.

which means that these claimants lose a significant portion of their awards as a result of non-economic damage caps. For example, in a sample of 28 gynecological malpractice cases in California, 83% of damages awarded by juries were non-economic.[49] Those claimants experienced a 64% reduction in average recoveries as a result of the damage cap.[50] Because proponents of non-economic caps have succeeded in convincing legislators that non-economic damages are "parasitic,"[51] women who face harms that are unique to their sex and that are only compensated for by non-economic damages are subject to drastically reduced damage awards.

Just as the types of harms women experience expose them to disproportionately adverse outcomes with respect to jury awards, the tendencies of juries in "gender-neutral" medical malpractice cases also adversely affect women. In those cases, the damage packages awarded to women are composed of a higher percentage of non-economic damages than those awarded to men.[52] That fact exacerbates the harm caused to women by non-economic damage caps. Finley's study of 131 "gender-neutral adult plaintiff" California cases, with sixty-seven female plaintiffs and sixty-four male,[53] revealed that MICRA damage caps depressed women's awards even further below men's awards than they already tended to be.

In addition, common methods of calculating economic damages deny minorities and women the ability to obtain gender- and race-neutral jury awards. The work-life expectancy of the claimant and the average wage the claimant would have earned absent the malpractice are the factors used to tabulate economic damages.[54] However, race- and gender-specific data frequently used in tabulating economic damages projects that white men are worth more economically than women or minorities.[55] For instance, race- and gender-specific work-life tables indicate that racial minorities and women will spend fewer years in the labor force.[56] Higher rates of incarceration and a shorter average life span for minorities and absence from the work force due to childbearing for women all amount to shorter work-life expectancies for individuals in these categories, which amounts to lower economic damages for the same harms.[57] According to the U.S. Bureau of Labor Statistics, the work-life expectancy for a white man injured at age thirty was estimated to be 4.7 years longer than that of a mi-

---

[49] *Id.* at 1296.

[50] *Id.*

[51] Martha Chamallas, *Vanished from the First Year: Lost Torts and Deep Structures in Tort Law*, *in* LEGAL CANONS 114 (J. M. Balkin & Sanford Levinson eds., 2000).

[52] Finley, *supra* note 43, at 1284.

[53] *Id.* at 1284–85 ("Before applying the MICRA cap, women's average jury awards were 52% of men's average awards. After the MICRA reduction, the women on average recovered only 45% of men's average recoveries.").

[54] Chamallas, *supra* note 51, at 109.

[55] *Id.* at 112.

[56] *Id.*

[57] *Id.* at 110.

nority man, 8.7 years longer than that of a white woman, and 9.2 years longer than that of a minority woman.[58] Moreover, disparities in criminal sentencing and a lack of job opportunities in minority communities further corrupt the objectivity of such tables.[59]

In sum, because of unique harms that affect women and minorities, jury tendencies, and reliance upon flawed economic tables, non-economic damages caps silently plague these groups.

IV. THE EMPTY SEARCH: WHY $250,000

Although the $250,000 non-economic damages cap reigns as the popular reform solution for many states and pending federal reform, the question still remains: Why $250,000?[60] Since those presently carrying the medical malpractice tort reform banner have failed to provide an explanation as to why $250,000 is the optimal cap, an examination of the origins of that number may shed light on this question.

A. Legislative History

Legislative records charting California's early tort reform bill, eventually known as MICRA,[61] revealed that the highly publicized cap on non-economic damages was not even a part of the initial bill—it was a last-minute addition.[62] As originally introduced, the bill limited compensation for certain non-economic losses, including pain and suffering, to $800 a month and provided that a claimant would not be entitled to non-economic losses if his earnings exceeded $1,500 a month.[63] At the request of the Assembly Judiciary Committee, the monthly restriction on non-economic losses was deleted,[64] prompting the California Physicians Crisis Committee to call the Judiciary Committee "no more than a lawyers' lobby."[65] Assembly Bill 1xx was passed by the Assembly on June 20, 1975, without any limit on the amount of damages that an injured party

---

[58] *Id.* at 110.

[59] *Id.* at 112.

[60] *See supra* note 2

[61] Files relating to A.B. 1xx from the Judiciary Committee, Governor Brown, and the bill's author (retained by the California Assembly) constitute the 1500+ pages of legislative records searched. The Senate Insurance and Financial Institutions Committee files from the California State Archives were unobtainable.

[62] *See infra* notes 63–69 and accompanying text.

[63] S. COMM. ON INSURANCE AND FINANCIAL INSTITUTIONS, ANALYSIS OF ASSEMB. B. NO. 1 AS AMENDED JUNE 27 1 (1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files).

[64] *Id.* ("At the request of the Assembly Judiciary Committee, this restriction on non-economic losses was deleted. Therefore A.B. 1, as passed by the Assembly, did not limit the amount of damages that an injured party may recover.").

[65] Press Release, Cal. Physicians Crisis Comm. (June 26, 1975) (on file with Herbert M. Baus).

could recover.[66] As of June 27, 1975, however, the Senate Insurance and Financial Institutions Committee had adopted significant amendments to the bill, which included the provision limiting non-economic damages to $250,000.[67] Assembly members did not abide by the Assembly Speaker's request to "non-concur" with the Senate version of the bill, which would have sent it back to the Assembly Judiciary Committee, where it had previously been gutted of its damages provision.[68] Instead, Assembly members accepted the Senate version of the bill, which included the non-economic damages cap.[69]

B. Casualties of the Special Session

Although Assembly Bill 1xx stole the spotlight during the special session, it did not stand alone as the only attempt at capping damages in medical malpractice tort reform. In addition to many amendments to Assembly Bill 1xx that were abandoned in the special session, a series of medical malpractice tort reform bills were also laid to rest. Senator Omer L. Rains's Senate Bill 7xx contained a measure that was identical to the original version of Assembly Bill 1xx.[70] It proposed an $800 to $1,500 per month limitation for pain and suffering and no limit on loss of earnings or necessary medical treatment costs.[71] A different bill, Senate Bill 1, allowed for the recovery of up to $250,000 in non-economic damages when the injury consisted of the loss of a body part, serious impairment of a bodily function, or serious and permanent disfigurement, whether as a direct result of negligent medical intervention or negligent intervention combined with the original injury necessitating medical care.[72]

Not all of the bills proposed during the special session resembled Assembly Bill 1xx. Senate Bill 13, presented by Senator Alfred Song, sought

---

[66] *See supra* note 30 and accompanying text.

[67] *Aug. 6 Hearing on A.B. 1 Before the Cal. S. Comm. on Insurance and Financial Institutions*, *supra* note 22, at 3("AB 1xx does not place limitations on the amount of damages that a claimant may recover for necessary medical treatment, loss of earnings, and expenses for obtaining a substitute to perform necessary services that would have been performed by the patient. 'Pain and suffering' would be limited to $250,000."). By August 11, the $250,000 limitation had been re-inserted onto the bill. S. COMM. ON INSURANCE AND FINANCIAL INSTITUTIONS, ANALYSIS OF ASSEMB. B. NO. 1 AS AMENDED AUG. 11 1 (1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files).

[68] *Perspectives: An Interview With MICRA Author Barry Keene*, CAPP NEWS (Californians Allied for Patient Protection, Sacramento, Cal.), Aug. 8, 2005, at 1, *available at* http://www.micra.org/augustnews.pdf [hereinafter *Perspectives 3*] (the last installment of an interview run over three issues).

[69] *Federal Update*, *supra* note 11, at 1.

[70] *Medical Injury Compensation Reform Act of 1975: Hearing on A.B. 1 as Amended in the Second Extraordinary Sess. Before the Cal. S. Comm. on Insurance and Financial Institutions.*, 1975-1976 Leg. 2 (June 26, 1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files).

[71] *Id.*

[72] S. COMM. ON INSURANCE AND FINANCIAL INSTITUTIONS, 1975-1976 Leg., MALPRACTICE BILLS INTRODUCED JUNE 2 (Cal. 1975).

to limit total damages to $500,000.[73] After going to the Senate Judiciary Committee, it failed to emerge as a front-runner in reform. In addition, Senator Alan Robbins's Senate Bill 521 sought to create a board that would be the sole recourse for patients seeking compensation for medical injuries and awards and to limit damages to loss of wages, medical expenses, and loss of bodily function.[74] This bill also died during the special session,[75] as did Senator Dennis E. Carpenter's Senate Bill 397, which provided a $500,000 maximum for the death of a minor or adult without dependents.[76]

In addition to those failed bills, two significant amendments to Assembly Bill 1xx also died during the special session. The first, proposed on August 11, 1975, sought to limit the provision of non-economic damages to ten percent of the economic losses.[77] The second, the extraordinary hardship exception, originally slated to coincide with the initial bill proposal, did not gain the support necessary for inclusion in MICRA.[78] For no apparent reason, these two amendments, which would have moderated the harsh effects of a pure damages cap, were not included in the final bill. The question remains: Why did these ideas fail while the hard $250,000 cap succeeded?

C. Discussion Inside the Assembly

In Chairman Keene's statement before the Senate Insurance and Financial Institutions Committee, he did not mention the $250,000 cap. In fact, in his synopsis of his "lower-cost, fairer resolution" to the insurance premium crisis, he mentioned only the proposal to allow "periodic payments to spread compensation amounts over a period of time, to preclude costly windfalls to heirs when an injured party dies before completing his actuarially expected life."[79]

On September 3, 1975, Alister McAlister, chairman of the Assembly Committee on Finance, Insurance, and Commerce, sent Chairman Keene suggested amendments to his bill. McAlister proposed the creation of a joint

---

[73] ASSEMB. COMM. ON HEALTH, 2d Spec. Sess., MEDICAL MALPRACTICE LEGISLATIVE REFERENCE CHART: LEGISLATION PRESENTLY BEFORE THE CAL. S. ON JULY 15 (Cal. 1975).

[74] ASSEMB. COMM. ON HEALTH, 2d Spec. Sess., MEDICAL MALPRACTICE LEGISLATIVE REFERENCE CHART: LEGISLATION PRESENTLY BEFORE THE CAL. S ON MAY 8 3 (Cal. 1975).

[75] *Id.*

[76] *Id.*

[77] CAL. ASSEMB., AMENDMENTS TO ASSEMB. B. NO. 1 AS AMENDED IN SENATE AUG. 11, 1975, 1975-1976 Leg., 2d Spec. Sess. (1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files).

[78] The extraordinary hardship exception would have applied when "non-economic loss substantially outweighs actual economic loss." Req. #12856, Interoffice Memorandum, Office of Assemblyman Barry Keene (date unknown) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files).

[79] *Aug. 6 Hearing on A.B. 1 Before the Cal. S. Comm. on Insurance and Financial Institutions*, *supra* note 22, at 4.

underwriting association and the establishment of three actuaries to assist California's Insurance Commissioner; he also did not mention the cap on non-economic damages.[80]

As the bill progressed through the State Senate, Senate Judiciary Committee consultant and later legislative counsel Bion Gregory suggested indexing the non-economic damages cap.[81] However, this suggestion was disregarded because the plaintiff lawyers' lobby would not support the idea.[82] Ironically, some of the representatives of the trial bar thought indexing the cap would improve the bill's overall chance for passage and increase the likelihood of the Governor signing it.[83] As a result, they withheld their support of the indexed cap in order to try to kill the bill altogether.[84] But their move backfired—the final bill included the non-indexed cap that adversely affects a number of their potential clients today. Still, their tactical error gives no indication as to why $250,000 was chosen as the magic number.

Nor did such an indication emerge from the records. There was, however, an actuarial appraisal of the potential effect of the $250,000 non-economic damages cap, which indicated the following:

> A limitation of damages for non-economic losses to $250,000 will probably result in approximately a 5% reduction in total premium. This item has been difficult to measure since the awards are not normally categorized by economic and non-economic losses. It should have a tendency to reduce some of the very large emotional awards.[85]

There was also an Insurance Commissioner's estimate, which projected that the damages cap would yield estimated premium savings of six to ten percent.[86] However, the Insurance Commissioner offered little explanation of those projected savings. Nevertheless, they were also employed by the Department of Consumer Affairs in its analysis of the bill.[87] In that analysis the Department acknowledged that the insurance industry had

---

[80] Memorandum from Carlyle R. Brakensiek, Chief Counsel, Cal. Assemb. Comm. on Finance, Insurance, and Commerce to Barry Keene, Chairman, Cal. Assemb. Comm. on Health (Sept. 3, 1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files).

[81] *Perspectives 3*, *supra* note 68, at 3.

[82] *Id.*

[83] *Id.*

[84] *Id.*

[85] COATES AND CRAWFORD, INC., ACTUARIAL APPRAISAL OF ASSEMBLY BILL 1 AS AMENDED ON JUNE 27, 1975 (1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files).

[86] THE INSURANCE COMMISSIONER, A.B. 1XX PREMIUM SAVINGS ESTIMATED BY THE INSURANCE COMMISSIONER (1975) (on file with the Cal. Assemb. B. Author A.B. 1xx 1975 files).

[87] *See* DEP'T OF CONSUMER AFFAIRS, ENROLLED BILL REPORT ON AB 1xx, at 15 (Sept. 22, 1975) (on file with Cal. Governor A.B. 1xx 1975 files).

failed to provide an estimate of how much the bill would reduce premiums.[88] Moreover, it even noted that it had relied upon the Insurance Commissioner's estimates.[89] However, no hard analysis beyond the Insurance Commissioner's projections indicated that a non-economic damages cap would translate into premium savings.

The lack of any clear reason for the value of the $250,000 cap is indicated by an exchange between Senator Newton R. Russell and Travelers Insurance. Senator Russell and Terry Miller, Consultant to the Senate Committee on Insurance and Financial Institutions, requested estimates from California's insurance companies of the effects the proposed legislation would have on 1976 medical malpractice insurance premiums.[90] Travelers Insurance responded by proposing that more stringent measures be incorporated into the bill.[91] However, far from providing rate estimates that would substantiate even their new proposal, Travelers stated that they could not be "responsive to [the committee's] requests for pricing information" because they did "not have the necessary data to make intelligent estimates" about the effect the bill would have on rates.[92] The insurance industry thus had no better estimate of the bill's potential effects than did the Department of Consumer Affairs.

Despite the espoused inability of insurers themselves to make such estimates, Chairman Keene urged Governor Brown to sign the bill because he projected that its provisions would reduce premiums by 18% to 24%.[93] The basis of those projections was absent from the letter.[94] Further, while Keene also failed to proffer an explanation as to why a $250,000 cap would be effective, he did state during his opening remarks to the Joint Health-Judiciary Committee Meeting on Medical Malpractice that "[trial lawyers] do not like limitations on pain and suffering even though there may be sound reasons of social policy for imposing them."[95] His statement suggests that he hoped to convince his peers that the cap might be worthwhile to adopt only because there might, theoretically, be sound policy

---

[88] *Id.*

[89] *Id.*

[90] *See* Letter from Roger J. Fisher, Second Vice President, The Travelers, to Newton R. Russell, Cal. S. (July 29, 1975) (on file with the Cal. Assemb. B. Author A.B. 1xx 1975 files) [hereinafter Letter to Newton R. Russell]; Letter from Roger J. Fisher, Second Vice President, The Travelers, to Terry J. Miller, Consultant, S. Comm. on Insurance and Financial Institutions (July 29, 1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files) [hereinafter Letter to Terry J. Miller].

[91] Letter to Newton R. Russell, *supra* note 90.

[92] *See* Letter to Terry J. Miller, *supra* note 90.

[93] Letter from Barry Keene, Chairman, Cal. Assemb. Comm. on Health, to Edmund G. Brown, Jr., Governor, State of California (Sept. 12, 1975) (on file with Cal. Governor 1975 A.B. 1xx files).

[94] *See id.*

[95] Barry Keene, Chairman, Cal. Assemb. Comm. on Health, Opening Remarks to the Cal. Assemb. Joint Comm. on Health & Judiciary Comm. Hearing on Medical Malpractice 2 (June 5, 1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files).

reasons to do so. However, Keene declined to specify precisely what those reasons were.

In exchanges since MICRA's adoption, Keene and other supporters of the non-economic damages cap have still failed to provide empirical support for the $250,000 limit. The bill author himself, when pressed about the number, said that the "selection of a quarter million dollars was subjective."[96] He went on to explain:

> The theory was that you could never really and adequately compensate for pain and suffering, no matter how much money you provided. Money just doesn't do it. But $250,000 (in addition to meeting the medical and other needs of the patient), properly invested to the extent that it elevated the quality of life over and above the post-injury status, was thought to be enough to do that job.[97]

Finally, a Department of Consumer Affairs report discussed whether $250,000 of non-economic damages would be an equitable amount for claimants.[98] In its analysis, the Department stated, "[t]he limit of $250,000 may be acceptable for consumers" because the limit "would not change the amount awarded to most claimants."[99] However, this analysis did not recognize the concentrated effect that this amount would have on those specific groups of claimants who would be especially affected, such as women and minorities.[100] Moreover, even the bill's author has explained that the damage cap constituted a "moral dilemma" between "denying patients access to reasonably-priced health care in California" or "undercompensating" some injured patients.[101]

That idea is confirmed by a suggested veto message that was sealed in Governor Brown's files. The message noted that the bill "did not deal effectively with the problem of the negligent practitioner, without whom there would be no malpractice crisis. . . . [The bill] fail[ed] to provide added disciplinary powers which are needed to protect the consumers of California from malpractice."[102] The suggested remarks may reflect Governor Brown's concern that the bill's primary harmful effects would be felt by patients first.

---

[96] E-mail from Barry Keene, Author of MICRA (A.B. 1xx), to Amanda Edwards, Author (Aug. 22, 2005, 16:43:23 EST) (on file with author).

[97] *Id.*

[98] DEP'T OF CONSUMER AFFAIRS, *supra* note 87, at 15

[99] *Id.*

[100] *Supra* notes 43–58 and accompanying text.

[101] Keene, *supra* note 23, at 31.

[102] Edmund G. Brown, Jr., Governor, State of Cal., Suggested Veto Message to the Members of the Cal. Assemb.: A.B. 1xx (on file with Cal. Governor A.B. 1xx 1975 files).

Governor Brown's predictions were correct. In fact, despite the controversy, lack of precedent, and uncertainty about the bill's efficacy,[103] health care providers,[104] officials throughout the state,[105] and insurance companies[106] pledged support for its passage. Still, the bill's supporters were concerned about the bill's fate. Consequently, they made several procedural maneuvers to avoid the Assembly Judiciary Committee, whose Chairman, they feared, might "amend [A.B. 1xx] into ineffectiveness."[107] Health Committee Chairman Keene, coordinating with the Senate Rules Committee, avoided referring the bill to the Judiciary Committee, where it would have been certain to face an uphill battle due to its unpopularity with committee members.[108] Instead, the bill was sent to the Senate Financial Institutions and Insurance Committee, which molded it into a form that included the $250,000 cap on non-economic damages for which the legislation has received so much attention over the past 30 years.[109] Ultimately, the bill passed in the California legislature because of procedural maneuvering rather than the logical or empirical force of its provisions.[110]

The non-economic damages cap thus quietly made its way into California's MICRA, hardly discussed by legislators, constituents, or even lob-

---

[103] Letter from Elmer Low, President, Cal. Trial Lawyers Ass'n., to John Miller, Chairman, Cal. Assemb. Comm. on the Judiciary (June 13, 1975) (on file with Cal. Comm. on the Judiciary A.B. 1xx 1975 files) ("Our committee believes this bill is unfair to injured patients and seriously hinders the fair administration of justice.").

[104] William M. Whelan, Executive Vice President, Cal. Hosp. Ass'n, Statement of Cal. Hosp. Ass'n (June 5, 1975) (on file with Cal. Comm. on the Judiciary A.B. 1xx 1975 files) ("I am here to express the California Hospital Association's general support for AB1.").

[105] *See* Letter from Marvin Freedman, Chief Admin. Officer, County of Los Angeles, to Barry Keene, Chairman, Cal. Assemb. Comm. on Health (July 28, 1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files).

[106] *See* Letter from Larry Harvey, Senior Vice President, Signal Ins. Co., to Members of the Assemb. Judiciary Comm. (June 2, 1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files) ("We can state unequivocally that we will as soon as this package is enacted, start to lower our rates."). However, in a June 11, 1975, letter to the Judiciary Committee, Signal Insurance Company stated that the sliding scale contingent fee, the collateral source component, the structured settlement payment provision, and the lack of a res ipsa component would not amount to sufficient enough changes to lower the cost of malpractice insurance but potentially could cause increases. Letter from Larry Harvey, Senior Vice President, Signal Ins. Co., to Members of the Assemb. Judiciary Comm. (June 11, 1975) (on file with Cal. Assemb. B. Author A.B. 1xx 1975 files).

[107] *Perspectives: An Interview With MICRA Author Barry Keene*, CAPP NEWS (Californians Allied for Patient Protection, Sacramento, Cal.), May. 26, 2005, at 1, *available at* http://www.micra.org/CAPPNews52605.pdf [hereinafter *Perspectives 2*] (the second installment of an interview run over three issues).

[108] E-mail from Barry Keene, Bill Author of MICRA (A.B. 1xx), to Author (Aug. 23, 2005, 11:48:10 EST) (on file with author).

[109] *Perspectives 2*, *supra* note 107, at 2.

[110] *See* Letter from Richard F. Mills, Law Corporation, to John J. Miller, Member, Cal. Assemb. (June 9, 1975) (on file with author) ("In the flurry of legislation presently pending, many people have lost sight of the fact that the so-called crisis was caused by a unilateral increase of malpractice insurance by the insurance industry, and as yet, no one has really *found out* why."); *see also* Letter from William J. O'Connell, Harrison & Watson Attorneys at Law, to John J. Miller, Member, Cal. Assemb. (July 2, 1975) (on file with author); *Perspectives 2*, *supra* note 107, at 2.

bying groups. The historical records reveal the behind-the-scenes priorities of lawyers' groups, legislators, health care providers, and active constituents; for all of them, the non-economic damages cap was hardly on the radar. The cap's low-profile enactment in 1975 seems surprising considering how it has become the centerpiece of current reform efforts.

V. AN ALTERNATIVE TO THE $250,000 CAP

Since the $250,000 cap on non-economic damages is not supported by a strong evidentiary foundation, alternative solutions should be examined for inclusion in federal legislation. A simple solution already exists to remedy the blockbuster cases that medical malpractice reform supporters continue to caricature and thrust into the spotlight. Federal Rule of Civil Procedure 59,[111] which governs motions for partial new trials, provides a solution without the heartache that the cap on non-economic damages presently imposes. A judge may engage the process of *nisi remittitur damna*,[112] by which a judge may reduce the jury award. In the *remittitur* process, the judge conditions the grant of a new trial on the plaintiff's refusal to consent to a reduction of damages.[113] Plaintiffs may elect to reduce damages through *remittitur* instead of undergoing a new trial, during which they may incur significant expenses and after which they may be left entirely without compensation.

The media rarely reports the regular practice of award reductions; although fairly common, these reductions are not sexy enough to garner public attention. Instead, huge pre-reduction jury awards are used as fodder in the depiction of a tort system gone mad. For example, the well-known McDonald's coffee case has long been used to characterize the American tort system as a three-ring circus.[114] In that case, an eighty-one-year-old woman was severely burned by hot coffee after McDonald's failed to explicitly warn customers of the dangers of its coffee.[115] After the trial judge reduced the $2.7 million punitive damage award to $480,000, the $2.86 million total award was reduced to $640,000.[116] However, those details were left lying on the editing room floor when the case was publicized nationwide.[117] One examination of 198 jury awards of $1 million or more revealed

---

[111] Fed. R. Civ. P. 59.

[112] This is the process by which parties negotiate for reduced jury awards after the trial's conclusion. *See id.*; *see also* Suja A. Thomas, *Re-examining the Constitutionality of Remittitur Under the Seventh Amendment*, 64 OHIO ST. L.J. 731 (2003).

[113] 65 A.L.R.2d. 1331 ("When jury is found to have erred in awarding excessive damages by reason of passion, prejudice, or sympathy, court may require winner of excessive verdict to chose between new trial and acceptance of reduced award.").

[114] Liebeck v. McDonald's Restaurants, P.T.S., Inc., No. CV-93-02419, 1995 WL 360309 (D.N.M. Aug. 18, 1994).

[115] *Id.*

[116] *Id.*

[117] *See* Neil Vidmar et al., *Jury Awards for Medical Malpractice and Post-Verdict Adjustments of Those Awards*, 48 DEPAUL L. REV. 265, 266 (1998).

that plaintiffs received the original jury award in just over a quarter of cases.[118] On average, the plaintiffs' final disbursements were fifty-seven percent lower than their original awards.[119] Another study of jury awards in California revealed that one in four jury verdicts was subject to post-verdict adjustment.[120] Those facts indicate that policy-makers are avoiding the real issue involved when they cite large jury awards as the cause of high premiums. Because judges readily exercise tools like *remittitur* and Rule 59 to mitigate outlier jury verdicts, those verdicts are not a significant cause of high premiums.[121]

The process of finding a solution to insurance premium rate instability should entail determining how to compensate currently injured patients while ensuring that adequate funds remain available for future claimants. However, the most important piece of this puzzle is to ensure that recoveries do not become too small to serve their purpose. As the Center for Public Interest Law has noted, "[A] person whose life is changed forever because of a provider's misconduct deserves fair compensation—and the current cap has been ravaged by inflation over the past 24 years, rendering it wholly inadequate."[122] Medical malpractice tort reform advocates often paint a portrait of claimants hitting the "jackpot" with meritless suits. However, given the odds of reduced awards and the severe price paid by valid claimants, signing up for this lottery hardly seems like a ticket to the good life.

VI. CONCLUSION

Despite the negative impact of the $250,000 non-economic damages cap on minorities and women, supporters in state and federal government seek to impose the cap on the entire nation through federal legislation. But for what reason? Instead of revealing sources of substantive support for the $250,000 cap, an examination of the origin of the cap only reveals a lack of legislative deliberation on the issue.

Determining non-economic damages in a medical malpractice case entails subjectively deriving just compensation based on less-than-concrete notions of fairness and magnitude of pain and suffering. Injuries compensated by non-economic damages, however, are very palpable. Whether it is lifelong trauma caused by a miscarriage or the loss of a body part, each injury is unique and affects not only a victim's body but also the intangi-

---

[118] *See* Ivy E. Broder, *Characteristics of Million Dollar Awards: Jury Verdicts and Final Disbursements*, 11 JUST. SYS. J. 349, 349–50, 353 (1986).

[119] *See* Broder, *supra* note 118, at 350, 353; Vidmar et al., *supra* note 117, at 279.

[120] *See* Vidmar et al., *supra* note 117, at 294 (citing a 1998 study).

[121] Vidmar et al., *supra* note 117, at 298.

[122] *Hearing on A.B. 1380 Before the Assemb. Comm. on the Judiciary*, *supra* note 20, at 10.

ble elements that make a person whole. In the case of non-economic damages, one size does not fit all.

Although the rationale underlying the $250,000 non-economic damages cap in California may appear flimsy because of scant records showing in-depth deliberation over the value, it is possible that particular number is far from arbitrary. Still, it seems unwise for proponents to fight to preserve the non-indexed $250,000 cap while the reason that this number was selected remains a mystery—especially since its effects on certain groups of Californians are stark.

*—Amanda Edwards**

* B.A., Emory University, 2004; J.D. Candidate, Harvard Law School, Class of 2007.