HARVARD JOURNAL
on
LEGISLATION

VOLUME 31, NUMBER 1 WINTER 1994

ARTICLES

NOTES

HARVARD JOURNAL ON LEGISLATION
Harvard Law School
Cambridge, MA 02138
Phone: (617) 495-4400

The HARVARD JOURNAL ON LEGISLATION (ISSN 0017-808x) is published twice each year (winter and summer) by Harvard Law School students. Third-class postage paid at Boston, Massachusetts, and at additional mailing offices.

Publication Policy

The JOURNAL specializes in the analysis of legislation and the legislative process. It focuses on legislative reform and on organizational and procedural factors that affect the efficiency and effectiveness of legislative decision-making. The JOURNAL is especially interested in publishing articles that examine public policy problems of nationwide significance and propose legislation to resolve them.

The JOURNAL welcomes unsolicited manuscripts. Authors interested in being published in the JOURNAL should send their manuscripts and all related correspondence to: Articles Editor, HARVARD JOURNAL ON LEGISLATION, at the above address. Manuscripts should be typed, double-spaced, and submitted in duplicate.

Views expressed in the JOURNAL are those of the authors and do not necessarily reflect the views of the JOURNAL's officers, editors, or other members.

Subscriptions

The JOURNAL's subscription rate for the 1994 volume (Volume 31) is $24.00 domestic rate; $28.00 foreign surface rate ($30.00 in Canada); $42.00 foreign air mail. All correspondence concerning subscriptions should be addressed to the Subscriptions Manager, Publications Center, Harvard Law School, Cambridge, MA 02138.

Unless otherwise requested by the subscriber, the JOURNAL is mailed at bulk rate. For some overseas subscribers, delivery via air mail may be more reliable than standard surface rate service. Claims for missing issues of Volume 31 will be honored only until January 1995.

Individual Issues

Copies of individual issues from the current volume can be obtained directly from the JOURNAL for $12.00 each (plus $2.00 per copy for foreign mailing).

Copies of issues from all past volumes of the JOURNAL are available from Fred B. Rothman & Co., 10368 West Centennial Road, Littleton, CO 80123, (303) 979-5657. Contact Fred B. Rothman & Co. for information on prices for bound volumes.

When purchasing copies of issues from either the JOURNAL or Fred B. Rothman & Co., please send a check for the purchase price when you place your order.

Indexes

Starting with Volume 23, the JOURNAL discontinued publication of volume-by-volume and five-year cumulative indexes. The JOURNAL is indexed in the *Index to Legal Periodicals.*

ARTICLE

AND WHERE YOU GO I'LL FOLLOW: THE CONSTITUTIONALITY OF ANTISTALKING LAWS AND PROPOSED MODEL LEGISLATION

ROBERT P. FAULKNER
DOUGLAS H. HSIAO*

In recent days, stalking has become a tremendous problem plaguing American society. But, while states have responded to this growing crisis by passing statutes to punish offenders, these legislative solutions are often ineffective. Many statutes have been deemed unconstitutional by state courts, while others are too narrowly drawn to encompass much dangerous stalking behavior.

In this Article, the authors examine the problems with current antistalking laws and offers some practical solutions. Part I discusses and disects states' present antistalking statutes, focussing on the language common to many of these laws as well as their unique exceptions. Part II focuses on the general constitutional infirmities plaguing almost all currently enacted antistalking statutes. Part III examines the NCJA model antistalking statutes, noting its constitutional deficiencies, and in Part IV the authors draft their own model antistalking legislation which they argue will both pass constitutional muster and sweep more broadly to encompass many dangerous stalking behaviors.

In the wake of the stalking and brutal murder of "My Sister Sam" co-star Rebecca Schaeffer,[1] and in response to the perception that existing civil and criminal remedies are ineffective in

---

* Mr. Faulkner is an attorney at the Washington, D.C. branch of Jenner & Block, where he practices civil litigation focusing on antitrust and First Amendment issues. Mr. Hsiao, also an attorney with Jenner & Block, authored Subsection II(D)(2), *infra* notes 175–197 and accompanying text, relating to pretrial detention. The authors would like to thank Sheri L. Metcalfe, Mitchell D. Klaich, Teri L. Slabinski, and Tara M. Ghannam, whose collective efforts immeasurably improved this Article. This work derives from research originally commissioned by Nkechi Taifa, Legislative Counsel, American Civil Liberties Union. The conclusions expressed herein are the authors' alone and should not be taken to reflect the position of the ACLU.

[1] *See* Matthew J. Gilligan, Note, *Stalking the Stalker: Developing New Laws to Thwart Those Who Terrorize Others*, 27 GA. L. REV. 285 (1992) (attributing explosion of antistalking legislation to the murder of the popular television actress); Robert A. Guy, Jr., Note, *The Nature and Constitutionality of Stalking Laws*, 46 VAND. L. REV. 991 (1993) (same); Christina Perez, Note, *Stalking: When Does Obsession Become a Crime?*, 20 AM. J. CRIM. L. 263 (1993) (same). Other famous cases include the stalking of entertainment personalities David Letterman, Johnny Carson, Olivia Newton-John, and Michael J. Fox, *see* Kathleen G. McAnaney et al., Note, *From Imprudence to Crime: Anti-Stalking Laws*, 68 NOTRE DAME L. REV. 819, 819-20 (1993), as well as novelist Stephen King. *See* 138 CONG. REC. S9527 (July 1, 1992) (statement of Sen. Cohen).

dealing with repeated threats and threatening conduct,[2] at least forty-three states have now passed "antistalking" statutes.[3] These statutes, which generally criminalize the repeated following or harassing of another person when combined with a credible threat of serious violence, are intended to close a gap in the criminal laws and to empower law enforcement officials to prevent threatening conduct that too often leads to actual violence.[4]

On the other hand, the desire to prohibit truly life-threatening behavior should not lead to the adoption of sweeping and indefinite legislation that criminalizes conduct protected by the Constitution[5] or that vests too much discretion in police officers to enforce vague criteria.[6]

Some of these statutes, which rely on ambiguous terms such as "emotional distress," "legitimate purpose," "annoy," and "constitutionally protected activity," have already been used as a weapon against political protesters and others engaging in conduct that appears on its face to be lawful and legitimate.[7] Indeed, at least seven trial-level courts have invalidated antistalking laws brought before them,[8] one remarking that the statute in question

---

[2] *See infra* notes 27–30 and accompanying text. For extensive discussions of the inadequacies of existing legal remedies, see Gilligan, *supra* note 1, at 288–99; Guy, *supra* note 1, at 997–1000; Silvija A. Strikis, *Stopping Stalking*, ____ GEO. L.J. ____ (forthcoming 1993).

[3] ALA. CODE §§ 13A-6-90–94 (1992); ALASKA STAT. § 11.41.270 (1993); ARIZ. REV. STAT. ANN. § 13A-2921 (1992); ARK. CODE ANN. § 5-71-229 (Michie 1993); CAL. PEN. CODE § 646.9 (Deering 1993); COLO. REV. STAT. § 18-9-111 (1992); CONN. GEN. STAT. §§ 53a-182b, -183 (1992); DEL. CODE ANN. tit. 11.1312A (1992); 1992 D.C. Stat. 269u; FLA. STAT. ch. 784.048 (1992); GA. CODE ANN. § 17-6-1 (Harrison 1993); HAW. REV. STAT. §§ 604-10.5, 711-1106.5 (1992); IDAHO CODE § 18-7905 (1993); ILL. REV. STAT. ch. 38, para. 12-7.3 (1992); IOWA CODE §§ 708.7, .11 (1992); KY. REV. STAT. ANN. §§ 508.130, .140, .150 (Baldwin 1992), 403.715 (Michie 1992); LA. REV. STAT. ANN. § 14.40.2 (West 1992); 1993 Md. Laws 205–206; MASS. ANN. LAWS ch. 265, § 43 (Law. Co-op 1993); MICH. COMP. LAWS §§ 750.411h–.411i (1992); 1992 Miss. Laws 532; 1993 Mont. Laws 292; NEB. REV. STAT. §§ 28-311.03–.04; 1993 Nev. Stat. 233; 1993 N.H. Laws 173; N.J. REV. STAT. § 2C:12-10 (1992); N.M. STAT. ANN. §§ 30-3A-1 to -4 (Michie 1993); N.Y. PEN. LAW §§ 240.25–.26 (Consol. 1993); N.C. GEN STAT. § 14-277.3 (1992); 1993 N.D. Laws 120; OHIO REV. CODE ANN. §§ 2903.211, .214 (Baldwin 1993); OKLA. STAT. tit. 21, § 1173 (1992); 1993 Pa. Laws 28; R.I. GEN. LAWS §§ 11-59-1 to -3 (1992); 1993 S.C. Acts 417; S.D. CODIFIED LAWS ANN. § 22-19A (1993); TENN. CODE ANN. § 39-17-315 (1992); 1993 Tex. Sess. Laws Serv. 10 (Vernon); UTAH CODE ANN. § 76-5-106.5 (1993); VA. CODE ANN. § 18.2-60.3 (Michie 1993); WASH. REV. CODE ANN. § 9A.46.110 (West 1993); W. VA. CODE § 61-2-9a (1993); WIS. STAT. §§ 947.012, -.013 (1992); WYO. STAT. §§ 6-2-506, -509 (1993).

[4] *See* 138 CONG. REC. S9527 (July 1, 1992) (statement of Sen. Cohen) ("Only recently have the States begun to enact legislation that gives law enforcement officials the power to act against stalkers before they reach their prey.").

[5] *See infra* notes 75–103 and accompanying text.

[6] *See infra* notes 104–131 and accompanying text.

[7] *See infra* notes 55–63 and accompanying text.

[8] *See* United States v. Smith, No. M-6400-93 (D.C. Super. Ct. Sept. 9, 1993); State

would conceivably make it "a crime for a person in Saltville to use voodoo dolls to place a curse of death on someone in Virginia Beach."[9]

Protest groups ranging from right-to-life volunteers to organized labor have expressed concern that these statutes could be used to prevent legitimate demonstrations and political activity.[10] Conversely, some complain that these laws are too narrow to be effective.[11] Herein lies the problem which this Article attempts to solve: how can legislators draft a statute that criminalizes "stalking," a concept not susceptible to precise definition, without trampling on constitutional values? In our system of government, the answer must be that some arguably dangerous behavior should be excluded, but only to the extent absolutely necessary. The key is careful drafting based on a clear understanding of the constitutional pitfalls.

---

v. Tremmel, No. 93-02769 (Fla. Cir. Ct. June 9, 1993); State v. Wallace, No. 93-087 CF (Fla. Cir. Ct. May 19, 1993); State v. Kahles, No. 92-022819MM10A (Fla. Cir. Ct., 17th Cir. Mar. 10, 1993); Commonwealth v. Camper, No. 93-2876 (Va. Gen. Dist. Ct. Apr. 9, 1993).

The press has reported that a fourth Florida court and an Illinois court have held their respective state antistalking laws unconstitutional. *See, e.g.*, Jill J. Spitz, *Decision Delivers Blow to Florida's Anti-Stalking Law*, ORLANDO SENTINEL, Aug. 3, 1993, at 1; *Judge Rules Stalking Law Unconstitutional*, UPI, Aug. 26, 1993, *available in* LEXIS, Nexis Library, Wires File (court ruled that Illinois stalking law in effect until August 20, 1993 was unconstitutional under the state and federal constitutions).

Another Illinois court has held that the "no bail" provision of the state's antistalking law violated the state constitution. *See* David Bailey, *Circuit Judge Finds No-Bail Provision Of Stalking Law To Be Unconstitutional*, CHI. DAILY L. BULL., Feb. 5, 1993, at 1. As set forth *infra*, notes 192–196 and accompanying text, Illinois has amended the statute to improve its no-bail provisions.

Four courts in three states have held antistalking statutes constitutional. One court read the statute to require specific intent for all aspects of the crime. *See The Week's Opinions*, MASS L. WKLY., July 5, 1993, at 23 (construing statute as requiring that the acts of repeated following and harassment be made with the intent to put another in fear of bodily harm). Another court held that the phrase "substantial emotional distress" in the Florida felony section was clarified by the further requirement of "reasonable fear" of death or bodily injury. *See* State v. Slater, No. CRC 92-15850 CFANO (Fla. Cir. Ct. Feb. 18, 1993). Other courts upheld stalking statutes through bench rulings with, apparently, little constitutional analysis. *See* Wes Platt, *Stalking Charge to Stand, Judge Says*, ST. PETERSBURG TIMES, Aug. 6, 1993, at 1 (the Florida law, according to the court as quoted in the article, is "long overdue"); Charles Mount, *Stalking Law Survives First Test*, CHI. TRIB. June 18, 1993, at 4 (noting that court upheld the Illinois law).

[9] Commonwealth v. Camper, No. 93-2876 (Va. Gen. Dist. Ct. Apr. 9, 1993).

[10] *See* Wayne Slater, *Abortion Foes, Labor Oppose Stalking Bill*, DALLAS MORNING NEWS, Mar. 5, 1993, at 22A.

[11] *See* Wayne E. Bradburn, Jr., Comment, *Stalking Statutes: An Ineffective Legislative Remedy for Rectifying Perceived Problems With Today's Injunction System*, 19 OHIO N.U. L. REV. 271 (1992); Miles Corwin, *When the Law Can't Protect*, L.A. TIMES, May 8, 1993, at A1; Richard Seven, *Stalking Law Too Narrow, Lawyers, Victims Complain*, SEATTLE TIMES, July 6, 1993, at B1.

Part I of this Article describes existing antistalking legislation, noting actual or potential applications of the new laws to political, religious or traditionally "legitimate" activity.[12] Part II discusses the numerous constitutional issues raised by such legislation, concluding that, unfortunately, existing antistalking legislation is facially unconstitutional.[13] Part III examines the model legislation recently proposed by the National Criminal Justice Association, concluding that the negligence standard advanced in that model should be removed to avoid constitutional infirmity.[14] Part IV proposes a model statute that will be both constitutionally enforceable and, hopefully, more effective in preventing stalking.[15] The assumption throughout is that an unconstitutional criminal statute is both unfair to the accused and useless to the victim, who needs the protection of a criminal law capable of being upheld in court.

I. The Development and Application of Antistalking Legislation

There are an estimated 200,000 stalkers in the United States.[16] According to a recent treatment of the subject, the "typical" stalker can belong to one of four psychological profiles.[17] He or she may suffer from "erotomania," which is the "erotic delusion that one is loved by another," or from "borderline erotomania," in which the stalker knows that the target does not return his or her intense emotional feelings.[18] Also common, and often more violent, are the "former intimate" stalkers, estranged former

---

[12] *See infra* notes 16–63 and accompanying text.

[13] *See infra* notes 64–197 and accompanying text.

[14] *See infra* notes 198–209 and accompanying text.

[15] *See infra* notes 210–227 and accompanying text

[16] Gilligan, *supra* note 1, at 286 (citing *Stalking Victims Turn to Congress for Help* (CNN television broadcast, Sept. 29, 1992)); Corwin, *supra* note 11; *see also* David Holmstrom, *Efforts to Protect Women From 'Stalkers' Gain Momentum at State, Federal Levels*, Christian Sci. Monitor, Dec. 22, 1992, at 1 ("About 200,000 stalking cases are reported each year nationwide, victims rights organizations estimate."). No attempt has been made to verify statistics reported in the various sources cited in this Article. Given the absence of a consensus definition of "stalking," such figures should be treated with skepticism. However, the anecdotal evidence noted herein amply demonstrates that stalking presents a national problem of tremendous importance.

[17] Kathleen G. McAnaney et al., Note *From Imprudence to Crime: Anti-Stalking Laws*, 68 Notre Dame L. Rev. 819, 831–43 (1993).

[18] *Id.* at 832–38.

partners who may have a history of violence.[19] The most deadly types are the "sociopathic stalkers," serial murderers and rapists who stalk their victims, although often not with the intent to make their surveillance known.[20]

Despite these rough categorizations of certain types of stalkers, there does not appear to be a precise definition of criminal stalking behavior. The term "stalking" should not be limited to individual (as opposed to group) behavior[21] or confined to particular classes of "stalkers" or "victims." "[S]talkers can be obsessed fans, divorced or separated spouses, ex-lovers, rejected suitors, neighbors, coworkers, classmates, gang members, former employees, [or] disgruntled defendants, as well as complete strangers."[22] To these categories we may also add political and religious extremists whose conduct crosses the hazy line between legitimate protest and criminal behavior.[23]

Although both stalkers and their victims can be male or female, women appear to be the primary target of stalking behavior. It may be that "at least half of the women who leave violent relationships are subsequently followed and threatened by their abusers."[24] An advocate against domestic violence has estimated "that as many as ninety percent of the women killed by (former) husbands or boyfriends were stalked before a murder occurred."[25] Current predictions suggest that "about five percent of

---

[19] *Id.* at 838–41.

[20] *Id.* at 842–43.

[21] For a brief discussion of the possibility of group stalking see *id.* at 829 n.42.

[22] *Id.* at 821–23 (footnotes omitted) (citing supporting press accounts); *see also* Guy, *supra* note 1, at 995 n.30 ("Statistics from Los Angeles show that stalking victims consist of 17% highly recognizable celebrities, 32% lesser known entertainment figures, 13% former employers or other professionals, and 38% other people.") (citing Maria Puente, *Legislators Tackling the Terror of Stalking*, USA TODAY, July 21, 1992, at 9A).

[23] An Operation Rescue volunteer was arrested on stalking charges for repeatedly following an abortion clinic director, telling her that she should get a bullet proof vest because she "might be next" (a reference to the murder of an abortion clinic doctor), encouraging others to rip off the arms of the director's daughter so that she could know how aborted fetuses feel, and disclosing an intimate knowledge of the layout of the director's home. *See* Sandra G. Boodman, *Abortion Foes Strike at Doctors' Home Lives*, WASHINGTON POST, Apr. 8, 1993, at A1; *Anti-Abortion Activist Charged With Stalking Clinic Director*, CHI. TRIB., Apr. 2, 1993, at 4. Although the arrest of this woman raises constitutional issues, there is little doubt that, in the right circumstances, this kind of conduct could be proscribed as a kind of "stalking."

[24] *Developments in the Law: Legal Responses to Domestic Violence*, 106 HARV. L. REV. 1498, 1534 (1993) (citing ANGELA BROWNE, WHEN BATTERED WOMEN KILL 101 (1987)).

[25] McAnaney et al., *supra* note 1, at 838 (citing Melinda Beck, et al., *Murderous Obsession*, NEWSWEEK, July 13, 1992, at 60).

women in the general population will be victims of stalking at some time in their lives."[26]

Despite these disturbing figures, existing legal remedies for stalking are grossly inadequate, either because they are too weak or because they involve dangerous delays.[27] Typical criminal statutes prohibiting stalking conduct either insufficiently cover the spectrum of stalking behavior or require a threat of immediate bodily harm.[28] Civil claims, including suits for trespass, invasion of privacy, and intentional infliction of emotional distress, are costly to the victim and usually fail to deter obsessive stalking behavior.[29] Similarly, restraining orders involve substantial legal costs, ask too much of police, are often too narrowly drawn, and carry relatively insignificant contempt penalties in the event of non-compliance.[30]

In light of these ineffective remedies, at least forty-three states have passed antistalking legislation,[31] and the remaining states, as well as the federal government,[32] are considering similar bills.

A. *Defining the Crime*

The crux of the controversy surrounding antistalking legislation lies in the definition of criminal conduct. Three statutes from highly populated states have been chosen for the purposes of this Article to show the difficulties inherent in crafting language that encompasses the diverse forms of stalking while simultaneously excluding constitutionally protected and otherwise legitimate conduct.

1. California

California led the battle to criminalize stalking and has one of the most detailed antistalking statutes. In California, a "stalker" is defined as:

---

[26] Guy, *supra* note 1, at 995 (citing Maria Puente, *Legislators Tackling the Terror of Stalking*, USA TODAY, July 21, 1992, at 9A.

[27] For extensive discussions of the inadequacy of existing remedies, see Gilligan, *supra* note 1, at 288–99; Guy, *supra* note 1, at 996–1000; Strikis, *supra* note 2, at [ ].

[28] Matthew J. Gilligan, Note, *Stalking the Stalker: Developing New Laws to Thwart Those Who Terrorize Others*, 27 GA. L. REV. 285, 295–99 (1992).

[29] *Id.* at 289–93.

[30] *Id.* at 293–94.

[31] *See supra* note 3.

[32] *See, e.g.*, S. 470, 103d Cong., 1st Sess. (1993); H.R. 840, 103d Cong., 1st Sess. (1993); H.R. 740, 103d Cong., 1st Sess. (1993).

> [a]ny person who willfully, maliciously, and repeatedly follows or harasses another person and who makes a credible threat with the intent to place that person in reasonable fear of death or great bodily injury or to place that person in reasonable fear of the death or great bodily injury of his or her immediate family . . . .[33]

The statute defines "harasses" as

> a knowing and willful course of conduct directed at a specific person which seriously alarms, annoys, harasses, or terrorizes the person, and which serves no legitimate purpose. The course of conduct must be such as would cause a reasonable person to suffer substantial emotional distress, and must actually cause substantial emotional distress to the person. "Course of conduct" means a pattern of conduct composed of a series of acts over a period of time, however short, evidencing a continuity of purpose. Constitutionally protected activity is not included within the meaning of "course of conduct."[34]

For a threat to be "credible," it must be "made with the intent and the apparent ability to carry out the threat so as to cause the person who is the target of the threat to reasonably fear for his or her safety or the safety of his or her immediate family."[35] The statute specifically excludes otherwise illegal stalking if it "occurs during labor picketing."[36]

Although the drafters of this statute clearly attempted to satisfy constitutional concerns, several problems are immediately apparent. The statute includes a number of vague concepts, including: (1) intent to carry out the threat;[37] (2) "no legitimate purpose"; (3) "harasses"; (4) "substantial emotional distress"; and (5) "constitutionally protected activity." In addition, the exemption for constitutionally protected activity applies only to harassment, not to "following." It might be stalking, for example, if anti-Ku Klux Klan protestors shouted credible threats

---

[33] CAL. CODE § 646.9(1) (West 1993).
[34] *Id.* § 646.9(e).
[35] *Id.* § 646.9(f).
[36] *Id.* § 646.9(g).
[37] This presents a different problem than law enforcement officers and juries usually face in determining intent. Ordinarily they are able to view the circumstances surrounding a concrete act and infer an intent to *do that same act*. With respect to the California antistalking law, they are asked to infer a culpable intent to commit a *future* act (e.g., murder) based on a verbal threat in the present. *See* Guy, *supra* note 1, at 1014 (attempting to construe this problem out of the statute). This raises the strong likelihood that two police officers would treat the same threat differently, one concluding that harm was imminent, the other that the actor "didn't mean it."

during an altercation with Klan marchers and then followed them for the rest of the day in an attempt to disrupt the march and other scheduled events.[38] Also, the specific intent to cause fear of violence applies only to the "credible threat." Accordingly, non-threatening acts of following, which by themselves cause no harm,[39] could be converted into illegal stalking by a subsequent threat. Moreover, years could pass between the requisite acts so long as there was a "continuity" of purpose. Finally, *acts* of stalking do not constitute the *crime* of stalking if they occur during labor picketing.

2. Florida

The Florida statute poses more significant problems. The pertinent provisions are as follows:

> (1) As used in this section:
>
> (a) "Harasses" means to engage in a course of conduct directed at a specific person that causes substantial emotional distress in such person and serves no legitimate purpose.
>
> (b) "Course of conduct" means a pattern of conduct composed of a series of acts over a period of time, however short, evidencing a continuity of purpose. Constitutionally protected activity is not included within the meaning "course of conduct." Such constitutionally protected activity includes picketing or other organized protests.
>
> (c) "Credible threat" means a threat made with the intent to cause the person who is the target of the threat to reasonably fear for his or her safety. The threat must be against the life of, or a threat to cause bodily injury to, a person.
>
> (2) Any person who willfully, maliciously, and repeatedly follows or harasses another person commits the offense of stalking, a misdemeanor of the first degree . . . .
>
> (3) Any person who willfully, maliciously, and repeatedly follows or harasses another person, and makes a credible threat with the intent to place that person in reasonable fear

[38] As discussed *infra* notes 129–131 and accompanying text, the requirement that the following be "willful" and "malicious" includes the concepts of "annoyance" and "vexing." Presumably the counterprotestors could intend to "vex" the Klansmen and still be protected by the First Amendment.

[39] The act of following, for example, could have taken the form of secret surveillance which became known to the "victim" only after the threat occurred.

> of death or bodily injury, commits the offense of aggravated stalking, a felony of the third degree . . . .
>
> . . . .
>
> (5) Any law enforcement officer may arrest, without a warrant, any person he or she has probable cause to believe has violated the provisions of this section.[40]

Like the California statute, the Florida law does not exempt constitutionally protected "following." Indeed, by its terms the law is broad enough to criminalize any repeated following by a group of protestors, *whether or not accompanied by a threat.* Even the most peaceful protest would be a misdemeanor if it involved repeated following. In addition, if credible threats of violence were shouted at any time before or after the acts of following, the protestors could be charged with a felony.

Outside the political context, police officers, private investigators, journalists, jealous spouses, or would-be paramours could be charged with a misdemeanor in Florida if evidence suggested that the defendant intended to annoy the complainant or had some other improper purpose. For example, a Florida college professor was recently charged with "stalking" a former student by "harassing" her with letters and phone calls in an attempt to dissuade her from dating someone else; the communications were apparently neither obscene nor overtly threatening.[41]

3. Michigan

Michigan has taken a much different approach than that adopted by California and Florida. Rather than speak of "credible threats," the statute's principal elements are "emotional distress" and "unconsented contact:"

> (1) As used in this section:
>
> (a) "Course of conduct" means a pattern of conduct composed of a series of 2 or more separate noncontinuous acts, evidencing a continuity of purpose.
>
> (b) *"Emotional distress" means significant mental suffering or distress that may, but does not necessarily require, medical or other professional treatment or counseling.*

---

[40] FLA. STAT. CH. 784.048 (1992).

[41] Laurin Sellers, *BCC Professor is Charged With Stalking an Ex-Student*, ORLANDO SENTINEL TRIB., Feb. 25, 1993, at B1.

(c) "Harassment" means conduct directed toward a victim that includes, but is not limited to, repeated or continuing unconsented contact, that would cause a reasonable individual to suffer emotional distress, and that actually causes the victim to suffer emotional distress. Harassment does not include constitutionally protected activity or conduct that serves a legitimate purpose.

(d) "Stalking" means a willful course of conduct involving repeated or continuing harassment of another individual that would cause a reasonable person to feel terrorized, frightened, intimidated, threatened, harassed, or molested.

(e) *"Unconsented contact" means any contact with another individual that is initiated or continued without that individual's consent, or in disregard of that individual's expressed desire that the contact be avoided or discontinued.* Unconsented contact includes, but is not limited to, any of the following:

(i) Following or appearing within the sight of that individual.

(ii) Approaching or confronting that individual in a public place or on private property.

(iii) Appearing at the workplace or residence of that individual.

(iv) Entering onto or remaining on property owned, leased, or occupied by that individual.

(v) Contacting that individual by telephone.

(vi) Sending mail or electronic communications to that individual.

(vii) Placing an object on, or delivering an object to, property owned, leased, or occupied by that individual.

(f) "Victim" means an individual who is the target of a willful course of conduct involving repeated or continuing harassment.

(2) An individual who engages in stalking is guilty of a misdemeanor punishable by imprisonment for not more than 1 year or a fine of not more than $1,000.00, or both.[42]

This is also a troubling statute. First, the "emotional distress" standard is circular and does not limit the statute's sweep to conduct and speech that threatens imminent physical harm. Second, although the statute's constitutional disclaimer applies to all major elements of the crime, the "constitutionally protected

[42] MICH. COMP. LAWS § 750.411h (1992) (emphasis added).

activity" remains undefined. Third, there is probably no requirement of specific intent to stalk or even to cause emotional distress; rather, the only intent necessary to establish the crime is "willfulness," or the intent to do the act.[43] Fourth, the definition of "unconsented contact" raises the possibility of conviction even though the actor was unaware that the victim did not desire the "contacts." Reduced to its essential terms the statute provides that stalking means any repeated communication or contact, not expressly consented to, which may or may not cause a reasonable person to seek psychiatric help and which may or may not have caused the victim to seek psychiatric help, unless the conduct is constitutionally protected or otherwise legitimate. In terms of clarity, therefore, the Michigan law receives the lowest marks of all.

In many respects, most other statutes are similar to the California and Florida legislation, but some include troublesome characteristics such as denial of bail;[44] the failure to include exemptions for any constitutionally protected activity;[45] the failure to require that the threat or threatening conduct be "credible;"[46] the definition of stalking as following, harassing, *or* making a credible threat;[47] a negligence standard;[48] and the failure to

---

[43] As noted *infra* notes 127–129 and accompanying text, a Michigan court probably, though not certainly, would conclude that no specific intent is required by this statute.

[44] *See* 725 ILL. ANN. STAT. 5/110-4 (Smith-Hurd 1993).

[45] *See* COLO. REV. STAT. § 18-9-111 (1992); IOWA CODE § 708.11 (1992); MASS. ANN. LAWS ch. 265, § 43 (Law. Co-op. 1993); N.Y. PENAL LAW § 240.25 (Consol. 1993); OHIO REV. CODE ANN. § 2903.211 (Baldwin 1993); 1993 PA. LAWS 28; TENN. CODE ANN. § 39-17-315 (1992); VA. CODE ANN. § 18.2-60.3 (Michie 1993); WASH. REV. CODE ANN. § 9A.46.110 (West 1992); WIS. STAT. § 947.013 (1991-92); *see also* 720 ILL. COMP. STAT. 5/12-7.3 (1993) (exempting only for lawful picketing). The Alabama law adds a unique provision requiring that the antistalking section "shall be construed and, if necessary, reconstrued to sustain its constitutionality." ALA. CODE § 13A-6-94 (1992). It is unlikely that this provision adds anything to the courts' general obligation to construe statutes consistently with the Constitution.

[46] *See* ARIZ. REV. STAT. ANN. § 13-2921 (1992); 1992 D.C. STAT. 269; KY. REV. STAT. ANN. § 508.130 (Baldwin 1992); N.Y. PENAL LAW § 240.25 (Consol. 1993); 1993 N.D. LAWS 120; VA. CODE ANN. § 18.2-60.3 (Michie 1993); WASH. REV. CODE § 9A.46.110 (1992); WIS. STAT. § 947.013 (1991–92) (no threat for misdemeanor harassment).

[47] DEL. CODE ANN. tit. 11, § 1312A (1992); 1992 MISS. LAWS 532; N.J. REV. STAT. § 2C:12-10 (1992); *see also* 1992 D.C. STAT. 269 (conduct with intent to cause emotional distress or fear of death or bodily injury); 1993 PA. LAWS 28 (intent to cause fear of bodily injury *or* substantial emotional distress); S.D. CODIFIED LAWS ANN. § 22-19A-1 (1992) (amended to change "and" to "or").

[48] *See* 1993 N.D. LAWS 120 ("[N]or is it a defense that the actor did not intend to frighten, intimidate, or harass the person."); WASH. REV. CODE § 9A.46.110 (1992) ("knows or reasonably should know that the person being followed is afraid, intimidated, or harassed even if the stalker did not intend to place the person in fear or intimidate or harass the person"). The Wyoming statute is ambiguous, defining "harass"

require that the target of the stalking be adversely affected.[49] These characteristics, some of which are discussed below in greater detail, further undermine the constitutionality of the individual statutes.

B. *Application of Existing Antistalking Legislation by Law Enforcement Authorities*

Where antistalking laws exist, police have arrested large numbers of people. Though there are no national statistics for stalking charges actually filed, one source reports that the Los Angeles "city attorney's office handles about 500 stalking cases a year."[50] From July 1992 to February 1993, prosecutors in Cook County, Illinois (encompassing Chicago) "posted about 110 stalking charges, and the number continues to climb each week . . . ."[51] In a nine-month period, thirty-six stalking arrests occurred in Fairfax County, Virginia, an area just outside of Washington, D.C.[52] The Florida Department of Law Enforcement reported that 468 people were charged with stalking in a six-month period.[53] Michigan charged 355 people during a similar time frame.[54]

Without doubt, these high numbers reflect the seriousness of the stalking problem and the need for an effective law enforcement remedy. However, they also raise concerns that the statutes are so broad or vague that people are being arrested for relatively harmless and constitutionally protected conduct. The press has reported the following antistalking charges:

---

to mean conduct "which the defendant knew or should have known would cause a reasonable person to suffer substantial emotional distress" and then defining "stalking" as "intent to harass." WYO. STAT. § 6-2-506 (1993). *See also infra* notes 198–209 and accompanying text (discussing model statute proposed by the National Institute of Justice).

[49] *See* 720 ILL. ANN. STAT. 5/12-7.3 (Smith-Hurd 1993); IOWA CODE § 708.11 (1992); MASS. ANN. LAWS ch. 265, § 43 (Law. Co-op. 1993); N.M. STAT. ANN. § 30-3A-3 (Michie 1993); N.C. GEN. STATS. § 14-277.3. (1992); VA. CODE ANN. § 18.2-60.3 (Michie 1992).

[50] Miles Corwin, *When the Law Can't Protect*, L.A. TIMES, May 8, 1993, at A1.

[51] Mike Truppa, *Suburban Prosecutors Call Stalking Law Too Strict*, CHI. DAILY L. BULL., Feb. 12, 1993, at A1.

[52] Patricia Davis, *New Stalking Law Flushing the Crime Into the Open in Va.*, WASH. POST, Jan. 24, 1993, at B1.

[53] *468 Charged with Stalking, But Victims Say Police Wait Too Long*, MIAMI HERALD, Jan. 18, 1993, at B5.

[54] *See 355 Charged Under Michigan Stalking Law*, CHI. TRIB., Aug. 3, 1993, at 3.

- A councilwoman was charged with stalking a female police dispatcher. According to the dispatcher, the councilwoman "was stalking her in an attempt to discredit her" in a lawsuit that the dispatcher had filed against a police sergeant. The councilwoman admitted to "check[ing] up on the dispatcher to see if the dispatcher was abusing sick leave."[55] No threats or threatening conduct occurred.
- A Georgia man was arrested for stalking because he showed up at city council meetings and stared at his brother-in-law, a councilman.[56]
- In the wake of the murder of Dr. David Gunn, many Operation Rescue right-to-life protestors have been arrested on stalking charges based in part on conduct that could be described as legitimate, though extreme, protest. For example,
  - Two abortion protestors were enjoined from "stalking" after they followed a Planned Parenthood director to a restaurant and one began yelling that the director killed babies. The same man had later informed the director that it would happen again and she was then followed by the two protestors. There was no suggestion that the two men intended to do anything other than repeat their earlier performance.[57]
  - A man was arrested for following a clinic security guard several times during one week. No overt threat was made, the follower would not roll down his window when confronted, and the victim was angered but not frightened.[58]
- Four prochoice volunteers were recently arrested under Minnesota's antistalking law for following an Operation Rescue group. It was an isolated incident, there was no evidence of a threat, and no evidence of emotional distress.[59]

---

[55] Norman Parish, *Official Pleads Not Guilty to Stalking*, PLAIN DEALER, June 18, 1993, at 1B.

[56] Anne Rochell, *Stalking Charges Against Official's Kin Dismissed*, ATLANTA J. & CONST., June 10, 1993, at H4.

[57] *See* Judith Gaines, *R.I. Abortion Rights Leader Challenges Alleged Stalking By Foes*, BOSTON GLOBE, Apr. 24, 1993 at 20.

[58] Maureen M. Smith, *Operation Rescue Trainee Arrested*, STAR TRIB., June 26, 1993, at 1B. There is some indication of a concerted effort by prochoice forces to invoke antistalking laws and Dr. Gunn in an effort to counter right-to-life tactics. Ron Fitzsimmons of the National Abortion Federation has been quoted as saying that, "[b]efore Dr. Gunn was killed, our folks were nervous, but it was hard to prove they were in fear. But soon after David was killed, I sent a notice out to all of my clinics saying, 'If you have stalking laws, this is the time to use them.'" *Stalking Laws: Pro-Choicers Look to Measures for Relief*, ABORTION REP., Apr. 22, 1993.

[59] Maureen M. Smith, *4 Abortion Rights Volunteers Arrested Under State's New Anti-Stalking Law*, STAR TRIB., July 12, 1993, at 5A.

- A jury deliberated for less than an hour before acquitting a Chicago man charged with threatening his former wife during a phone conversation and "intimidating" her at a pizza parlor where she worked. The charge was based on the allegation of the ex-wife, who was apparently angry over their recent divorce. The man had spent 132 days in jail,[60] presumably under the Illinois statute's no-bail provision.
- A Florida teen was charged with stalking his ex-girlfriend based on coming to her home three times in one morning, knocking on the door and looking in the windows.[61]
- A former Chicago police officer was arrested and convicted of stalking his estranged wife based primarily on a single bizarre incident in which he was admitted into his wife's home by his son, barged into his wife's bedroom, and videotaped her in bed with another man. By all accounts, his primary intent was to gather evidence in an upcoming divorce proceeding. The videotape that led to his conviction disclosed threats to the man, but none were directed at his wife.[62]
- "In its bill of particulars relating to the stalking charge, the government alleged that on one occasion the defendant went to the complaining witness's place of employment, a radio station, and asked an employee of the station to deliver an envelope full of materials to the complaining witness, and later on was seen posting flyers containing derogatory comments about the complaining witness on the front door and window of the radio station. The government further alleged that on other occasions the defendant made numerous unwanted telephone calls to the complaining witness and stood outside of the radio station while the complaining witness was inside."[63]

These accounts are troubling because they indicate that antistalking legislation is being applied to conduct that cannot easily

[60] Curtis Lawrence, *1st Stalking Trial Results in Acquittal*, CHI. TRIB., Dec. 19, 1992, at 5.

[61] Rene Stutzman, *Teen Charged With Stalking Ex-Girlfriend*, ORLANDO SENTINEL TRIB., July 28, 1992, at 1.

[62] Robert Enstad, *Tape Leads to Stalking Conviction*, CHI. TRIB., Dec. 8, 1992 (Chicagoland) at 1; Robert Enstad, *Videotape Stalker Gets Probation*, CHI. TRIB., Jan. 26, 1993, (Lake) at 1; Steven R. Lake, *Stalking Law Sometimes Becomes a Family Matter*, CHI. DAILY L. BULL., Sept. 4, 1992, at 5.

[63] United States v. Smith, No. M-6400-93 (D.C. Super. Ct. Sept. 9, 1993) (noting that "posting of derogatory materials" and protest "following" may be protected by the Constitution).

be described as erotomaniacal, sociopathic or, in many cases, even overtly aggressive. Worse, it appears that political and religious speech, movement and activity have been found sufficient to trigger arrest under the new laws.

II. The Constitutionality of Existing Antistalking Legislation

With the possible exceptions of Alaska,[64] Connecticut,[65] Maryland[66] New Mexico,[67] and North Carolina,[68] no state has a constitutionally valid antistalking law. Depending on their exact wording, the laws sweep too broadly, suffer from acute cases of vagueness, and/or employ content-based distinctions based on the perceived "message" of the stalker. These defects, furthermore, have not been cured by any "procedural safeguards" designed to protect First Amendment concerns. Indeed, some antistalking statutes are doubly defective: by authorizing war-

---

[64] *See* discussion *infra* notes 212–215 in connection with the authors' proposed model legislation.

[65] The Connecticut law only covers following and lying in wait and requires (1) a specific intent to cause fear for one's physical safety and (2) reasonable fear for physical safety in the victim. Conn. Gen. Stat. § 53a-181d (1992). The high mens rea requirement, combined with "reasonable person" causation may save the statute from unconstitutionality, *see* Guy, *supra* note 1, at 1020–21, but it also excludes a substantial amount of stalking behavior.

[66] The Maryland law criminalizes "a malicious course of conduct that includes approaching or pursuing another person with intent to place that person" in reasonable fear of his own or another's serious bodily injury. 1993 Md. Laws 205. This law may be constitutional if it is construed by the courts as requiring a specific intent to cause fear of death or bodily harm for any "malicious course of conduct." Without that saving construction, the statute would clearly be unconstitutional for failing to define "malicious conduct." On the other hand, such a construction would not permit the statute to reach the optimum amount of stalking behavior. *See infra* notes 206–210 and accompanying text.

[67] The New Mexico law prohibits "knowingly pursuing a pattern of conduct that poses a credible threat . . . and that is intended to place that person in reasonable apprehension of death" *and* requires that the following, surveillance, or harassment be repeated and "in furtherance of the threat." N.M. Stat. § 30-3A-3 (1993). If "in furtherance of the threat" means that the threat must precede acts of following and harassment, then this statute is probably constitutional, though more precision would be desirable and though it, too, may cover less stalking than is constitutionally permissible. *See infra* notes 206–210 and accompanying text.

[68] North Carolina only criminalizes "following" that is specifically intended to cause reasonable fear of death or bodily injury *after* a reasonable warning or request to desist. N.C. Gen. Stat. § 14-277.3 (1992). The North Carolina statute is probably constitutional but the requirement of prior notification of the police makes it woefully underinclusive. *See generally* Melissa Perrell Phipps, Note, *North Carolina's New Anti-Stalking Law: Constitutionally Sound, But is it Really a Deterrent?*, 71 N.C. L. Rev. 1933 (1993).

rantless arrests and the denial of bail, they actually weaken safeguards that operate to protect speech and fundamental liberties.

A. *Facial Invalidity Under Overbreadth and Vagueness Doctrines*

The First and Fourteenth Amendments generally prevent government from "abridging the freedom of speech."[69] Similarly, the Supreme Court has long recognized a constitutional right to liberty of movement.[70] Even a cursory review of existing antistalking laws and their application compels the conclusion that they directly restrict speech and movement to a large degree. Indeed, one may be convicted of stalking solely on the basis of communication—for example, through telephone calls and letters containing express or implied threats—and several persons have been arrested solely on the basis of movement on public grounds and roadways.

This is not to say, however, that antistalking laws are invalid simply because they criminalize the content of speech or the character of movement. It is well settled that a statute may restrict speech if it serves "a compelling state interest and is narrowly drawn to achieve that end."[71] The somewhat lesser "intermediate scrutiny" standard probably applies to liberty of movement, requiring only that the statute be "narrowly tailored to serve significant government interests."[72] Moreover, certain

---

[69] U.S. CONST. amend. I. The First Amendment applies to the states through the due process clause of the Fourteenth Amendment. Gitlow v. New York, 268 U.S. 652, 666 (1925).

[70] *See* Attorney Gen. of N.Y. v. Soto-Lopez, 476 U.S. 898, 901 (1986) (holding that freedom of travel throughout the United States is a basic right); Kolender v. Lawson, 461 U.S. 352, 358 (1983) (statute which required loiterers to produce "credible and reliable" identification implicated First Amendment and freedom of movement freedoms); Papachristou v. City of Jacksonville, 405 U.S. 156, 163–64 (1972) (holding vagrancy statute void for vagueness where it implicated liberty of movement); Shapiro v. Thompson, 394 U.S. 618, 629 (1969), *overruled on other grounds by* Edelman v. Jordan, 415 U.S. 651 (1974); Spencer v. Casavilla, 903 F.2d 171, 174 (2d Cir. 1990) (holding intrastate travel to be a basic constitutional right); Lutz v. City of York, 899 F.2d 255, 268 (3d Cir. 1990) (freedom to move about neighborhood or town is a constitutional right).

[71] Simon & Schuster, Inc. v. New York Crime Victims Bd., 112 S.Ct. 501, 509 (1991) (quoting Arkansas Writers' Project, Inc. v. Ragland, 481 U.S. 221, 231 (1987)).

[72] *Lutz*, 899 F.2d at 269 (right of intrastate movement exists in Fourteenth Amendment substantive due process and requires only intermediate scrutiny). *But see* Gayle v. Governor of Guam, 414 F. Supp. 636, 639 (D. Guam 1976) (citing Carroll v. United States, 267 U.S. 132 (1924)); People *ex rel.* J.M., 768 P.2d 219, 221 (Colo. 1989) ("[T]he state must establish a compelling interest before it may curtail the exercise of such rights.").

types of speech, among them "fighting words" which "by their very utterance inflict injury" may be proscribed[73] under most circumstances.[74] Therefore, there is a relatively broad zone of speech and movement which legislatures may constitutionally restrict.

1. Overbreadth

Antistalking legislation serves a compelling government interest to the extent that it targets stalkers who threaten serious violence, whether or not they ever follow through on the threat. The Supreme Court has never suggested that government does not have a paramount interest in prohibiting the threat of force.[75] The controversial question, however, is whether these statutes have been narrowly drawn to effectuate their salutary purpose without unnecessarily infringing on First Amendment and liberty of movement rights. In First Amendment parlance, the issue is whether the legislation suffers from overbreadth.

Ordinarily, a party may not challenge the application of a statute solely because it might be unconstitutional as applied to someone else.[76] However, since 1940[77] the courts have made an exception in the case of the First Amendment, principally because of the danger that an overly broad statute would "chill" protected speech.[78] Under the "overbreadth" doctrine,

> [w]hen speech or expressive activity forms a significant part of a law's target, the law is subject to facial challenge and

---

[73] Gooding v. Wilson, 405 U.S. 518, 525 (1972); Chaplinsky v. New Hampshire, 315 U.S. 568, 572 (1942).

[74] R.A.V. v. City of St. Paul, 112 S.Ct. 2538, 2543–50 (1992).

[75] The same may not be true of other "stalking"-type behavior, such as sending excessive numbers of flowers and messages of love. Therefore, Florida and Michigan-type statutes, which can be construed as forbidding nonthreatening behavior, may also be required to justify that legislative goal as "compelling" or "significant." The more a statute criminalizes traditionally legitimate behavior, the more difficult it will be to establish a valid and important government purpose in enacting the statute.

[76] Richard A. Fallon, *Making Sense of Overbreadth*, 100 YALE L.J. 853, 859–60 (1991).

[77] Scholarly commentary generally traces the origins of overbreadth doctrine to Thornhill v. Alabama, 310 U.S. 88 (1940). *See* Fallon, *supra* note 76, at 863 (citing Martin H. Redish, *The Warren Court, the Burger Court and the First Amendment Overbreadth Doctrine*, 78 NW. U. L. REV. 1031, 1038–39 (1984); Henry P. Monaghan, *Overbreadth*, 1981 SUP. CT. REV. 1, 11 (1973)).

[78] *See, e.g.*, New York v. Ferber, 458 U.S. 747, 772 (1982) ("[A] sweeping statute . . . has the potential to repeatedly chill the exercise of expressive activity by many individuals."); United States Civil Serv. Comm'n v. National Ass'n of Letter Carriers, 413 U.S. 548, 598 (1973) (Douglas, J., dissenting) ("Overbreadth in the area of the First Amendment has a peculiar evil, the evil of creating chilling effects which deter the exercise of those freedoms.").

> invalidation if: (i) it is "substantially overbroad"—that is, if its illegitimate applications are too numerous "judged in relation to the statute's plainly legitimate sweep," and (ii) no constitutionally adequate narrowing construction suggests itself.[79]

Because overbreadth is "strong medicine,"[80] courts are reluctant to invalidate statutes on this ground.

Nevertheless, "[c]riminal statutes must be scrutinized with particular care"[81] when challenged for overbreadth because the consequences of imprecision are so great.[82] Courts have been particularly willing to invalidate statutes which tend to criminalize behavior based on emotional distress, as opposed to physical violence. Harassment statutes, for example, have been held to be overbroad where their terms readily encompassed constitutionally protected speech.[83] Likewise, statutes criminalizing "abusive language"[84] and "disorderly conduct"[85] have also received close

---

[79] Fallon, *supra* note 76, at 863 (footnotes omitted). For the proposition that the statute must be "substantially overbroad," see Gentile v. State Bar, 111 S.Ct. 2720 (1991); FW/PBS, Inc. v. City of Dallas, 493 U.S. 215, 259 (1990) ("[W]e cannot sustain the present facial attack unless the ordinance is *substantially* overbroad, that is, unless it reaches a substantial number of impermissible applications judged in relation to the statute's plainly legitimate sweep.") (citations and quotations omitted); Board of Trustees v. Fox, 492 U.S. 469, 485 (1989); Hoffman Estates v. Flipside, Hoffman Estates, Inc., 455 U.S. 489, 494 (1982); *Ferber*, 458 U.S. at 770–71. For the requirement that a court look for an appropriate narrowing construction, see Erznoznik v. City of Jacksonville, 422 U.S. 205, 216 (1975); Dombrowski v. Pfister, 380 U.S. 479, 491 (1965).

[80] Osborne v. Ohio, 495 U.S. 103, 122 (1990) (citing Broadrick v. Oklahoma, 413 U.S. 601, 613 (1973)).

[81] City of Houston v. Hill, 482 U.S. 451, 459 (1987) (citing Winter v. New York, 333 U.S. 507, 515 (1948)).

[82] *Hoffman Estates*, 455 U.S. at 498–99 ("The Court has also expressed greater tolerance of enactments with civil rather than criminal penalties because the consequences of imprecision are qualitatively less severe.").

[83] *See* Dorman v. Satti, 862 F.2d 432, 436–37 (2d Cir. 1988) (invalidating Connecticut Hunter Harassment Act as overbroad and vague where statute failed to define "interference" and "harassment," and because "acts in preparation" clause was not limited in time, place and manner), *cert. denied*, 490 U.S. 1099 (1989); State v. Dronso, 279 N.W.2d 710 (Wis. 1979) (telephone harassment statute overbroad). *But see* Gormley v. Director, 632 F.2d 938, 941–42 (2d Cir. 1980) (upholding telephone harassment statute where act penalized the making of a call with intent to annoy, whether or not conversation ensued).

[84] Gooding v. Wilson, 405 U.S. 518, 519–20 (1972) (invalidating statute criminalizing "opprobrious words or abusive language, tending to cause a breach of the peace"); United States v. Sturgill, 563 F.2d 307, 309–10 (6th Cir. 1977) (invalidating as overbroad Kentucky harassment statute criminalizing "abusive language" uttered "with intent to harass, annoy, or alarm another person").

[85] *See* R.A.V. v. City of St. Paul, 112 S.Ct. 2538, 2558–60 (1992) (White, J., concurring) (four justices agreeing that bias-motivated disorderly conduct statute was overbroad where the state court's narrowing construction permitted the statute to proscribe expression that "'by its very utterance' causes anger, alarm or resentment'");

scrutiny because such statutes are inherently difficult to draft without infringing on protected First Amendment activities.

The closest analogy to antistalking laws are the so-called "intimidation" statutes and other laws designed to prohibit threats. With respect to the constitutionality of these statutes, several principles apply. First, as noted above, threats of violence fall outside the traditional protection of the First Amendment.[86] However, "a statute, . . . which makes criminal a form of pure speech, must be interpreted with the commands of the First Amendment clearly in mind. What is a threat must be distinguished from what is constitutionally protected speech."[87] Thus, the threat, to be prohibited, must be of a violent nature, and "implicit in the nature of such punishable threats is a reasonable tendency to produce in the victim a fear that the threat will be carried out."[88] Therefore, intimidation statutes have been declared overbroad where they failed to limit the statute to threats of violence[89] and where they failed to require that the threat reasonably tend to produce actual fear in the victim.[90]

a. *Statutory exceptions for "constitutionally protected activity" and "legitimate purpose."* Turning to the statutes in question, several states have attempted to avoid an overbreadth challenge by expressly exempting "constitutionally protected activity" or by expressly limiting illegality to conduct that "serves no legitimate purpose." It is questionable whether these phrases cure the overbreadth problem. Although courts traditionally strive to construe statutes so as to avoid overbreadth,[91] two Florida lower courts have ruled that the Florida

---

Beckerman v. City of Tupelo, 664 F.2d 502, 508 (5th Cir. 1981) ("disorderly conduct" as used in parade licensing statute and defined in state penal code was overbroad).

[86] Watts v. United States, 394 U.S. 705, 707 (1969) (true threats are not constitutionally protected speech); Wurtz v. Risley, 719 F.2d 1438, 1441 (9th Cir. 1983) ("[T]hreats have traditionally been punishable without violation of the first amendment."); United States v. Mitchell, 463 F.2d 187, 191 (8th Cir. 1972) (violent threats are devoid of constitutional protection), *cert. denied*, 410 U.S. 969 (1973).

[87] *Watts*, 394 U.S. at 707 (construing presidential threat statute).

[88] *Wurtz*, 719 F.2d at 1441; *see also* United States v. Orozco-Santillan, 903 F.2d 1262, 1265–66 (9th Cir. 1990) ("A 'true' threat, where a reasonable person would foresee that the listener will believe he will be subjected to physical violence upon his person, is unprotected by the first amendment.").

[89] *Wurtz*, 719 F.2d at 1442 (intimidation statute criminalized any threat to commit any crime, serious or not); Landry v. Daley, 280 F. Supp. 938, 964 (N.D. Ill. 1968) (per three-judge district court) (same), *rev'd on other grounds*, 401 U.S. 77 (1971).

[90] *Wurtz*, 719 F.2d at 1440–41.

[91] *See, e.g.*, Arnett v. Kennedy, 416 U.S. 134, 162–63 (1974) (construing "such cause as will promote the efficiency of the service" to exclude constitutionally protected conduct); Dunkel v. Elkins, 325 F. Supp. 1235, 1241 (D. Md. 1971) (construing "no

antistalking law is nonetheless overbroad. These courts found that, because these provisions were included without definition, "the vagueness of the statute merges with its overbreadth and violates both constitutional precepts."[92]

In any event, several antistalking laws contain no exception for "constitutionally protected activity" or conduct undertaken for a "legitimate purpose."[93] These statutes, therefore, are much more likely to be declared overbroad. For example, in *United States v. Smith*,[94] the District of Columbia's temporary antistalking law, which contains no such exception,[95] was declared unconstitutionally overbroad because "it could, without any definite guidelines, prohibit constitutionally protected activity" such as right-to-life protest tactics.[96] Although the District of Columbia statute would nevertheless have been unconstitutionally vague, it is unlikely that the *Smith* Court would have reached the overbreadth question if a constitutional exception had been included.

b. *A "credible threat" of serious violence.* A better defense against overbreadth is the requirement, found in many antistalking laws, that there be a credible threat of serious violence. Usually, this is measured by a reasonable person standard. Limiting a statute to cover such threats helps ensure that only threats that are constitutionally unprotected are proscribed. However, several statutes do not contain a credible threat requirement,[97] and two others do not define

---

lawful business to pursue" language in statute as exclusion of constitutionally protected conduct).

[92] Florida v. Kahles, No. 92-022819 MM10A (Fla. Cir. Ct. Mar. 10, 1993); *see also* Florida v. Wallace, No. 93-087 CF (Fla. Cir. Ct. May 19, 1993) (both clauses vague *and* overbroad).

[93] *See* ALA. CODE § 13A-6-94 (1992); COLO. REV. STAT. § 18-9-111 (1992); 1992 D.C. Stat. 269; IOWA CODE § 708.11 (1992); MASS. GEN. L. ch. 265, § 43 (1993); N.Y. PENAL LAW § 240.25 (Consol. 1993); OHIO REV. CODE ANN. § 2903.211 (Baldwin 1993); 1993 Pa. Laws 28; TENN. CODE ANN. § 39-17-315 (1992); VA. CODE ANN. § 18.2-60.3 (Michie 1993); WASH. REV. CODE ANN. § 9A.46.110 (1992); WIS. STAT. § 947.013 (1991-92). *See also* 720 ILL. ANN. STAT. 5/12-7.3 (Smith-Hurd 1993) (exemption only for lawful picketing).

[94] No. M-6400-93 (D.C. Super. Ct. Sept. 9, 1993).

[95] The law prohibits "conduct with the intent to cause emotional distress to another person or places another person in reasonable fear of death or bodily injury by willfully, maliciously, and repeatedly following and harassing that person . . . ." 1992 D.C. Stat. 269.

[96] United States v. Smith, No. M-6400-93 (D.C. Super. Ct. Sept. 9, 1993).

[97] ARIZ. REV. STAT. ANN. § 13-2921 (1992); KY. REV. STAT. ANN. § 508.130 (Baldwin 1992); N.Y. PENAL LAW § 240.25 (Consol. 1993); 1993 N.D. Laws 120; VA. CODE ANN. § 18.2-60.3 (Michie 1993); WASH. REV. CODE § 9A.46.110 (1992); WIS. STAT. § 947.013 (1991-92) (no threat for misdemeanor harassment). Many statutes prohibit

"credible threat" in a way that requires "a reasonable tendency to produce in the victim a fear that the threat will be carried out."[98] Without a carefully defined credible threat element, and without a general exclusion of constitutionally protected conduct, these statutes will probably be found substantially overbroad.

c. *Special overbreadth problems.* Even those statutes that exempt constitutionally protected activity and provide for a credible threat may well be overbroad due to loose drafting of the constitutional activity exception. As set forth above, the California law, which provides the model for many statutes, contains an exception for constitutionally protected activity. However, that exception is part of the definition of "course of conduct," which in turn is part of the definition of "harass." There is no exception applicable to "following." This may be fatal, depending on whether a court will apply "overbreadth" analysis to an issue involving liberty of movement, which only indirectly implicates the First Amendment.[99]

First, following often occurs during an organized protest, usually the epitome of First Amendment activity. The California statute, however, makes it a crime for counterprotestors, who often hurl threats of violence, to follow marchers down a public street.[100] Following, moreover, is often used in the exercise of the First Amendment right to receive and gather information.[101]

---

following, harassment, *or* a credible threat. *See* DEL. CODE ANN. tit. 11, § 1312A (1992); 1992 Miss. Laws 532; N.J. REV. STAT. § 2C:12-10 (1992); S.D. CODIFIED LAWS ANN. § 22-19A-1 (1992) (amended to change "and" to "or"); *see also* 1992 D.C. Stat. 269 (conduct with intent to cause emotional distress or fear of death or bodily injury); 1993 Pa. Laws 28 (intent to cause fear of bodily injury *or* substantial emotional distress).

[98] ARK. CODE ANN. § 5-71-229 (Michie 1993) ("terroristic threat" undefined); MASS. GEN. L. ch. 265, § 43 (1993) ("threat" undefined); *see also* Wurtz v. Risley, 719 F.2d 1438, 1441 (9th Cir. 1983).

[99] *See* Lutz v. City of York, 899 F.2d 155, 270–71 (3d Cir. 1990) (holding that overbreadth does not apply to liberty of intrastate travel, which is a Fourteenth Amendment concept). *But see* Stotland v. Pennsylvania, 398 U.S. 916, 921 (1970) (Douglas, J., dissenting from per curiam dismissal) (noting that "one's constitutional right to freedom of movement . . . of course is essential to the exercise of First Amendment rights"). It is unclear whether the Supreme Court would carefully parse the allegations of First and Fourteenth Amendment overbreadth, or whether, once significant First Amendment concerns were raised, pendent liberty of movement overbreadth questions would be considered.

[100] Also disturbing is the arrest, noted *supra* note 57 and accompanying text, of Operation Rescue volunteers who followed someone in order to publicly chastise her for performing abortions.

[101] For the proposition that the First Amendment encompasses the right to receive information, see Renne v. Geary, 111 S.Ct. 2331, 2343–44 (1991) (White, J., dissenting) (citing Virginia Pharmacy Bd. v. Virginia Citizens Consumer Council, Inc., 425

Thus, the following "legitimate" conduct may be subject to arrest under the California law:

- Three days after a face to face confrontation at an abortion clinic in which violent threats are exchanged on all sides, prochoice volunteers follow Operation Rescue volunteers with the dual intent to ensure that clinic workers are not harassed and to annoy the right-to-life group;[102]
- A friend tells Alice that she saw Betty and Alice's husband, Jack, coming out of a motel. For several days thereafter, Alice follows Betty to make sure Betty is not having an affair with Jack, but also making sure that both know they are being followed. When confronted by Betty, Alice screams, "I know you're having an affair with my husband! He won't think you're so pretty after I throw acid in your face!"
- After an altercation between a photojournalist and a Hollywood star, in which threats are exchanged, the photojournalist follows the star even more often than before.

Under any of these situations, the state can legitimately proscribe the "threat" but probably cannot prohibit the "malicious" and "willful" following without infringing on First Amendment values.[103]

U.S. 748, 756–57 (1976)); Kleindienst v. Mandel, 408 U.S. 753, 762–63 (1972); Red Lion Broadcasting Co. v. FCC, 395 U.S. 367, 390 (1969); Stanley v. Georgia, 394 U.S. 557, 564 (1969). For the right to gather information, see Richmond Newspapers, Inc. v. Virginia, 448 U.S. 555, 586–87 (1980); Note, *The Rights of the Public and the Press to Gather Information*, 87 HARV. L. REV. 1505 (1974).

[102] As noted *supra* note 59 and accompanying text, prochoice volunteers were arrested in Minnesota on a stalking charge after engaging in this kind of surveillance.

[103] The requirement that the defendant act "willfully" and "maliciously" does not save the statute from overbreadth. "The addition of 'willful' to the statutory prohibition . . . adds no greater precision, since this element of intent is not proved separately but [is] inferred from the conduct constituting the violation." Reynolds v. Tennessee, 414 U.S. 1163, 1167 (1974) (Douglas, J., dissenting from denial of certiorari). Courts ordinarily do not give substantial weight to intent requirements in the overbreadth context unless the statute provides for "specific intent." *See, e.g.*, Von Lusch v. State, 387 A.2d 306, 310 (Md. 1978); City of Takoma v. Luvene, 827 P.2d 1374, 1383 (Wash. 1992). Neither "willful" nor "malicious" are equated with specific intent under California law. *See* People v. Sears, 401 P.2d 938, 943 (Cal. 1965), *overruled on other grounds by* People v. Cahill, 853 P.2d 1037 (Cal. 1993); People v. Mooney, 59 P. 761, 762 (Cal. 1899) ("The words 'willfully, unlawfully, feloniously, and maliciously' . . . import only that criminal intent which is a necessary part of every felony or crime."); People v. Glover, 285 Cal. Rptr. 362, 366 (Cal. Ct. App. 1991) (stating that neither "willful" nor "malicious" in a criminal statute establish a specific intent requirement). Further, the term "maliciously" includes merely the "wish to vex, annoy, or injure another." BLACK'S LAW DICTIONARY 958 (6th ed. 1990). Given that much vexing and annoying activity is protected by the First Amendment, a generalized "malicious intent" requirement does not sufficiently narrow the sweep of California-type antistalking statutes.

More importantly, any threat can operate to *convert* constitutionally protected following into an element of the crime, so long as the threat and the following had a continuity of purpose. For example, assume that gay rights activists follow a local community leader (seen as anti-gay) to his home, loudly and vexatiously disturbing his peace and causing him emotional distress. The next day they follow him back to his city office, seriously annoying him. On day three, he gets into a shouting match with the group, at which time one protestor yells, "Maybe you should feel what it's like to get beat up for no reason. Maybe we should come over tomorrow night and show you!" Without the threat, the activists' conduct was clearly constitutionally protected. Under the California statute, however, the addition of the threat operates to transform all previous constitutionally protected following into elements of the crime of stalking. Accordingly, this and similar antistalking statutes sweep much too broadly into protected areas.

To summarize this discussion of overbreadth, existing antistalking laws are or could be "substantially overbroad" because they (1) fail to exempt a large amount of constitutionally protected activity; (2) fail to limit the statute to "fighting words" by requiring a credible threat of serious violence, as defined by a reasonable person; or (3) fail to exempt constitutionally protected "following." Nevertheless, the most significant constitutional defect of these statutes rests not in their overbreadth, but in their vagueness.

2. Vagueness

The Fifth and Fourteenth Amendment Due Process Clauses require "that a statute may not forbid conduct in terms so vague that people of common intelligence would be relegated to differing guesses about its meaning . . . ."[104] Thus, courts insist "that laws give the person of ordinary intelligence a reasonable opportunity to know what is prohibited, so that he may act accordingly."[105]

> The requirement that government articulate its aims with a reasonable degree of clarity ensures that state power will be

---

[104] Schad v. Arizona, 111 S. Ct. 2491, 2497 (1991) (citing Lanzetta v. New Jersey, 306 U.S. 451, 453 (1939)).

[105] Grayned v. City of Rockford, 408 U.S. 104, 108 (1972).

> exercised only on behalf of policies reflecting an authoritative choice among competing social values, reduces the danger of caprice and discrimination in the administration of the laws, enables individuals to conform their conduct to the requirements of law, and permits meaningful judicial review.[106]

The elimination of guesswork among citizens and reducing "caprice and discrimination" by law enforcement officials are considered the principal aims of the vagueness doctrine.[107]

The constitutional requirement of clarity "applies with particular force in review of laws dealing with speech."[108] As in overbreadth analysis, for example, statutes alleged to inhibit First Amendment interests may be challenged for facial invalidity on the basis of vagueness.[109] The importance of free speech also heightens the level of scrutiny for vagueness: "[b]ecause First Amendment freedoms need breathing space to survive, government may regulate in the area only with narrow specificity."[110] Accordingly, any substantial ambiguity will void a statute that applies significantly to speech.

As currently drafted, the vast majority of antistalking laws are facially invalid for vagueness. In *Commonwealth v. Camper*,[111] for example, a Virginia trial court invalidated that state's antistalking statute on overbreadth and vagueness grounds, concluding that, taken together, the undefined phrases "engage in conduct" and "intent to cause emotional distress" were unconstitutionally vague.[112] Likewise, in *State v. Tremmel*,[113] the court held that the phrases "substantial emotional distress," "serves no legitimate purpose," and "constitutionally protected activities" made the entire statute impermissibly vague and overbroad.[114] Another Florida court held that the undefined phrases "substantial emo-

---

[106] Roberts v. United States Jaycees, 468 U.S. 609, 629 (1984) (citing, inter alia, Kolender v. Lawson, 461 U.S. 352, 357–58 (1983)).

[107] Michael S. Kelley, *"Something Beyond": The Unconstitutional Vagueness of RICO's Pattern Requirement*, 40 CATH. U. L. REV. 331, 370–71 (1991).

[108] Hynes v. Mayor of Oradell, 425 U.S. 610, 620 (1976); *accord* Smith v. California, 361 U.S. 147, 151 (1959) ("[A] man may the less be required to act at his peril here, because the free dissemination of ideas may be the loser.").

[109] *See* Maynard v. Cartwright, 486 U.S. 356, 361 (1988).

[110] Keyishian v. Board of Regents, 385 U.S. 589, 603–04 (1967) (quoting N.A.A.C.P. v. Button, 371 U.S. 415, 432–33 (1963)).

[111] No. 93-2876 (Va. Gen. Dist. Ct. Apr. 9, 1993).

[112] *Id.* at 6–7. The court implied that, had the Virginia statute had a "credible threat" element like the California statute, the vagueness and overbreadth might have been cured. *Id.* at 7.

[113] No. 93-02769 (Fla. Cir. Ct. June 9, 1993).

[114] *Id.* at 3–4.

tional distress," "repeatedly," "course of conduct," and "serves no legitimate purpose" were vague and that, with respect to the term "constitutionally protected conduct," the "vagueness of the statute merges with its overbreadth and violates both constitutional precepts."[115] A third Florida trial court focused on "emotional distress," "serves no legitimate purpose," and "constitutionally protected activity" in declaring the statute vague.[116] Recently, the District of Columbia Superior Court ruled that the District's temporary antistalking law was overbroad and vague.[117] With respect to the vagueness challenge, the court held that the statute

> fails to define terms such as "emotional distress" and "harassing" with particularity even though such terms are essential to understanding what conduct the statute makes criminal. Under this version of the statute, law enforcement personnel are left to their own subjective interpretations as to what constitutes "emotional distress" or "harassing." What may constitute emotional distress or harassment to one person may be different to another person and could depend on the tolerance level of the persons enforcing the statute or the person to which such conduct is directed.[118]

Finally, according to press accounts, an Illinois court has held that the former Illinois antistalking statute was unconstitutionally vague under the state and federal constitutions because it failed to distinguish between legal and illegal conduct.[119]

Although some trial-level courts have upheld antistalking laws against vagueness challenges, it is likely that appellate courts will follow the example of the cases set forth above.[120] Applying normal vagueness principles, reasonable citizens and law en-

---

[115] State v. Kahles, No. 92-022819MM10A (Fla. Cir. Ct. Mar. 10, 1992).

[116] State v. Wallace, Nos. 93-087 CF, 93-104-31 MM (Fla. Cir. Ct. May 17, 1993). None of these cases distinguished between misdemeanor and felony stalking even though the latter includes a "credible threat" requirement; rather, they declared the entire statute vague.

[117] United States v. Smith, No. M-6400-93 (D.C. Super. Ct. Sept. 9, 1993).

[118] *Id.*

[119] *See Judge Rules Stalking Law Unconstitutional*, UPI, Aug. 26, 1993, *available in* LEXIS, Nexis Library, Wires File.

[120] *See supra* notes 111–119. The only decision requiring extended comment is State v. Slater, No. 92-15850 CFANO (Fla. Cir. Ct. Feb. 18, 1993). There, the court began with the premise that a statute is not impermissibly vague "'merely because prohibited conduct is described in general language.'" *Id.* (quoting Wilkerson v. State, 401 So. 2d 1110, 1112 (Fla. 1981)). *Wilkerson*, however, did not involve a statute prohibiting speech and did not refer to the special vagueness scrutiny to which such statutes are subject. *See* 481 So. 2d at 1111–12; *See also supra* notes 108–110 and accompanying text.

forcement authorities are left to guess at the meaning of anti-stalking laws, including California's. Specifically, the California law (1) fails to define following in a way that provides guidance to the average citizen or excludes constitutionally protected activity; (2) fails to require a temporal connection between the threat and the acts of following or harassment; (3) uses the term "harasses" to define "harasses"; (4) uses vague terms like "alarms" and "annoys" in the definition of "harasses"; (5) defines "harasses" as a course of conduct which causes "substantial emotional distress," which in turn is undefined;[121] (6) fails to define "constitutionally protected activity" and "legitimate purpose"; and (7) may require police officers to decide whether the suspect intends to carry out the threat.[122]

Of the defects listed, by far the most important is the exemption for constitutionally protected activity. Although courts often treat overbreadth and vagueness together, it is probably more analytically correct to say that statutes expressly exempting constitutionally protected activity "simply exchange[] overbreadth for vagueness."[123] According to Professor Laurence Tribe:

> The risk of introducing vagueness when attempting to [judicially] reconstruct statutes reveals a structural relationship of

---

Proceeding from its flawed premise, the *Slater* court then rejected the claim that "substantial emotional distress," as used in the definition of "harass," was vague:

> When read *in toto*, the statute relates the victim's substantial emotional distress caused by harassment to the fear arising from the perpetrator's acts. In a very real and common sense fashion, anyone subjected to the acts and conduct of another sufficient to cause them to suffer a "reasonable fear" will experience "substantial emotional distress . . . . [A] 'credible threat' sufficient to place a person in 'reasonable fear' will invariably cause them 'substantial emotional distress.'"

Id. This argument is specious. In the first place, the Florida law does not require that the credible threat "cause" reasonable fear; it requires merely the "intent" to cause reasonable fear. Of course, the court might have construed the Florida "credible threat" definition to contain a requirement of causation akin to California's, but the court did not make this explicit.

Such a construction, moreover, would not eliminate the more serious analytical problem. Both the California and Florida statutes appear to require a credible threat *and* repeated acts of harassment. The concepts are separate, though sometimes overlapping. For example, two credible threats would also constitute harassment, but two acts of harassment would not necessarily amount to a credible threat. The court's analysis, therefore, would only be correct if the statute said that the *separate* acts of harassment must each be sufficient to cause a reasonable person to fear death or great bodily injury. That statute, of course, would be constitutionally valid, but neither the Florida nor the California law even remotely fits that description.

[121] Particularly confusing is the definition of "harasses" as conduct which "alarms or annoys" while at the same time causing "substantial emotional distress."

[122] *See* Guy, *supra* note 1, at 1013 (noting the ambiguity).

[123] LAURENCE H. TRIBE, AMERICAN CONSTITUTIONAL LAW § 12-29, at 1031 (2d ed. 1988).

> general importance in the interplay of overbreadth and vagueness. This relationship is most sharply focused in a hypothetical statute: *"It shall be a crime to say anything in public unless the speech is protected by the first and fourteenth amendments."* This statute is guaranteed not to be overbroad since, by its terms, it literally forbids nothing that the Constitution protects. The statute is nonetheless patently vague . . . .[124]

The conclusion that such an exception introduces a substantial lack of clarity seems unassailable; we cannot expect our citizenry and law enforcement officers to be constitutional scholars.

The same conclusions apply to exceptions for conduct serving a "legitimate purpose," at least where, as here, free speech interests are implicated. In *Bolles v. Colorado*,[125] the Colorado Supreme Court rejected the lower court's attempt to save a harassment statute by

> engrafting onto it a statement that it applied only to conduct engaged in for "no legitimate purpose." We think even with such limitation the subsection of the statute in question remains facially invalid.
>
> We again point out that the subsection in question deals with speech. Adding a phrase "without any legitimate purpose" to that subsection injects a vagueness into the statute which cannot withstand First Amendment scrutiny. Such a "limiting construction" disguises the constitutional difficulties of the statute but does nothing to resolve them.[126]

---

[124] *Id.*

[125] 541 P.2d 80 (Colo. 1975).

[126] *Id.* at 83; *see also* Papachristou v. City of Jacksonville, 405 U.S. 156, 157 (1972) (invalidating for vagueness an ordinance prohibiting, inter alia, "wandering or strolling around from place to place without any lawful purpose"). This is not to say that such a phrase automatically renders the statute vague. *Cf.* Illinois v. Williams, 551 N.E.2d 631, 633 (Ill. 1990) (citing cases where "unlawful purpose" was not vague where construed to mean an intent to violate a criminal code). Such a construction could not easily be invoked with respect to antistalking laws, however, because the whole idea behind them is to prohibit conduct that is intended *not* to break existing criminal laws. Moreover, the special nature of the conduct to be prohibited indicates that a "legitimate purpose" clause will be found impermissibly vague. As one writer aptly noted:

> In western culture, the difference between harassment and courtship too often turns on subjective interpretations of otherwise constitutional conduct. In other words, the same conduct which causes "substantial emotional distress" on one occasion might serve a "legitimate purpose" on another. Telephone calls are the best example. The distinction between "harassment" and an attempt at "reconciliation" is entirely subjective.

Hugh L. Korner, *Criminal Law: 1992 Survey of Florida Law*, 17 NOVA L. REV. 189, 220 (1992). Given this fact, and given the substantial amount of expressive activity involved, such clauses render the statute just as vague as the constitutional exemptions.

Thus, although such phrasing probably eliminates overbreadth considerations, it magnifies the vagueness of the statute.

The Michigan statute presents a special problem because, as written, it may not require that the defendant intend to cause any harm to the victim. The statute provides that it is a crime to "willfully" initiate or continue contact with an individual without his consent if the conduct serves no legitimate purpose and causes emotional distress. There does not appear to be a requirement that the defendant intend to stalk or cause emotional distress.[127] If that is the case, the statute covers erotomaniacal stalking better than the California-type legislation, but its broader language also contributes to its unconstitutional vagueness.[128] Without some requirement of specific intent to cause severe emotional harm or achieve some other wrongful result, and without a useful definition of "emotional distress," both individuals and law enforcement authorities will find it extremely difficult to gauge when "willful" conduct violates the law.[129]

Numerous press accounts demonstrate that law enforcement authorities have too much discretion under these statutes. These reports describe instances in which individuals were charged

---

[127] This conclusion depends on the premise that a Michigan court would construe "willful" as not requiring a specific intent to cause emotional distress. This is an open question in that state. "The word 'willfulness' has been called a 'word of many meanings, depending upon the context in which it is used.'" People v. Culp, 310 N.W.2d 421, 423 (Mich. Ct. App. 1981) (quoting People v. Cook, 279 N.W.2d 579 (Mich. Ct. App. 1979)). Thus, it may or may not imply a requirement of specific intent, depending primarily on whether the term refers to the wilful doing of an act or the wilful achievement of a result. Thus, specific intent is required where the statute prohibits "wilfully endangering lives," *see* People v. Beaudin, 339 N.W.2d 461, 463–64 (Mich. 1983), but not required for the wilful doing of an act, such as assault. *See, e.g.*, People v. Gleisner, 320 N.W.2d 340, 341–42 (Mich. Ct. App. 1982). Here, the statute proscribes wilfully engaging in a "course of conduct" involving "unconsented contact" with another. It does not appear to require an intent to cause a result (i.e., emotional distress). This conclusion appears to be consistent with the intent of the Michigan legislature, which reportedly designed the statute "to encompass all harassment and following that causes the victim to have reasonable fear." Guy, *supra* note 1, at 1009 n.162 (citing statement of Michigan State Representative Perry Bullard).

[128] *See generally* Gilligan, *supra* note 1, at 315–17 (scienter requirement is "critical" to sustain stalking laws against challenges for vagueness). Gilligan cites Screws v. United States, 325 U.S. 91 (1945), for the proposition that a "wilfulness" requirement is not vague, Gilligan, *supra* note 1, at 316 nn.208, 211, but the Court reached that conclusion only after it construed the term to require proof of specific intent. *Screws*, 325 U.S. at 101–03.

[129] This is not to say, as Guy, *supra* note 1, suggests, that the "wilfulness" requirement requires no harmful intent at all, or that it creates a "strict liability offense." Guy, *supra* note 1, at 1009. Rather, the requirement establishes a mens rea consisting of a generalized bad intent, *see Culp*, 310 N.W.2d at 423 ("willfully" encompasses "evil intent," "bad purpose," or "guilty knowledge"), such as an intent to annoy, vex, or embarrass another.

based on extremely slim evidence of actual stalking or based on activity that appears to be protected by the First Amendment. Particularly disturbing are the arrests in different states of both right-to-life and prochoice volunteers who did nothing more than follow members of the "other side."[130] Similarly, arrests for staring at someone in a public meeting, knocking on a door and looking in a window, and monitoring how an employee uses sick leave show that the statutes can be construed to permit applications far beyond the probable intent of the legislature.

Thus, although it is probably impossible to eliminate all broad phrases from this kind of statute without robbing it of necessary flexibility, the aggregate effect of numerous ambiguous terms in these laws is to create unconstitutional vagueness.[131]

---

[130] One writer has said that "the [California] statute requires that the stalker make, rather than merely pose, a credible threat. Therefore, the statute seems to require a verbal or written threat, rather than mere threatening conduct." Guy, *supra* note 1, at 1001 n.91. Apparently, law enforcement officers interpret similar statutory language differently, arresting any person who follows another.

[131] Three law review notes have attempted to address the constitutionality of existing antistalking legislation, each concluding that many such laws are constitutionally valid. Gilligan, *supra* note 1, at 304–20; Guy, *supra* note 1, at 1010–21; Strikis, *supra* note 2. In light of the contrary conclusions expressed herein, these pieces deserve a careful response.

The three notes share several common features that undermine the conclusion that existing antistalking laws are constitutional. First, none of them reflect an awareness of the numerous trial-level decisions invalidating these laws. *See supra* notes 111–119 and accompanying text. Second, none of them take note of the numerous disturbing instances of arrest under these statutes for conduct that is either difficult to call "stalking" or that is clearly protected by the First Amendment. *See supra* notes 55–63 and accompanying text. Third, only Gilligan mentions (but does not address) Professor Tribe's compelling argument that an exception for constitutionally protected activity "exchang[es] overbreadth for vagueness." Gilligan, *supra* note 1, at 309 n.160. Fourth, all of them emphasize the "credible threat" requirement as saving the statutes; however, none mentions the constitutional difficulties presented by the fact that a substantial amount of constitutionally protected conduct may be converted into illegality by a subsequent threat or tainted by a prior threat. *See supra* notes 39, 104 and accompanying text.

Individually, the notes are subject to some small criticisms. Gilligan suggests that a "willfulness" scienter element helps a statute avoid constitutional infirmity. Gilligan, *supra* note 1, at 315–20. But that is true only if "willfulness" requires specific intent to cause a reasonable fear of substantial violence. *See supra* note 103. The note also assumes that a high scienter element will limit the discretion of law enforcement authorities. *Id.* at 319–20. That point is arguable even for the run-of-the-mill criminal statute. *Cf.* Reynolds v. Tennessee, 414 U.S. 1163, 1167 (1974) (Douglas, J., dissenting from denial of certiorari). Where, as here, the mens rea element may be inferred from pure expression unaccompanied by concrete acts, law enforcement discretion is extremely broad.

Guy expressly eschews any First Amendment analysis, focusing mainly on due process vagueness concerns. Guy, *supra* note 1, at 994 n.21. This is its principal shortcoming. *See supra* note 110 and accompanying text (discussing heightened standard of vagueness scrutiny where First Amendment values are at stake). The note also assumes that the phrases "constitutionally protected activity," "no legitimate

C. *"First Amendment Due Process"*

The term "First Amendment Due Process" is generally attributed to Professor Henry Monaghan, whose 1970 article of that title sought to describe the various procedural protections which give special protection to First Amendment concerns.[132] Because current antistalking laws encompass a significant amount of speech, states may be required to provide certain procedural guarantees not ordinarily required of a criminal statute. Even if not required, these safeguards could operate to save an otherwise overbroad or vague statute.

The law in this area is unsettled, and it is by no means clear that procedural guarantees are either required or that they will cure an otherwise facially defective statute. Furthermore, if the statutes are drawn more narrowly to eliminate the overbreadth and vagueness described above, these guarantees may no longer be necessary.

1. First And Fourth Amendment: Warrantless Arrests

From the victim's standpoint, one of the more beneficial provisions in antistalking legislation is the authorization of warrantless arrests if the police officer has probable cause to believe that stalking has been or is occurring. Unfortunately, a further consequence of the overbreadth and vagueness described above is that, by encompassing a substantial amount of conduct protected by the *First* Amendment, the statutes heighten the *Fourth* Amendment standard to be applied to provisions authorizing warrantless arrests. Such provisions in antistalking laws are probably unconstitutional, at least until the statutes are redrafted to reduce First Amendment concerns.

The Fourth Amendment provides:

---

purpose," "emotional distress," "substantial emotional distress," and "follow" pose no vagueness problems, either individually or in the aggregate.

Strikis operates on a relatively high level of generality and does not attempt to analyze particular statutes. Strikis, *supra* note 2. Nevertheless, it makes some troubling assumptions, including (1) that the state interest in prohibiting harassing conduct that amounts to "protected expression" may outweigh the individual's free speech interests in this area; (2) that prosecutors will decline to bring charges based on constitutionally protected activities; (3) that a requirement of recklessness is sufficient to satisfy vagueness standards; and (4) that "a generally recognized notion of 'stalking' is evolving." Strikis, *supra* note 2.

[132] Henry P. Monaghan, *First Amendment "Due Process"*, 83 HARV. L. REV. 518 (1970).

> The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated; and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.[133]

Without doubt, "the arrest of a person is 'quintessentially a seizure.'"[134] Ordinarily, an arrest can be made without a warrant if the police officer has probable cause to believe that the "suspect has violated or is violating the law."[135] Accordingly, a statutory authorization to allow warrantless arrests would not, under normal circumstances, violate the Fourth Amendment.[136]

However, that conclusion does not follow where substantial free speech interests are at stake. It is well settled that the Fourth Amendment must be applied with "scrupulous exactitude" where First Amendment values are involved.[137] Although the Supreme Court has not ruled directly on the application of the "scrupulous exactitude" standard to warrantless arrests,[138] it is generally assumed that arrests without a warrant are subject to a higher level of scrutiny when the First Amendment is implicated.[139] Lower courts have held that such warrantless arrests violate the First and Fourth Amendment in cases where obscenity is involved:

> [Appellant county solicitor general] initially contends that the warrantless arrests were proper because his officers viewed a violation of the Georgia obscenity statute in their

---

[133] U.S. CONST. amend. IV.

[134] Payton v. New York, 445 U.S. 573, 585 (1980) (quoting United States v. Watson, 423 U.S. 411, 428 (1976) (concurring opinion)).

[135] Criss v. City of Kent, 867 F.2d 259, 262 (6th Cir. 1988) (citing Michigan v. DeFillipo, 443 U.S. 31, 36 (1979)).

[136] *Cf.* United States v. Valenzuela-Bernal, 458 U.S. 858, 864 (1982) (discussing federal statute authorizing warrantless arrests of suspected illegal immigrants); *Payton*, 445 U.S. at 587 (warrantless arrests in public place are valid).

[137] Maryland v. Macon, 472 U.S. 463, 468 (1985); Zurcher v. Stanford Daily, 436 U.S. 547, 564 (1978); Stanford v. Texas, 379 U.S. 476, 485 (1965).

[138] *See Macon*, 472 U.S. at 471 (declining to decide issue).

[139] *See id.* at 473–74 (Brennan, J., dissenting) ("[A] warrantless arrest involves the same difficulties and poses the same risks as does a warrantless seizure of books, magazines, or films."); Henry P. Monaghan, *A Stealthy Encroachment: Obscenity and the Fourth Amendment Under Maryland v. Macon*, 36 AM. U. L. REV. 773, 796 (1987) (noting "the obvious prior restraint implications" of the warrantless arrest in *Macon*); Monaghan, *supra* note 132, at 538 ("Functionally, an arrest resembles a non-judicially imposed injunction against certain conduct; and where timing is important, as in many demonstrations, an arrest, like an ex parte injunction, can wholly frustrate the exercise of first amendment rights without any searching inquiry into the merits of the first amendment claim.").

> presence. He relies on Georgia Code § 27-207 which provides that "[a]n arrest for a crime may be made by an officer . . . without a warrant if the offense is committed in his presence." The problem with appellant's claim is that the ability to make a warrantless arrest for an offense committed in the officer's presence contemplates the officer's ability to determine that an offense has actually been committed. Appellant is attempting to apply a statute normally appropriate for the case of a fleeing robber to items presumptively protected by the First Amendment. Appellant is incorrect in his belief that he or his agents may properly make the initial determination concerning the obscenity of a publication and that he may make a warrantless arrest if he determines that the subject matter of a publication is obscene.[140]

Of course, obscenity differs from stalking, principally in its potential for immediate violence. For this reason alone, the Constitution would not prohibit all warrantless arrests in stalking situations. For example, a police officer is presumably capable of determining whether a "credible threat" has been made in his presence, and the exigencies of that circumstance could justify immediate arrest.

Nevertheless, as long as the statutes remain broadly worded, a warrantless arrest provision cannot survive the presumptively applicable "scrupulous exactitude" test. For example, the Florida statute authorizes the police officer to make a warrantless arrest of someone who vexatiously follows a prochoice rally down the street on two separate occasions (remember: misdemeanor stalking in Florida does not require a threat). Moreover, as in the case of an obscenity statute, the Florida law requires the police officer to make a determination as to whether constitutionally protected speech is involved. This extremely delicate constitutional analysis is compounded by the fact that the officer need not have witnessed the threat or the following or harassing conduct; he may rely instead on a third-party description of the conduct and then make the constitutional determination. Finally, since the "credible threat" necessary to establish felony stalking is tied to vague concepts, and since the threat could have occurred in the distant past, it is a close question as to whether a warrantless arrest would be appropriate even if there were substantial third-party evidence that such a threat had been made.

---

[140] Penthouse Int'l, Ltd. v. McAuliffe, 610 F.2d 1353, 1359 (5th Cir. 1980); *see also* Goodman v. City of Dallas, 73 F.R.D. 642, 646 (N.D. Tex. 1977) (there must be a prior *adversary* hearing before obscenity arrest can be made); Goodwin v. Morris, 318 F.Supp. 1325, 1326–27 (N.D. Ohio 1970) (same).

At least as construed in connection with the statutes' other constitutionally suspect provisions, warrantless arrest sections in antistalking laws probably violate the First and Fourth Amendments.[141] As set forth below, this defect can only be cured through a more narrow redefinition of the substantive crime or, perhaps, by enumerating specific dangerous instances in which a warrantless arrest is permissible.

2. Immediate Hearing

Whether or not a warrant issues, a person charged with stalking under a broadly worded statute may also be constitutionally entitled to an immediate judicial hearing if she claims that she was engaging in constitutionally protected conduct. Professor Monaghan has argued as follows:

> The other major teaching of the obscenity cases is that in the first amendment area judicial review must either precede final government action or expeditiously follow it. Both *Freedman* and *Teitel Film Corp. v. Cusak*, [sic] invalidated statutes which did not provide for immediate judicial review of the administrative determination. In part, this result seems predicated on the belief that delay in the availability of judicial relief differs only in degree, and sometimes not at all, from the complete absence of judicial review.[142]

Under this reasoning, a statute that potentially swept up constitutionally protected expression and movement along with unprotected speech and activity must provide a speedy judicial determination on the constitutional issues, particularly where the statute expressly exempts constitutionally protected activity and where other provisions (e.g., authorizing warrantless arrests or the denial of bail) weakened normal procedural protections.

Although the case law is sparse, a Ninth Circuit decision may

---

[141] Like the Florida statute, the New Hampshire law is defective, providing for a warrantless arrest when there is probable cause to believe that the offense occurred within six hours. 1993 N.H. Laws 173. However, Maryland provides for a warrantless arrest only where there is probable cause supported by credible evidence other than the statements of the "victim," combined with a reasonable belief of imminent bodily harm to the victim. MD. ANN. CODE art. 27 § 594B (1993). So limited, the Maryland warrantless arrest provision probably survives constitutional scrutiny.

[142] Monaghan, *supra* note 132 at 532 (footnotes omitted) (citing Freedman v. Maryland, 380 U.S. 51, 59–59 (1965) (requiring adversary hearing before final restraint of motion picture); Teitel Film Corp. v. Cusack, 390 U.S. 139 (1968) (invalidating ordinance which did not provide for speedy judicial determination of whether a film was constitutionally protected).

be instructive. In *Golden Eagle v. Johnson*,[143] the appellant was arrested for possession of peyote, which he claimed was used for religious purposes. In holding that the First Amendment did not require an adversary hearing *before* arrest, the Ninth Circuit first disagreed with the assertion that procedural safeguards developed for obscenity cases (such as a preseizure adversary hearing) automatically extend to other First Amendment areas:

> We believe, however, that the unique and sensitive procedures applicable to seizures of allegedly obscene material were designed more for the special purpose of protecting a right of the public than to protect the rights of the person from whom the allegedly obscene material was seized. Seizure of allegedly obscene material prior to any determination of obscenity by means of an adversary hearing will result inescapably in the suppression of large numbers of nonobscene books. The public has a First Amendment right to the unimpeded flow of nonobscene books.[144]

Finding no equivalent rights of the public in the case at hand, the court then noted that the California Penal Code already provided for "an expeditious hearing before a magistrate . . . without regard to whether the arrest was with or without a warrant . . . ."[145] This and other procedures, the court held, "afford assurance that a determination of the [constitutionally protected] good faith of a member of the Native American Church and a return of any sacred objects seized from him can be speedily resolved when the facts permit."[146] Accordingly, the court concluded that "[t]he process that provides these opportunities for reconciling an accused's First Amendment rights . . . should not be further refined" by requiring a pre-arrest adversary hearing or other procedures suggested by the appellant.[147]

Assuming that the *Golden Eagle* distinction based on the public's First Amendment rights is correct, it is reasonably likely that the Ninth Circuit would find such a right in the case of a person charged with stalking. Like obscenity, and unlike the religious use of peyote, stalking as currently defined may restrict the public's access to protected speech—namely, vigorous political and religious protest. It may also prevent large groups from engaging in protests, picketing, lawful surveillance, and other

---

[143] 493 F.2d 1179 (9th Cir. 1974), *cert. denied*, 419 U.S. 1105 (1975).
[144] *Id.* at 1184.
[145] *Id.* at 1185.
[146] *Id.*
[147] *Id.*

forms of protected speech and movement that border on harassment. In light of the nature of the crime of stalking, and in light of the *Golden Eagle* court's emphasis on the availability of immediate post-arrest judicial review, a court could hold that such review is required to save the statute from constitutional infirmity.

Whether or not an antistalking law should be amended to include a specific provision for immediate post-arrest review will depend on a number of factors, including (1) the scope of any overbreadth; (2) the degree of vagueness; (3) the existence of a clause exempting constitutionally protected activity; (4) procedural safeguards already contained in the state's penal code; and (5) the extent to which other provisions in the antistalking statute operate to weaken those safeguards. For example, an immediate hearing may be required in order to counteract the effect of a warrantless arrest.

3. Constitutionally Protected Activity as Evidence Supporting a Stalking Conviction

The Kentucky statute contains a hearing requirement and, though it does not require that the hearing be immediate, it does point to an important evidentiary issue: "If the defendant claims that he was engaged in constitutionally protected activity, the court shall determine the validity of that claim as a matter of law and, if found valid, shall exclude that activity from evidence."[148]

The question is whether language similar to that found in the Kentucky statute is required by the Constitution. Although no cases directly address this point, there are several indications in the case law that a heightened evidentiary standard is required, though exclusion is not mandated.

As a threshold matter, it is clear that evidentiary protections are included in the "procedural safeguards" which courts may require in the First Amendment context.[149] On the other hand, *per se* exclusion of evidence of constitutionally protected activity is probably not required by analogous cases. In *Dawson v.*

---

[148] KY. REV. STAT. ANN. § 508.130(2) (Baldwin 1992). North Dakota's antistalking law contains a similar provision. N.D. CENT. CODE § 12.1-17-07.1.

[149] *See, e.g.*, New York Times Co. v. Sullivan, 376 U.S. 254, 286–87 (1964) (requiring clear and convincing evidence that libel defendant acted with "actual malice").

Delaware,[150] the Court held that "the Constitution does not erect a *per se* barrier to the admission of evidence concerning one's beliefs and associations at a sentencing simply because those beliefs and associations are protected by the First Amendment."[151] Nevertheless, the Court subjected the admission of such evidence to a searching relevance inquiry which, as the dissent noted, allowed the defendant to present evidence of constitutionally protected activity while making it difficult for the prosecution to present similar evidence.[152] Commenting on the case, Justice Stevens has said that "the trial court erred by allowing the jury to base a death sentence, in part, on the fact that Dawson had engaged in constitutionally protected conduct."[153] Thus, *Dawson* may stand for the proposition that, though not subject to automatic exclusion, evidence of constitutionally protected conduct is subject to a higher constitutional scrutiny.

This conclusion finds support in cases decided under the *Noerr-Pennington* doctrine, which, based on First Amendment concerns, immunizes from criminal or civil antitrust liability most attempts to influence political bodies in a way that harms competitors or competition.[154] In a footnote to the *Pennington* decision, the Court stated in dicta that evidence of such protected activity could be introduced, if relevant and not unduly prejudicial, to help show the "purpose and character" of other, unprotected conduct.[155] However, some courts and commentators have determined that the First Amendment underpinnings of the *Noerr-Pennington* doctrine require that such evidence be inadmissible[156] or, at least, "presumptively prejudicial."[157] In light of

---

[150] 112 S.Ct. 1093 (1992).

[151] *Id.* at 1097.

[152] *Id.* at 1097–99; *id.* at 1102 (Thomas, J., dissenting).

[153] John P. Stevens, *Address: The Freedom of Speech*, 102 YALE L.J. 1293, 1294 (1993).

[154] *See* Eastern R.R Presidents Conference v. Noerr Motor Freight, Inc., 365 U.S. 127 (1961); United Mine Workers of Am. v. Pennington, 381 U.S. 657 (1965).

[155] *Pennington*, 381 U.S. at 670 n.3.

[156] *See* Colorado Interstate Gas Co. v. Natural Gas Pipeline Co., 661 F. Supp. 1448, 1481 (D. Wyo. 1987) (*Noerr* "bars" evidence of a threat to intervene in FERC proceeding), *rev'd in part and aff'd in part*, 885 F. 2d 683 (10th Cir. 1989); United States v. Johns-Manville Corp., 259 F. Supp. 440, 453 (E.D. Pa. 1966) (inferring predatory intent from protected activity would be illogical and would infringe on First Amendment rights); Daniel Fischel, *Antitrust Liability for Attempts to Influence Government Action: The Basis and Limits of the Noerr-Pennington Doctrine*, 45 U. CHI. L. REV. 80, 121 (1977).

[157] United States Football League v. National Football League, 634 F. Supp. 1155, 1181 (S.D.N.Y. 1986) (presumptively prejudicial), *on appeal from jury verdict*, 842 F.2d 1335, 1373–75 (2d Cir. 1988); PHILLIP AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW ¶ 203.7 (1992).

the Court's reasoning in *Dawson*, a presumption of prejudice is probably the most that the courts can require.

The fact that a court may be required to exclude evidence of protected conduct does not, of course, mean that an antistalking statute must contain an express directive on the subject. However, two characteristics peculiar to existing antistalking statutes suggest that a "Kentucky" provision may be constitutionally required. First, many statutes expressly exclude "constitutionally protected activity" from the definition of "harasses," indicating perhaps that the element of harassment *must* be proved without reference to such protected activity. Second, as currently drafted, antistalking statutes create the danger that the jury will find the existence of one element of the crime (the act of following) based solely on constitutionally protected activity *or* will find that a credible threat operated to transform such protected conduct into unprotected illegal stalking. Arguably, that evidence is "presumptively prejudicial" absent a prior judicial determination that there is a constitutionally sufficient "nexus" between the threat and the act of following.

As with the requirement of an immediate hearing, the need for a "Kentucky" provision declines as overbreadth and vagueness concerns are addressed. Nevertheless, such a provision, or one that creates a presumption of prejudice, increases the likelihood that the law will be upheld.

D. *Miscellaneous Constitutional Infirmities*

In addition to the broader concerns outlined above, several statutes have constitutional defects particular to themselves.

1. Content-Based Distinctions and Exemptions for Labor Picketing

Several antistalking statutes, including California's, prohibit one kind of stalking while permitting another. This decision to exempt the latter but not the former is based on the message of the speaker. This sort of content-based distinction probably violates both the First Amendment and the Equal Protection Clause of the Fourteenth Amendment.

a. *Content-based distinctions and the First Amendment.* In *R.A.V. v. City of St. Paul*[158] the Supreme Court held that a city ordi-

nance prohibiting bias-motivated disorderly conduct was unconstitutional. The statute provided as follows: "Whoever places on public or private property a symbol, object, appellation, characterization or graffiti . . . which one knows or has reasonable grounds to know arouses anger, alarm or resentment in others on the basis of race, color, creed, religion or gender commits disorderly conduct . . . ." The petitioner's claim that the statute was overbroad was rejected by the Minnesota Supreme Court, which construed the statute as being limited to "fighting words" not protected by the First Amendment.

The United States Supreme Court, unanimously finding the statute to be unconstitutional, split in its reasoning. The majority opinion, written by Justice Scalia, ignored the overbreadth claim and instead held that the statute, by prohibiting a subset of fighting words based on the content of the speaker's message, was facially unconstitutional:

> [T]he ordinance applies only to "fighting words" that insult, or provoke violence, "on the basis of race, color, creed, religion or gender. Displays containing abusive invective, no matter how vicious or severe, are permissible unless they are addressed to one of the specified disfavored topics . . . . The First Amendment does not permit St. Paul to impose special prohibitions on those speakers who express views on disfavored subjects.[159]

The concurring opinions would have held the statute overbroad.[160]

Although the jurisprudential dust has not settled around the *R.A.V.* decision, the California statute and others like it are probably invalid under the new decision. The statute provides that it "shall not apply to conduct which occurs during labor picketing." In other words, California has forbidden stalking *except* when the stalkers are picketing laborers. Thus, the same "fighting words" and stalking activity that would land an employer in jail during an ordinary labor dispute have been expressly sanctioned in the case of a union member during a labor picket.[161] This is arguably content-based discrimination forbidden by the *R.A.V.* decision because the classification appears to be related to "the sovereign's agreement with what a speaker may intend

---

[158] 112 S. Ct. 2538 (1992).

[159] *Id.* at 2547.

[160] *Id.* at 2550–71.

[161] Of course, it is conceivable that an employer could invoke this section if he stalked a union member who was picketing, but it strains credence to say that this exemption evenhandedly protects workers and employers. Even if it did, it is still a content-based distinction between employer-employee stalking and all other types.

to say"[162] and because the state appears to be "licens[ing] one side of a debate to fight freestyle, while requiring the other to follow Marquis of Queensbury Rules."[163] Accordingly, statutes that contain this kind of exception should be redrafted to conform to the reasoning of the *R.A.V.* decision.

b. *Equal protection.* Whether or not *R.A.V.* applies, it is relatively well settled that irrational underinclusiveness violates the Equal Protection Clause of the Fourteenth Amendment. In *Carey v. Brown*,[164] the Supreme Court struck down an ordinance which prohibited picketing in residential neighborhoods, except that which occurred during a labor dispute.[165] Noting that "[t]he permissibility of residential picketing . . . is thus dependent solely on the nature of the message being conveyed," the Court held that the Equal Protection Clause, as informed by First Amendment considerations, prevented the city from discriminating among speech-related activities in a public forum.[166]

That the conduct occurred in a public forum was important to the decision because it established that the ordinance impinged on "a fundamental right protected by the Constitution."[167] Under that standard, "the Equal Protection Clause mandates that the legislation be finely tailored to serve substantial state interests, and the justifications . . . must be carefully scrutinized."[168] Where such a fundamental right is not involved, the legislation "need only rationally further a legitimate state purpose."[169] Antistalking statutes would probably be judged under the higher standard, either because the statutes restrict a significant amount of con-

---

[162] *R.A.V.*, 112 S. Ct. at 2547 (quoting Metromedia, Inc. v. City of San Diego, 453 U.S. 490, 555 (1981)) (Stevens, J., dissenting in part) (citation omitted).

[163] *R.A.V.*, 112 S. Ct. at 2548. Conceivably, the California statute could be distinguished on the theory that rather than proscribe certain fighting words based on content, the statute creates an *exception* based on content. However, the distinction would be extremely difficult to maintain. To put the issue in perspective, California would be forbidden under *R.A.V.* from proscribing only those fighting words "directed against the Republican Party." Presumably, it also could not enact a statute proscribing all fighting words "*except* those directed at the Democratic Party." In either situation, one form of fighting words will be advanced over another form based on content.

[164] 447 U.S. 455 (1980).

[165] *Id.* at 459–71.

[166] *Id.* at 461; *see also* Police Dep't of Chicago v. Mosley, 408 U.S. 92 (1972) (invalidating ordinance prohibiting picketing in front of school except those for labor disputes); Blair v. Shanahan, 775 F. Supp. 1315, 1325 (N.D. Cal. 1991) (antibegging statute unconstitutionally discriminates based on the content of the communication).

[167] Perry Educ. Ass'n v. Perry Local Educators' Ass'n, 460 U.S. 37, 54 (1983) (explaining *Carey*).

[168] *Carey*, 447 U.S. at 461–62.

[169] *Perry*, 460 U.S. at 54.

duct in public forums (e.g., city streets) or because they directly impinge on free speech and fundamental liberties.[170]

Under either standard, however, the labor picketing exemption would violate the Equal Protection Clause because it does not even rationally further a legitimate purpose. The broad goal underlying antistalking legislation (i.e., preventing threatening behavior) is not served in any way by the exemption, which in fact thwarts that goal. Nor can the goal of protecting labor protest justify the content-based distinction because the exemption ostensibly protects labor stalking, not legitimate protest.[171] In fact, the only conceivable justification for the exemptions—that they are necessary because the statutes threaten the legitimate protest of favored groups—cannot be made without admitting that the laws are fatally overbroad. Thus, the exception is wholly irrational.

2. Shifting the Burden of Proof

The Michigan statute contains a unique provision that operates to shift the burden of proof with respect to causation onto the defendant.

> In a prosecution for a violation of this section, evidence that the defendant continued to engage in a course of conduct involving repeated unconsented contact with the victim after having been requested by the victim to discontinue the same or a different form of unconsented contact, and to refrain from any further unconsented contact with the victim, shall give rise to a rebuttable presumption that the continuation of

---

[170] This is apparently the case with obscenity statutes, which, like antistalking legislation, implicate First Amendment rights without reference to a "forum" analysis. *See* Kucharek v. Hanaway, 902 F.2d 513, 517 (7th Cir. 1990) (in dicta, citing *Carey* as prohibiting a statute "forbidding the sale of obscene books unless the book confined its depictions to procreative sex").

[171] Nor can the exemption be justified as a rational legislative finding that threats of physical violence during labor disputes are always harmless. Stalking in the employer-employee context is a significant problem, *see* Guy, *supra* note 1, at 995 n.30 (former employers often stalked), and hardcore stalking activity appears to be common in the kind of strike disputes the California legislation would exempt. *See, e.g., Freedom of Access to Clinic Entrances Act of 1993: Hearings on S. 636 Before the Senate Labor and Human Resources Committee*, 103rd Cong., 1st Sess. (1993) (question from Senator Judd Gregg (R-NH)) ("I've got a list here of over . . . 5,000 violent acts in the area of labor strikes that occurred . . . since 1975."); Stephen Franklin, *Bitter Road to Reconciliation*, CHI. TRIB., May 16, 1993, at C1 (describing Greyhound strike) ("[S]niper shots hit passenger-filled buses across the country, drivers were shot in Tennessee and Florida . . . . Greyhound received more than 100 bomb threats in the early days of the strike."); Lawrence H. Kaufman, Abstract, *FBI Probes Derailments*

> the course of conduct caused the victim to feel terrorized, frightened, intimidated, threatened, harassed, or molested.[172]

This provision probably violates the due process clause, which prohibits irrational presumptions created by a criminal statute.[173]

A statutory presumption is irrational "unless it can at least be said with substantial assurance that the presumed fact is more likely than not to flow from the proved fact on which it is made to depend."[174] Here, the fact that a person continued a particular contact after being asked to stop does not reasonably lead to the conclusion that the victim suffered significant mental suffering unless that vague concept is taken to mean mere annoyance or frustration, in which case the statute is plainly overbroad and vague. Since there is no substantial assurance that significant emotional harm is "more likely than not" to flow from proof of unconsented contact, the presumption is irrational and, hence, a violation of due process.

3. Pretrial Detention Without Bail

Pretrial detention without bail is a mechanism developed to ensure the safety of stalking victims and other members of the community. This mechanism itself raises constitutional concerns separate from those set forth above.

Legislators and law enforcement agencies justify pretrial detention without bail upon myriad grounds, including risk of flight, intimidation of witnesses, and danger to the victims of the crime and the community. Stalking particularly raises the last concern.[175] The media portrays and the public perceives the stalker as an obsessed pursuer who abides by no legal restraint and must be completely disabled in order to protect the victim. No-bail provisions are in part a legislative response to those fears.

However, "[i]n our society liberty is the norm, and detention

---

in Coal Strike Area, J. OF COM., July 27, 1993, at A1, *available in* LEXIS, Nexis library (reporting that FBI authorities believe vandals linked to United Mine Worker strikers caused derailment of train).

[172] MICH. COMP. LAWS § 750.411h (1992).

[173] Leary v. United States, 395 U.S. 6, 36 (1969); Tot v. United States, 319 U.S. 463, 467–68 (1943).

[174] *Leary*, 395 U.S. at 36.

[175] The very nature of the crime of stalking is a repetition of contacts with the victim. Victims sometimes obtain restraining orders against their stalkers, but the stalkers often violate these orders with relative impunity. *See supra* notes 27–30 and accompanying text.

prior to trial or without trial is the carefully limited exception."[176] To allow pretrial detention without bail gives the state the power to detain a presumptively innocent person in jail for an indeterminate amount of time based on predictions of future criminal conduct.[177] The Constitution contains two potential limitations on the ability of legislatures to take such a drastic step: the bail clause of the Eighth Amendment and the due process clause of the Fourteenth Amendment.

a. *The Bail clause of the Eighth Amendment.* The relevant portion of the Eighth Amendment reads: "Excessive bail shall not be required . . . ."[178] The first question one must ask, therefore, is whether that clause prohibits a provision precluding any bail at all where the primary justification is the potential danger to the public.[179]

In *United States v. Salerno*,[180] the Supreme Court addressed a similar problem with respect to the Bail Reform Act of 1984, which provides for pretrial detention without bail for those who are charged with capital offenses, those who pose a risk of flight, and those who pose a danger to the safety of the community.[181] The first two categories fall into the traditional reasons for denying bail. But "danger to the community" presents a wholly different rationale. In effect, the defendant could be charged with one crime but could be held before trial by the government for an entirely different reason.

Salerno was a facial attack on the Bail Reform Act based on the Fifth and Eighth Amendments. In its decision, the Court minimized the importance of the Eighth Amendment in the challenge to the federal no-bail provision. The Court, quoting the

---

[176] United States v. Salerno, 481 U.S. 739, 755 (1987).

[177] *See id.* at 760 (Marshall, J., dissenting).

[178] U.S. CONST. amend. VIII. The courts have applied the Bail Clause of the Eighth Amendment to the states, under the Fourteenth Amendment. *See* Baker v. McCollan, 443 U.S. 137, 144 n.3 (1979) (quoting Schilb v. Kuebel, 404 U.S. 357, 365 (1971); *see also* Faheem-El v. Klincar, 841 F.2d 712, 718 n.8 (7th Cir. 1988).

[179] Historically, bail requirements were used to ensure the defendant's appearance at trial and to aid and protect the administration of justice, e.g., to protect witnesses from intimidation. Stack v. Boyle, 342 U.S. 1, 4–5 (1951); *see also* United States v. Simpkins, 826 F.2d 94, 97–98 (D.C. Cir. 1987). Bail must be set at an amount reasonably calculated to ensure the defendant's appearance and bail set at figures in excess of that amount violates the Eighth Amendment. *Stack v. Boyle*, 342 U.S. at 6. Thus, until recently it was understood that the Bail Clause dictated that bail be exclusively directed at ensuring the appearance of the defendant and protecting the judicial process. However, bail provisions have increasingly been used to advance other societal goals.

[180] 481 U.S. 739 (1987).

[181] 18 U.S.C.A. § 3141, 3142 (e) (West 1993).

observation in *Carlson v. Landon* that "the very language of the [Eighth] Amendment fails to say all arrests must be bailable," held that the Bail Clause did not apply to complete denial of bail.[182]

> Nothing in the text of the Bail Clause limits permissible Government considerations solely to questions of flight. The only arguable substantive limitation of the Bail Clause is that *the Government's proposed conditions of release or detention not be "excessive" in light of the perceived evil . . . .* We believe that when Congress has mandated detention on the basis of a compelling interest other than prevention of flight . . . the Eighth Amendment does not require release on bail.[183]

The Court's literal reading of the Bail Clause left little substantive power in the Eighth Amendment to prohibit a no-bail provision, certainly no more than that provided by the Due Process Clause, discussed below.

b. *Due Process under the Fourteenth Amendment.* The *Salerno* Court clearly thought that the more important constitutional protection for those detained without bail is provided by the Due Process Clause of the Fourteenth Amendment. Due process has been defined as having two components, encompassing both the "legitimacy of the government interest, i.e., substantive due process; and procedural due process."[184] The test for substantive due process appears to be two-pronged. The *Salerno* Court stated:

> Unless Congress expressly intended to impose punitive restrictions, the punitive/regulatory distinction turns on "whether an alternative purpose to which [the restriction] may rationally be connected is assignable for it, and whether it appears excessive in relation to the alternative purpose assigned [to it]."[185]

---

[182] 481 U.S. at 754 (quoting *Carlson*, 342 U.S. 524, 545–46 (1952)). But this holding would seem to violate the spirit, if not the letter of the Bail Clause. Justice Marshall argued vociferously in dissent that a no-bail provision was tantamount to allowing a judge to set bail at an infinite amount, completely out of proportion to the amount necessary to ensure the detainee's presence at trial. *Id.* at 760–61 (Marshall, J., dissenting). To the dissent, once the bail inquiry strayed from the question of risk of flight, the statute began to infringe on the presumption of innocence. In an earlier case, Justice Black called a similarly limiting interpretation of the Bail Clause a "weird, devitalizing interpretation." Carlson v. Landon, 342 U.S. at 556 (Black, J., dissenting).

[183] 481 U.S. at 754–55 (emphasis added).

[184] United States v. Simpkins, 826 F.2d 94, 97 (D.C. Cir. 1987).

[185] *Salerno*, 481 U.S. at 747 (quoting Kennedy v. Mendoza-Martinez, 372 U.S. 144, 168–69 (1963)).

Under this standard, the first question is whether the government is seeking to impose an impermissible pretrial punishment. Substantive due process requires that the bail provision show that the government is not engaging in conduct that would shock the conscience or interfere with rights implicit in the concept of ordered liberty.[186] Obviously, if the government denied bail in an attempt to punish a presumptively innocent person, this attempt would violate substantive due process. The Court in *Salerno*, however, found that the government presented a compelling case for protecting the public safety and was thus not seeking to punish the detainees.[187] In light of *Salerno*, and given that the government interest in protecting stalking victims from threatened violence is equally compelling, it is highly unlikely that a court would consider a no-bail provision unduly punitive.

The second prong of substantive due process requires the court to look at whether "the incidents of pretrial detention [are] excessive in relation to the regulatory goal [the government] sought to achieve."[188] Looking at the Bail Reform Act of 1984, the Court held:

> When the Government proves by clear and convincing evidence that an arrestee presents an identified and articulable threat to an individual or the community, we believe that, consistent with the Due Process Clause, a court may disable the arrestee from executing that threat. Under these circumstances, we cannot categorically state that pretrial detention "offends some principle of justice . . . rooted in the traditions and conscience of our people . . . ."[189]

A court following *Salerno* would probably consider pretrial detention a non-excessive means to disable the arrestee from executing a threat.

Finally, the no-bail provision must pass the procedural due process test, which provides that when the government deprives one of life, liberty, or property, it must do so in a "fair manner."[190] The *Salerno* Court described the statutory criteria under which detainees would be entitled to a hearing before a neutral judicial officer, and the procedures the officer was required to

---

[186] *See, e.g.*, United States v. Salerno, 481 U.S. at 746.
[187] *Id.* at 747.
[188] *Id.*
[189] *Id.* at 751 (quoting Snyder v. Massachusetts, 291 U.S. 97, 105 (1934)).
[190] *Salerno*, 481 U.S. at 746.

follow.[191] It concluded that these procedural safeguards satisfied the requirements of procedural due process.

c. *No-bail clauses in antistalking legislation.* Of the state antistalking legislation we examined, only statutes in Georgia and Illinois provided law enforcement agencies the power to detain defendants without bail before trial. These statutes may be constitutionally deficient under the Fourteenth Amendment's Due Process Clause because they lack sufficient procedural safeguards for the accused.

Most illustrative is the former Illinois no-bail provision, which was revised after being struck down by a trial court on state constitutional grounds.[192] The original statute contained few procedural safeguards.[193] Its procedures were not nearly as comprehensive as those provided in the federal statute, upheld in *Salerno*, which included appearance before a neutral magistrate, an adversarial hearing with an opportunity to present evidence, the right to counsel, a written decision with findings of fact, statutory criteria to be considered by the officer, and a decision backed by clear and convincing evidence.[194] Although the Illinois statute provided for a bail hearing, it did not establish guidelines for the judicial officer to follow and did not dictate the form of the hearing. Most crucially, the original Illinois statute allowed the detention to occur without clear and convincing evidence of any threat to public safety or the safety of the victim.[195]

The revised Illinois statute contains additional procedural requirements, thereby eliminating the deficiencies of the prior ver-

---

[191] *Id.* at 751–52.

[192] David Bailey, *Circuit Judge Finds No-Bail Provision of Stalking Law to Be Unconstitutional*, CHI. DAILY L. BULL., Feb. 5, 1993, at 1. An appellate court decision on the constitutionality of the bail aspects of the law is pending. This Article does not take any position as to the constitutionality of no-bail provisions under state constitutions, but instead addresses the federal constitutional minimum.

[193] The law stated: "(a) All persons shall be bailable before conviction, except the following offenses where the proof is evident or the presumption great that the defendant is guilty of the offense: . . . stalking or aggravated stalking, where the court, after a hearing, determines that the release of the defendant would pose a real and present threat to the physical safety of the alleged victim of the offense and denial of bail is necessary to prevent fulfillment of the threat upon which the charge is based." ILL. ANN. STAT. ch. 725, para. 5/110-4 (Smith-Hurd 1993).

[194] 18 U.S.C. § 3142 (e)–(f) (Supp. II 1988); *Salerno*, 481 U.S. at 751–52. While the Court stated that "[t]he numerous procedural safeguards detailed . . . *must* attend this adversary hearing," *Salerno*, 481 U.S. at 755 (emphasis added), it later held that inadvertent omissions of some of those procedures were not automatic grounds for freeing a detainee. *See* United States v. Montalvo-Murillo, 495 U.S. 711, 717 (1990). *But see* United States v. Mantecon-Zayas, 949 F.2d 548, 551 (1st Cir. 1991) (holding that absence of written bail decision violated bail statute).

[195] The state had the burden to show only that the proof was evident or the

sion. The statute now calls for a full adversarial hearing and establishes specific criteria for a judicial officer to consider.[196] The Illinois statute explicitly requires the judicial officer to make his or her decision upon clear and convincing evidence. Its explicit bail hearing procedures could serve as a valid model for other state no-bail provisions.

By contrast, the Georgia law offers no procedural safeguards at all. It states that "the judge of a court of inquiry may impose such restrictions on the defendant which may be necessary to deter further stalking of the victim, including but not limited to denying bail or pretrial release."[197] While the reasons for no-bail provisions are compelling, this law fails to establish any procedures to guide the judicial officer's discretion. Consequently, the Georgia statute would probably violate the standards for procedural due process established under the Fourteenth Amendment.

If a state is to take the drastic approach of pretrial detention without bail of those accused under antistalking laws, due process requires that the statute have procedural and substantive safeguards in order to protect the presumption of innocence in our criminal justice system.

III. Analysis of NCJA Model Legislation

In September 1993, the National Criminal Justice Association ("NCJA"), acting pursuant to legislative mandate, released proposed model antistalking legislation. The NCJA's intent was to present states with a "constitutional and enforceable" antistalking law.[198] The analysis accompanying the model legislation, particularly with respect to conditions of pretrial release and other measures to protect the victims of stalking, is well-re-

---

presumption great that the defendant was guilty of the offense. ILL. ANN. STAT. ch. 725, para. 5/110-4 (Smith-Hurd 1993).

[196] ILL. ANN. STAT. ch. 725, para. 5/110-6.3 (Smith-Hurd 1993).

[197] GA. CODE ANN. § 17-6-1 (Harrison 1993).

[198] NAT'L CRIM. JUST. ASS'N, PROJECT TO DEVELOP A MODEL ANTI-STALKING CODE FOR STATES (Sept. 1993) [hereinafter *NCJA Report*]. The NCJA described its purpose in the Preface to the report:

> The model code development project was undertaken to satisfy § 109(b) of the U.S. Departments of Commerce, Justice, State, the Judiciary and Related Agencies Appropriations Act for Fiscal Year 1993 (Pub. L. 102-395) in which the Congress charged the U.S. attorney general, through the NIJ, to develop and distribute among the states a "constitutional and enforceable" model antistalking code.

Id. at i.

searched and should be extremely helpful to states.[199] The NCJA model legislation itself, however, does not fully solve the constitutional problems.

Under the NCJA legislation,[200] felony liability attaches if an actor "purposefully" engages in a course of conduct that would cause a reasonable person to fear bodily injury, and the actor "has knowledge or should have knowledge" that his or her target will suffer such fear.[201] According to the Commentary accompanying the NCJA statute, "if a defendant consciously engages in conduct that he knows or should know would cause fear in the person at whom the conduct is directed, the intent element of the model code is satisfied."[202]

This language amounts to a negligence standard. As the discussion set forth above makes clear,[203] such a low *mens rea* requirement raises serious constitutional concerns, particularly when, as in the NCJA Report, state legislatures are encouraged to make such negligent conduct a felony. Looking specifically at the NCJA model statute, the "has knowledge or should have

---

[199] *Id.* at 49–90.

[200] The NCJA Model Antistalking Code for the States [hereinafter *NCJA Model Statute*]:

> Section 1. For purposes of this code:
>
> (a) "Course of conduct" means repeatedly maintaining a visual or physical proximity to a person or repeatedly conveying verbal or written threats or threats implied by conduct or a combination thereof directed at or toward a person;
>
> (b) "Repeatedly" means on two or more occasions;
>
> (c) "Immediate family" means a spouse, parent, child, sibling, or any other person who regularly resides in the household or who within the prior six months regularly resided in the household.
>
> Section 2. Any person who:
>
> (a) purposefully engages in a course of conduct directed at a specific person that would cause a reasonable person to fear bodily injury to himself or herself or a member of his or her immediate family or to fear the death of himself or herself or a member of his or her immediate family; and
>
> (b) has knowledge or should have knowledge that the specific person will be placed in reasonable fear of bodily injury to himself or herself or a member of his or her immediate family or will be placed in reasonable fear of the death of himself for herself or a member of his or her immediate family; and
>
> (c) whose acts induce fear in the specific person of bodily injury to himself or herself or a member of his or her immediate family or induce fear in the specific person of the death of himself or herself or a member of his or her immediate family; is guilty of stalking.

[201] *Id.*

[202] *NCJA Report*, *supra* note 198, at 47.

[203] *See supra* notes 127–129 (arguing that a "wilfulness" standard is not constitutionally acceptable without a narrow construction supplying specific intent). *See also infra* notes 212–215 and accompanying text (discussing constitutional difficulties with a "recklessness" standard).

knowledge" standard renders the phrases "maintaining a visual or physical proximity to a person" and "threats implied by conduct" substantially overbroad and vague.[204] The following scenarios could trigger criminal liability under the model statute and enforcement would depend on the personal judgment of the observer:

- Repeated following, such as that associated with Operation Rescue, would subject the volunteer to arrest because, arguably, the volunteer should have known that such following would cause reasonable fear in light of violence sometimes associated with the abortion issue. An intent to protest at the location where the putative victim was going (to shout "Baby Killer" in a restaurant, for example, or to picket the victim's home) would not exculpate the accused under the model law.[205]
- Striking Greyhound workers could have been liable if, after the first reports of sniper fire, they repeatedly followed "scab" bus drivers on the highways. The strikers' intent to peaceably ask the drivers not to break the strike would be irrelevant.[206]
- A police officer, private investigator, or journalist could break this law if, with the intent to engage in secret surveillance, he or she repeatedly follows a target, the target realizes that he or she is being followed, and the target's ensuing fear is that of a reasonable person.
- In the heat of an argument over a poker game, Jim threatens to hit Bill. The next day in his car, Jim sees Bill on the street. He honks the horn and slows down, intending to apologize. Bill runs away and calls the police.
- The following scenario probably would not result in an arrest, but it highlights the broad scope of the model legislation. Carl spends several years in jail for sexually assaulting his next-door neighbor, Mary. When he gets out of

---

[204] The NCJA's negligence standard would require a narrow construction in order for the phrase "conveying verbal or written threats" to mean direct verbal or written threats of bodily injury or death. *NCJA Model Statute*, *supra* note 200. Otherwise, a person who made a succession of threats to commit suicide, or to "see you in hell," or to "make you pay for what you did," or "to use all means necessary to stop the abortion Holocaust," could be arrested and might even be convicted if a jury were convinced that the actor should have known that, under the circumstances, a reasonable person would be afraid.

[205] *See supra* note 58 and accompanying text.

[206] *See supra* note 171.

prison, he moves back into the same house, his family home. They see each other on the street, and Mary has noticed Carl staring at her house from time to time. She looked up once and saw Carl looking at her through her window. He averted his eyes quickly. She called the police.

By purposefully moving back into his old home, presumably Carl's constitutional right after serving a debt to society, he has "maintain[ed] a visual or physical proximity" to Mary which he arguably should have known would cause a reasonable person to fear another assault. Excluding constitutional arguments, his only hope at trial would be to convince a jury that his acts were not "directed" specifically at Mary or that Mary's fear of him was not reasonable.

Furthermore, the language of the NCJA model law practically guarantees its abuse in acrimonious family, divorce, and custody battles:[207]

- The model law would make it easier to convict the man in Virginia who was arrested for stalking based on appearing at City Council meetings, sitting in the front row, and staring menacingly at his brother, a councilman.
- Jill calls her husband at work and tells him that she is filing for divorce and that she has changed the locks on the doors to the family home. After he gets off work, Jim comes to the door and pounds on it, demanding that he be allowed in. The next night, Jim comes back, pounds on the door again and, when there is no answer, breaks the glass and enters the home, intending to get his things. Jill calls the police, who arrest Jim for stalking.
- A vicious custody battle involves allegations that a father sexually abused his daughter. The proof is weak. On several occasions, the father watches his child from a distance. Later, he tells the mother that he is monitoring the child,

[207] Existing stalking laws have been applied to family law matters. *See* Steven R. Lake, *Stalking Law Sometimes Becomes a Family Matter*, CHI. DAILY L. BULL., Sept. 4, 1992, at 5 ("If the present trend is any indication of the future of Illinois' stalking law, family law practitioners will have to familiarize themselves with the ramifications of the law in order to effectively advise clients about certain courses of conduct."). The model legislation advanced by this Article could easily be applied to divorce or custody matters involving private investigation, for example. The problem with the NCJA model is that, by proscribing purposeful conduct but not more, the law could allow arrests for whole classes of formerly legitimate, usually harmless conduct. The NCJA standards are so loose as to create the risk of exaggerated or fraudulent application by vindictive or interested witnesses.

saying "You can't keep me from seeing my own child." The police arrest him for stalking the next day.

If the NCJA model legislation were adopted by a state, one of two results would occur. Either the courts would strike down the law on overbreadth and vagueness grounds, or that state would have created felony liability for negligent infliction of emotional distress, a tort concept that has received skeptical treatment by judges and legislatures.[208] The negligence standard renders the NCJA model law unconstitutional as well as an example of bad public policy.

Despite these problems, removal of the "has knowledge or should have knowledge" standard would adequately restrict the model statute to a scope well within constitutional limits. While this modification would satisfy the Constitution, it would leave too much proscribable behavior unproscribed. Should a state adopt the model legislation advanced in this Article, however, almost all constitutionally protected conduct would remain unproscribed. If individuals and law enforcement officials clearly understand the type of conduct that is proscribed, antistalking laws will be more fairly enforced while providing adequate victim protection.[209]

IV. Model Legislation

The three critical problems with existing antistalking laws are vagueness, overinclusiveness (reaching too much constitutionally protected conduct), and underinclusiveness (excluding some types of behavior that is often called "stalking").[210] Concern about underinclusiveness led us to reject the California statute

---

[208] *See generally* Nancy Levit, *Ethereal Torts*, 61 Geo. Wash. L. Rev. 136 (1992); Virginia E. Nolan & Edmund Ursin, *Negligent Infliction of Emotional Distress: Coherence Emerging From Chaos*, 33 Hastings L.J. 583 (1982).

[209] Should a state adopt the NCJA model legislation, procedural mechanisms to protect constitutional rights would become even more essential.

[210] Because there is no clear psychological or behavioral definition of "stalking," nor a dependable means of predicting when obsessive behavior may turn violent, any criminal statute addressing this problem includes a necessary element of intuition. Nevertheless, we conclude that, at a minimum, antistalking laws should cover threats and coercive "erotomaniacal" conduct that would place a reasonable person in fear of serious violence. This Article demonstrates that support for antistalking laws among judges, legislatures, and advocates dissolves as the laws are broadened to encompass obsessive or annoying behavior that includes no independent, outward indication that the actor will become violent.

as the starting point for model legislation. The California statute requires an express verbal threat with a specific intent to cause fear of death or great bodily injury to trigger criminal liability, thereby excluding more erotomaniacal conduct than is necessary.[211] While the Michigan and Alaska statutes were potential models, they too contain troublesome elements. In Alaska, a person commits stalking if she "knowingly engages in a course of conduct that recklessly places another person in fear of death or physical injury, or in fear of the death or physical injury of a family member."[212] "Course of conduct" means repeated acts of "non-consensual contact," including acts that are simply initiated or continued without the victim's consent.[213]

These definitions present a close constitutional question. On the one hand, Alaska's requirement of recklessness without a specific intent to achieve a bad result may be sufficient to satisfy overbreadth and vagueness concerns, at least where there is a narrowly defined act requirement.[214] On the other hand, these definitions of "course of conduct" become the rough criminal law equivalent to "reckless infliction of emotional distress," a still controversial tort concept which involves problems of proof and can raise special First Amendment concerns.[215]

While the proposed model statute resembles the Alaska and Michigan statutes in certain respects, it includes modifications of those states' approaches to remove constitutional problems.

---

[211] For example, the California statute cannot easily be refitted to protect people like Erin Tavegia, who at age fifteen was stalked by a man who slowly followed her in his car as she walked to and from school every day, "sometimes offering her money to get in his car, and often saying that he would see her after school" but without ever making an overt verbal threat. Guy, *supra* note 1, at 1006 (citing Constance L. Hays, *If That Man is Following Her, Connecticut Plans to Follow Him*, N.Y. TIMES, June 5, 1992, at B1). The central feature of the California law—a verbal threat—is therefore too limiting. A better approach, adopted by the model legislation, is to proscribe threatening conduct accompanied by a specific intent to coerce or cause a fear of substantial violence.

[212] ALASKA STAT. § 11.41.270 (1993).

[213] *Id.*

[214] *See* Osborne v. Ohio, 495 U.S. 103, 112 n.9 (1990) (rejecting a claim of overbreadth based on lack of *any* scienter requirement in child pornography statute used to convict viewer of four photographs where separate default statute specified "recklessness" as the appropriate mens rea). *But see* United States v. X-Citement Video, Inc., 982 F.2d 1285, 1292 (9th Cir. 1992) (requiring knowledge that at least one performer in a pornographic photograph or film was a minor).

[215] *Cf.* Hustler Magazine v. Falwell, 485 U.S. 46 (1988) (denying recovery to a public figure for *intentional* infliction of emotional distress without a showing of actual malice).

MODEL ANTISTALKING LEGISLATION

(1) Stalking. Definitions. As used in this section

(a) "Fear of substantial violence" means a fear that a person or his or her immediate family member or housemate will suffer, at the hands or direction of the person causing the fear, any of the following: death, physical torture or serious bodily injury; rape or forced sexual contact; kidnapping; or the bombing, burning, destruction or unlawful entry of the person's dwelling place;

(b) "Forced contact" means any contact with an individual that is initiated or continued when the defendant knows or has been informed that the contact or communication is unwanted. "Contact" includes, but is not limited to, any of the following:

(i) Following or appearing within the sight of that individual.

(ii) Approaching or confronting that individual in a public place or on private property.

(iii) Appearing at the workplace or residence of that individual.

(iv) Entering onto or remaining on property owned, leased, or occupied by that individual.

(v) Contacting that individual by telephone, whether or not conversation ensues.

(vi) Sending mail or electronic communications to that individual.

(vii) Placing an object on, or delivering an object to, property owned, leased, or occupied by that individual.

No act occurring at or within 100 yards of an organized protest, demonstration, picket, rally or public meeting shall be included in the meaning of forced contact.

(c) "Course of Conduct" means a pattern of conduct composed of a series of two or more separate, noncontinuous acts involving forced contact and evidencing a continuity of intent. A verbal or written communication is not included in the meaning of course of conduct unless: (1) the communication, considered individually, would cause a reasonable person to suffer fear of substantial violence, *or* (2) the communication follows an act or communication which, considered individually, would cause a reasonable person to fear substantial violence;

(d) "Intent to coerce" shall mean an intent to force another person to engage in conduct from which he or she has a legal right to abstain, or to abstain from conduct in which he or she has a legal right to engage, when the accused knows or has been informed that the person is unwilling to comply.

(2) Stalking. Misdemeanor and Felony Offenses.

(a) A person is guilty of Misdemeanor Stalking who, with the intent to coerce another individual, engages in a course of conduct that would cause a reasonable person to fear substantial violence and that actually causes such fear in the individual.

(b) A person is guilty of Aggravated Stalking who: (1) with the intent to cause another individual to fear substantial violence, engages in a course of conduct that would cause a reasonable person to fear substantial violence and that actually causes such fear in the individual, *or* (2) in violation of a restraining order, purposefully engages in any act which, if done repeatedly, would constitute misdemeanor stalking.

(c) A person is guilty of Recidivist Stalking who, upon being convicted of any stalking charge under this section, thereafter: (1) knowingly engages in any act directed at the same individual which would cause a reasonable person to fear substantial violence, *or* (2) commits aggravated stalking against another individual.[216]

(3) Stalking. Special Procedures.

(a) Any law enforcement officer may arrest, without a warrant, any person he or she has probable cause to believe has engaged in a course of conduct in violation of the provisions of this section, *PROVIDED* that the officer has a good faith belief that delay will endanger the life of the complaining witness *and PROVIDED* that at least one act which, if repeated, would support a conviction under this Section is: (a) observed by a police officer or district attorney in the course of his or her duties; or (b) established, in whole or in part, by physical or third-party evidence in

---

[216]The question of appropriate punishment for Misdemeanor, Aggravated, and Recidivist Stalking is not addressed by this Article.

addition to the statement of a complaining witness. Where a decision to seek a warrant is made, the complaining witness shall have a right to police protection until such time as a warrant is issued or denied by the magistrate.

(b) Whether or not a warrantless arrest is made, the accused shall be entitled to an adversarial hearing before a magistrate within 72 hours of arrest, at which time the magistrate shall:

(i) make an initial determination as to any claims that the accused was engaged in activity not covered by this Section or protected by the Constitution, dismissing the charges if, excluding any such conduct, they are legally insufficient to support conviction;

(ii) upon petition of the State, determine whether bail should be denied to the accused. Bail shall not be denied unless the magistrate, after a hearing comporting with the procedures set forth in subdivision (ii)(A), makes a written finding, supported by clear and convincing evidence, that the accused committed the crime of stalking and that no other condition or combination of conditions pursuant to subdivision (ii)(B) can ensure the safety of the victim or of the community.

(A) Conduct of the Hearing.

(1) The defendant shall have the right to counsel and, if indigent, shall have counsel appointed by the court.

(2) The defendant shall have the right to proffer evidence at the hearing and shall have the right to testify and to present and cross-examine witnesses.

(3) The court shall consider the following available information in determining the threat to the safety of the victim and of the community: (a) the nature and circumstances of the stalking offense charged, including, but not limited to, the age and physical condition of the victim, and the nature of any threats or statements upon which the charged offense is based; (b) the history and characteristics of the defendant, including, but not limited to, defendant's prior criminal history indicative of violent or abusive behavior, or lack thereof; defendant's psychological, psychiatric, or other similar social his-

tory indicative of violent or abusive behavior, or lack thereof; or any other past history indicative of violent or abusive behavior, or lack thereof; and (c) whether the stalking offense charged occurred while the defendant was on probation, on parole, or on any other release pending trial, sentencing, or appeal for an offense under federal, state or local law.

(B) Release on Conditions. If the safety of the victim or of the community can be ensured by the imposition of any of the following conditions of release or combination thereof, the court shall release the defendant on bail or personal recognizance with the conditions or combination of conditions the court deems necessary:

(1) an order directing the accused to undergo psychological examination and counselling;

(2) a restraining order directing the accused to

(a) stay away from the home, school, business or place of employment of the victim of the alleged crime and any other location specifically named by the court;

(b) refrain from contacting, intimidating, threatening or otherwise interfering with the victim of the alleged crime and any other person specifically named by the court.

(c) If the defendant claims that he or she was engaged in constitutionally protected activity or that any part of the charges are based on conduct occurring within 100 yards of an organized protest, demonstration, picket, rally or public meeting, the presiding judge shall, prior to trial, determine the validity of that claim as a matter of law and, if found valid, shall exclude that activity from evidence unless its probative value in showing the purpose and character of other conduct substantially outweighs its prejudicial effect. If evidence of such conduct is admitted, the court shall instruct the jury that said conduct is legal and does not establish any element necessary to convict the defendant.

A. *Scope of the Statute*

The danger inherent in drafting a statute conservatively to avoid constitutional problems is that the result may be too narrow to achieve the original purpose. That is not the case here. Although the model legislation does not go as far as the Florida or Michigan statutes, it goes further than the California-type legislation, and the legislation adopted in most states, by addressing threatening conduct and speech when undertaken with a *coercive intent*, as opposed to an intent to cause fear of serious violence.[217] In fact, the model legislation covers the following activity that arguably falls outside California's law:

- A man slowly drives behind a teenage girl on her way to school every day, sometimes offering her money to get in the car, but never making any threatening or obscene statements.[218] He ignores pleas to stop following her. This attempt to convince the teenager to get into his car through conduct that would cause a reasonable person to fear violence or sexual assault satisfies the elements of misdemeanor stalking. In contrast, California would have to wait for a threat.
- In an attempt to force his ex-girlfriend to return to him, a man vandalizes the ex-girlfriend's car three times, steals her dog, and sends pictures of the dog.[219] Again, the absence of a threat would not prevent his arrest under the legislation.
- "He would come up and ring the doorbell and just stand there sometimes for as long as 20 minutes at a time. He phoned, on average, 100 to 120 times a day. He left bizarre, obscene notes in our door, in our mailbox."[220] Here, the addition of fear of sexual assault, combined with the intent to coerce element, allow for a stalking conviction under the model law whereas the California police could not act without stretching the language of that state's statute.

---

[217] In only one respect does the model statute exclude conduct proscribed by the California law: it does not affect a nonthreatening communication uttered *prior* to a truly threatening statement or act. Any attempt to sweep up such communications in a prohibition against stalking would likely be held unconstitutional.

[218] Guy, *supra* note 1, at 1006 (citing Constance L. Hays, *If That Man Is Following Her, Connecticut Plans to Follow Him*, N.Y. TIMES, June 5, 1992, at B1).

[219] James Quinn, *'Stalking' Law Violator Jailed a 2nd Time*, L.A. TIMES, Mar. 18, 1992, at B3.

[220] *Larry King Live* (CNN television broadcast, May 12, 1993) (interviewing stalking victims).

- An obsessed fan, after being told that a Hollywood star does not want to see him, follows her repeatedly and sends her "love" letters revealing his detailed knowledge of the layout and contents of her bedroom, saying "You won't meet with me, but we will be in heaven together soon."[221] Again, this potentially dangerous person could be locked up with or without an actual threat.

The above situations are proscribed by the model legislation in a constitutionally valid manner.

B. *Constitutional Analysis*

We contend that this statute will pass constitutional muster while encompassing more intuitively "dangerous" stalking behavior than the California law. The principal advantage of the model legislation over antistalking laws such as those in California, Florida, and Michigan, is the removal of the following ambiguous phrases: "harass," "alarm," "vex," "annoy," "emotional distress," "substantial emotional distress," "constitutionally protected conduct," and "legitimate purpose." These terms are replaced by more carefully defined terms that remove much of the vagueness discussed above while excluding the vast majority of constitutionally protected activity. For example, "emotional distress," an element of the Michigan statute, has been replaced by "fear of substantial violence," which is defined as a fear of specified acts of substantial violence to person or property, including fear of unlawful entry of a dwelling place. Likewise, "unconsented contact," which included the possibility that the accused was unaware of the lack of consent, has been replaced by "forced contact." Under the definition of "forced contact" in the model legislation, the accused is not liable unless the accused knows or has been informed that victim does not desire the contact.

Similarly, vague mens rea elements such as "maliciously" and "willfully" have been replaced by requirements of specific intent to achieve a bad result through knowing "forced contact" that would cause fear of substantial violence in a reasonable person.

---

[221] The example is hypothetical, but the quoted statement was used by Deputy Los Angeles District Attorney Rhonda Saunders as an example of language *not* covered by the existing California law. Scott Armstrong, *States Crack Down On 'Stalking'*, CHRISTIAN SCIENCE MONITOR, May 19, 1993, at 7.

Intent to cause fear of substantial violence is essentially the same as the constitutionally valid "credible threat" intent discussed above. Statutes prohibiting conduct with intent to coerce, moreover, have been upheld against overbreadth and vagueness challenges on many occasions even where the term has not been further defined.[222] Requirements that the coercion and the forced contact be "knowing" further limit the offense of stalking to inherently wrongful conduct.

Finally, "constitutionally protected activity" and "legitimate purpose" have been replaced by two specific definitional exclusions. The first, model statute section (1)(b)(vii), excludes all conduct, including following, that occurs within 100 yards of a demonstration or public meeting. The other, section (1)(c), exempts communications which do not amount to an unprotected "threat" or precede such a threat or threatening conduct. This exemption eliminates the most difficult problem with most antistalking statutes: the possibility that legitimate or constitutionally protected speech will be made illegal by a *subsequent* threat.

While these specific definitions and exclusions should resolve the most troublesome constitutional problems, three overbreadth concerns remain to be addressed. First, the model statute contains no express exception for constitutionally protected following. However, the "intent to coerce" and "intent to cause fear of violence" scienter elements, combined with the requirement that the "course of conduct" be of such a character that a reasonable person would suffer fear of substantial violence,[223] serve to exclude most constitutionally protected following. Secret surveil-

---

[222] In United States v. Mussry, 726 F.2d 1448, 1455 (9th Cir. 1984), the court held:

> The fact that the statute requires that an individual must have acted with *the intent of coercing* another into his service *goes a long way toward alleviating any vagueness problems*. It is difficult to argue that a person did not have notice that certain conduct was illegal when the offense requires that the conduct be improper or wrongful and that the actor intend that the conduct have a coercive effect.

(emphasis added); *accord* C.I.S.P.E.S. v. F.B.I., 770 F.2d 468, 472–77 (5th Cir. 1985) (holding that statute criminalizing willful coercion of foreign officials was not unconstitutionally vague or overbroad); Perkins v. Commonwealth, 402 S.E.2d 229, 234 (Va. Ct. App. 1991) ("By requiring that the language [of the state telephone harassment statute] be used with the intent to 'coerce, intimidate, or harass,' the statute gives a person of ordinary intelligence a reasonable opportunity to know what is prohibited."). *Compare* United States *ex rel.* Holder v. Circuit Court of 17th Judicial Circuit, 624 F. Supp. 68, 70–71 (N.D. Ill. 1985) (holding that Illinois statute defining intimidation as a threat to commit "any criminal offense," without restricting intimidation to behavior that had a "reasonable tendency to coerce," was overbroad).

[223] This standard for "course of conduct" is narrower than that of the California law, which requires only that the following be malicious and willful. Such broad terms could include the idea of an intent to vex or annoy, thereby risking overbreadth.

lance, for example, is not prohibited by the model statute because it would not cause a reasonable person to have any reaction at all, it would be difficult to define as "contact," and it would lack the requisite level of intent. The exemption for conduct occurring within 100 yards of a demonstration, moreover, further reduces the possibility that protected following could be proscribed. While it is possible to imagine a person legitimately intending to force someone to do something against his or her will by following him or her in a manner that would "scare the daylights" out of a reasonable person, such instances would be rare and could be challenged on an "as applied" basis.[224]

A second area of concern from an overbreadth analysis is the status under the model statute of "communications" that are not intrinsically threatening. Under the statute, communication and stalking are related in the following manner:

$$\text{COMMUNICATION} + \text{COMMUNICATION} \neq \text{STALKING}$$
$$\text{COMMUNICATION} + \text{THREAT} \neq \text{STALKING}$$
$$\text{BUT}$$
$$\text{THREAT} + \text{COMMUNICATION} = \text{STALKING}$$

Thus, there is no danger that two statements, neither of which amounts to a threat, can be collectively treated as a threatening "course of conduct." A thousand love letters, without more, would not constitute stalking under the model statute even if, considered together, they would lead a reasonable person to suffer fear of violence. Nor can an otherwise innocent comment be transformed into an element of stalking by a subsequent threat of death. Under the model statute, a love letter does not become an act of stalking because it is followed by a hate letter. This exclusion of certain types of speech will allow some acts

[224] Even under the model statute's "course of conduct" standard and demonstration exemption, extreme protest tactics may present close cases. For example, Operation Rescue volunteers have followed clinic workers with, at a minimum, the intent to force them to stop performing abortions. *See supra* notes 57–59 and accompanying text. Since that kind of following has been associated with vandalism of clinics and the deaths of doctors who perform abortions, a reasonable person might be put in fear of violence by such acts of repeated following. An Operation Rescue volunteer, therefore, could be charged with misdemeanor stalking under the model legislation. Presumably, an Operation Rescue volunteer who was so charged would claim that these acts were a form of legitimate religious protest shielded by the Constitution. Such close constitutional questions are best resolved through case by case analysis, as opposed to a facial overbreadth inquiry, where the defendant could argue that such conduct deserved protection, and the state could maintain that it has a compelling interest in prohibiting this conduct through criminal legislation.

of true stalking to go unpunished, but it is necessary to ensure that the statute does not impermissibly infringe First Amendment freedoms.

However, the model statute does allow for conviction based on an act that would cause a reasonable person to suffer fear of violence *followed* at some later time by a communication that would otherwise be constitutionally protected. For example, assume that A brandishes a knife and threatens B with mutilation if she continues to have an affair with A's husband. That night, A puts a note in B's mailbox saying "I love my husband more than life itself. Keep away." By itself, the note is probably a legitimate communication that would be protected by the Constitution. Viewed in connection with the earlier threat, however, it becomes the requisite second act of misdemeanor stalking. The connection between the two acts, the "continuity of intent," allows the communication to be treated as illegal.[225] Nevertheless, there could be rare situations in which a communication that follows a threat could not constitutionally be considered an element of the offense even though it shared the same "continuity of intent." This could be a particular problem when a substantial amount of time passes between the threat and the communication. However, to simply exclude all communications unless they independently amount to a credible threat of substantial violence would immunize otherwise prohibited stalking behavior. Accordingly, any claims that a particular communication is constitutionally protected should be considered in an "as applied" challenge.

The final problem is that the exceptions related to communications do not include symbolic acts, such as burning a cross or flag, pointing a finger as if it were a gun, burning a person in effigy, or chopping off the head of a voodoo doll. Depending on the circumstances, such acts may deserve constitutional protection, but they are not specifically exempted from the definition of "communication" in the model statute. As in the case of certain kinds of following, therefore, the possibility arises that constitutionally protected activity could support an arrest for stalking based on a literal reading of the statute. It is likely, however, that instances in which the statute would truly cover such conduct would be extremely rare in light of the statute's requirements that the act and communication share a continuity

[225] A threatening hate letter, followed by obsessive love letters, could constitute stalking under the model statute if the other statutory requirements are met.

of intent, that the intent be to knowingly coerce or cause fear of substantial injury, and that the intent be accomplished through forced contact occurring more than 100 yards away from any organized protest. The "illegitimate applications," therefore, are insignificant "judged in relation to the statute's plainly legitimate sweep"[226] and may be challenged on a case by case basis.

C. *Procedural Safeguards for the Accused and the Victim*

The model legislation set forth above should satisfy constitutional requirements while proscribing a substantial range of stalking behavior. The procedural requirements of this statute will further insulate the law from constitutional challenge while giving law enforcement officials a swift and effective response to most dangerous forms of stalking. Victims will also receive some protection under these procedures, as a police officer must either arrest the stalker or provide police protection if there is probable cause to believe that stalking has occurred.[227] The magistrate who presides over the postarrest hearing, moreover, will either deny bail where there is clear and convincing evidence that the accused committed the crime, or impose a restraining order which, if violated, will support a felony stalking charge without the need for the prosecution to prove an intent to cause fear of substantial violence.

The accused is also substantially protected. First, law enforcement officials do not have carte blanche to arrest someone without a warrant based solely on the statement of the complaining witness. Once arrested, moreover, the accused will have an immediate adversarial hearing where his constitutional claims can be initially tested.

V. Conclusion

The constitutional flaws in existing antistalking statutes are primarily a result of the hasty and highly politicized manner in

---

[226] Fallon, *supra* note 83, at 860.

[227] This protection is not absolute since the stalker may go into hiding while the warrant is sought. A requirement that police provide protection until the accused is caught, rather than allowing the police to provide such protection on a discretionary basis, would be unworkable and could create a disincentive for the police to find probable cause.

which these laws were enacted rather than from any inherent constitutional protection for true stalking behavior. The model statute presented here should be precise enough to satisfy the most exacting constitutional scrutiny. At a minimum, the defendant must intend a bad result, know or be told that his contacts are unwanted, and engage in conduct that any reasonable person would know causes substantial fear in another. The statute further protects the accused by providing exceptions for conduct that is often protected by the Constitution, and by providing procedural safeguards guaranteeing speedy and just treatment following an arrest.

The statute protects victims' interests by encompassing a significant core of stalking behavior and by empowering law enforcement authorities to act quickly to arrest stalkers and keep them separated from their victims. Finally, a statute drafted conservatively to satisfy potential constitutional objections serves the accused and the victims alike by producing fair arrests and sustainable convictions.

ARTICLE

FOREIGN TAKEOVERS OF UNITED STATES AIRLINES: FREE TRADE PROCESS, PROBLEMS, AND PROGRESS

THOMAS D. GRANT*

Civil aviation has become an international industry, and many foreign air carriers seek access to the lucrative U.S. aviation market by purchasing shares of U.S. airlines. However, the Federal Aviation Act imposes limits on foreign ownership. In this Article, Mr. Grant argues that although the Department of Transportation previously exercised its discretion by interpreting liberally its statutory mandate, recently the agency has appeared reluctant to encourage globalized aviation and, in fact, has reversed the trend toward "open skies."

Mr. Grant analyzes the "changed landscape" of civil aviation, noting that troubled U.S. carriers welcome capital from any source. He advocates a strategy of "liberalizing for leverage" in aviation diplomacy; the United States should implement a policy of explicit linkage between more liberal treatment of U.S. carriers overseas and permissive allowances for foreign investment in U.S. airlines. Mr. Grant recommends amending the Federal Aviation Act so that it authorizes, and even directs, the Department of Transportation to push for more open skies.

Foreign takeovers of United States airlines currently exist in the forefront of media attention, political debate, economic development, and the law. This Article discusses foreign takeovers of U.S. airlines in a number of contexts: in the context of free trade and the global integration of commerce and economics; of bilateral aviation diplomacy; of airline and airline-related interest groups; of an evolving administrative case law; and of proposals for legislative change to the statute governing such transactions. This Article presents an extensive narrative of the most

---

* B.A., Harvard University, 1991; Yale Law School Class of 1994; law clerk (prospective) for the Honorable Conrad K. Cyr, United States Court of Appeals for the First Circuit, 1994–95. The following persons provided helpful guidance and information for this Article: at the U.S. Department of Transportation, Donald H. Horn and Christopher T. Tourtellot (Office of General Counsel for International Law) and Patricia Szrom, Edward Oppler, and Terri Bingham (Office of International Aviation); Clark H. Onstad, former General Counsel of the Federal Aviation Administration (1977–81) and vice president of Texas Air (1981–89); Peter Passell at the Financial Section of the *New York Times*; Edward G. Fedderman, staff member of the House of Representatives Subcommittee on Aviation; the chief executive officers and their staffs at American Airlines, Delta Airlines, United Airlines, Southwest Air, Federal Express, and USAir; and the staffs of the Airlines Pilots Association and other aviation-related unions. The author especially thanks Stephen Fraidin of Fried, Frank, Harris, Shriver, & Jacobson and Professor Alan Schwartz, who co-taught the course on corporate takeovers in which work on this Article began.

important and controversial foreign airline takeover to date, the British Airways ("BA")/USAir deal, aiming to place that deal in the context of past transactions, trade liberalization, and the overall landscape of present-day airline ownership law. The Article also analyzes the current and future directions of the law and policy governing foreign ownership of air carriers.

A major theme of this Article is that while the law governing foreign takeovers previously evolved to remain consistent with a changed landscape in civil aviation, the latest developments in American regulation of airline ownership have stalled—or even reversed—that evolution. If, as is commonly assumed, competition energizes industry and commerce, then policy makers should be concerned that the progress toward free trade in U.S. airline law has abated.

Continued and prudent progress toward free trade in aviation can generate gains not only by invigorating competition within the United States but also by opening perennially protected aviation markets abroad. In order to promote free trade in aviation to its optimal extent, American foreign ownership rules must evolve beyond bilateral treaty negotiations and administrative case law; this Article argues that effective promotion of free trade requires Congressional revision of the Federal Aviation Act. In fact, the current rules have engendered a great deal of confusion and hesitation precisely because the governing statute will not support further evolution of foreign ownership rules, even though the changed context of civil aviation logically demands such evolution. As the administrative case law has approached the limits of permissible interpretation under the Federal Aviation Act, doubt and reversal have infected the rule and policy making of the U.S. Department of Transportation ("DOT"), the agency responsible for reviewing foreign takeover proposals. To remedy this doubt and reversal—manifest in recent cases and in initiatives such as the Open Skies Order—Congress should clarify the aviation statute. But simultaneously, DOT must continue in its role as a rule-changing institution. The difficult political balance (and perhaps stalemate) between interests in favor of and opposed to free trade underscores the importance—even decisiveness—of an aggressive DOT role in accelerating progress toward free trade. Like most administrative agencies, DOT is a repository of formidable technical expertise; as a result, DOT's voice may be especially convincing in a debate where the answers to close technical questions prove to be decisive.

This Article proceeds in six Parts. Part I begins by exploring the setting in which foreign takeovers occur. It introduces the statute governing foreign ownership of U.S. air carriers, the functions and mandate of DOT, and the sources and forms of agency policy enunciation. Part I proceeds to describe the features of civil aviation that have impelled the evolution of U.S. ownership rules. It further describes how and why the legal restrictions on foreign activity in U.S. aviation have driven foreign air carriers to pursue ownership. Ownership is attractive to foreign airlines for two major reasons: first, foreign airlines perceive lucrative opportunities in the huge U.S. domestic aviation market; and second, although U.S. law bars foreign carriers from flying domestic U.S. routes, it does permit them to own stock in U.S. airlines.

With the setting and rationale for foreign takeovers established, Part II addresses the diplomacy of American ownership regulation. The clash between U.S. domestic interests and foreign ownership, coupled with the limitations inherent in the statutory mandate, have necessitated an indirect approach to U.S. aviation diplomacy; administrative law interpretation and diplomatic initiatives have intertwined. Under the current scheme of U.S. aviation diplomacy, gains via treaties between the United States and foreign countries are repaid with favorable U.S. administrative rulings for foreign airlines rather than by concessions express in the treaties themselves. Part II suggests—and Part IV further explores—the possibility that the Open Skies Order, despite its promising title, in fact has subtly reversed some of the progress toward free trade in civil aviation and thus is symptomatic of the doubt and reversal infecting recent developments in aviation law and policy. Part II then focuses on the Dutch and Swedish bilateral treaties (those in which the U.S. has advanced furthest toward free trade in air commerce) and argues that the favorable treatment awarded U.S. carriers in these treaties has been repaid by liberal administrative interpretation of U.S. ownership restrictions, thereby giving Dutch and Swedish carriers generous allowances to buy into U.S. airlines. Part II concludes that because the open skies definition fails to give official gloss to the successful reciprocal regime that has arisen *de facto* in U.S. aviation diplomacy, it represents a reversal for free trade.

Part III first explains the legal significance of the pivotal terms "citizenship" and "control," then discusses their definitions in

contexts outside civil aviation in order to illustrate the general structure of ownership and control controversies. In light of this framework, Part III summarizes the administrative precedents regarding foreign takeovers of U.S. air carriers. The administrative decisions of DOT demonstrate a trend toward progressively loosened ownership and control requirements—in short, a trend toward free trade. However, the 1989 Northwest Airlines/KLM[1] case (the final case before the BA/USAir transaction) provoked aggressive opposition by U.S. airlines and other aviation-related interests. This set the stage for a slowed free trade trend. The Northwest/KLM case pushed agency interpretation close to the limits of the statutory mandate, spurring concerted opposition to trade liberalization in air commerce. This opposition crystallized in the subsequent BA/USAir case and contributed to DOT's slowed promotion of free trade.

Part IV fully develops the theme tentatively introduced in Parts I through III—that, upon approaching the interpretive frontier of the Federal Aviation Act, DOT has been beset by uncertainty and hesitation. Part IV examines informal agency statements by the Reagan and Bush Transportation Secretaries and suggests that these statements exhibit inconsistencies in foreign takeover law and policy. To illustrate recent tensions in U.S. foreign ownership policy, Part IV discusses private involvement in aviation diplomacy, suggesting that such involvement is beneficial and that DOT's hesitance to acknowledge it overtly is yet another indication of the stalled trend toward trade liberalism. Part IV revisits the Open Skies Order and argues that this DOT initiative marks the most serious indication yet of slowed liberalization in the face of an exhausted statutory mandate and energized opposition. The Open Skies Order has effectively foreclosed the fruitful diplomatic linkage between free air commerce abroad and liberal ownership rules at home—a linkage that, as argued in Part II, has been a useful vehicle for liberalization in air commerce and should therefore be formalized. To emphasize how the Open Skies Order may constitute a reversal in free trade principles, it is contrasted with European efforts toward open skies. Part IV concludes by discussing the interest groups whose influence prevents linkage from becoming a more openly stated element of U.S. aviation policy. These groups represent an in-

[1] Koninklijke Luchtvaart Maasschappij ("KLM") is the major Dutch air carrier.

creasingly concerted opposition force, and as such they are important players in the political conflict over trade liberalization.

Part V describes the BA/USAir transaction, in which the tension in ownership policy and the conflict among aviation interest groups came to a head.[2] Major U.S. carriers waged an unprecedented campaign to stop the first attempt at a BA/USAir deal, highlighting the underlying diplomatic, political, and statutory constraints that doomed the deal to failure. DOT has pushed its administrative precedents close to the statutory frontier, and subsequent angry reactions from the private sector have eroded confidence at DOT as to the future direction of foreign air commerce policy. Part V concludes with a summary of the second, successful BA/USAir proposal, but emphasizes that, given the avowed ambition of BA to use its stake in USAir as a stepping stone to greater control, the debate surrounding U.S. foreign takeover policy remains timely.[3]

Part VI proposes that explicitly liberalizing U.S. ownership and control restrictions would provide American negotiators with the leverage they need to open protected foreign aviation markets. First, Part VI analyzes the mechanics and potential of trade liberalization within the confines of the current statutory regime. Then, based upon earlier discussions about the problems encountered when administrative law reaches the frontier of statutory interpretation, Part VI argues that amending the Federal Aviation Act would be the better course for an American policy of "liberalizing for leverage." An existing proposal for a new ownership and control regime, the Clinger bill, is discussed in detail. This proposal to codify a policy of trade liberalism and reciprocity is supported by a strong state law precedent—the

---

[2] British Airways ("BA"), the leading British carrier, endeavored to purchase a sizeable stake in USAir in July 1992. This failed attempt, accompanied by a great deal of publicity and political debate, was followed by a successful proposal to purchase a somewhat smaller stake in March 1993. This transaction in discussed in great detail *infra.*

[3] Indeed, the debate has recently intensified. On November 17, 1993, a U.K. spokesperson stated that the United Kingdom wishes to gain for its carriers more liberal access to equity in U.S. airlines; if the Department of Transportation (DOT) fails to extend code-sharing permission between USAir and BA, effective January 12, 1994 the British government will strip a number of landing slots at London Heathrow Airport from U.S. carriers. Adam Bryant, *Britain Warns of Flight Bans If U.S. Retains Equity Limit*, N.Y. TIMES, Nov. 18, 1993, at D4. Permission to code-share is a step toward allowing integration of two carriers and thus is perceived as progress toward general trade liberalization when it is granted to a pair of carriers from different nations. *See infra* notes 32 and 48.

reciprocal interstate banking statutes in force between many states. These statutes provide a working illustration of how an international reciprocity rule might work in the context of civil aviation. Part VI concludes by distinguishing "liberalizing for leverage" from past, faulty exercises in strategic trade.

Part VII, a Conclusion, briefly summarizes the chief arguments of the Article and offers some further reflections.

I. Foreign Airline Takeovers: The Administrative and Global Setting

During the last several years, foreign airlines have appeared particularly eager to gain access to the U.S. air passenger market. Under current U.S. law, the sole device available to them to achieve this aim has been to purchase stock in American air carriers. However, the Federal Aviation Act of 1958[4] [hereinafter "the Act"] prevents a foreign citizen[5] from controlling a domestic U.S. air carrier.[6] The desire of foreign citizens to acquire stakes in U.S. airlines has placed great responsibility upon DOT.

---

[4] The Federal Aviation Act of 1958 ("the Act"), Pub. L. No. 85-726, 72 Stat. 731 (codified as amended at 49 U.S.C. app. §§ 1301–1557 (1988)), succeeded the Civil Aeronautics Act of 1938, Pub. L. No. 75-706, 52 Stat. 973 and the Air Commerce Act of 1926, Pub. L. No. 254, 44 Stat. 568, 49 U.S.C. app. §§ 171–184 (partially repealed 1938, repealed 1958). DOT assumed the power to review airline takeovers on January 1, 1985, pursuant to 49 U.S.C. app. § 1551(b)(1)(c) (1988), which transferred the functions of the Civil Aeronautics Board ("CAB") to DOT.

[5] The adjective "foreign" and noun "foreigner" will recur frequently in the following pages, and a caveat is in order regarding the meaning of these words. One of DOT's most significant tasks in evaluating and ruling upon proposed takeovers of U.S. airlines by foreign entities is to decide what constitutes a foreign entity.

[6] Under section 101(16) of the Act,

> "Citizen of the United States" means (a) an individual who is a citizen of the United States or one of its possessions, or (b) a partnership of which each member is such an individual, or (c) a corporation or association created or organized under the laws of the United States or of any State, Territory, or possession of the United States, of which the president and two-thirds or more of the board of directors and other managing officers thereof are such individuals and in which at least 75 per centum of the voting interest is owned or controlled by persons who are citizens of the United States or of one of its possessions.

49 U.S.C. app. § 1301(16) (Supp. 1993). Section 501(b)(1) of the Act provides that an aircraft must be registered in order to operate in the United States and that, in order to be eligible for registration, the aircraft's owner must be a "United States citizen." Coupled with a strict reading of section 413 (which provides for no distinction between direct and indirect control of a U.S. airline), *see* Appendix A, *infra*, this law would appear to forbid a foreigner from owning any stock in excess of 25% of the total value of the U.S. carrier. In fact, the statute has been interpreted to limit foreign ownership to 25% of voting stock and 49% of total equity.

All proposed foreign takeovers of American air carriers are governed by the Act,[7] under which DOT reviews all takeover proposals.[8] The statutory mandate affords DOT substantial discretion in disposing of each proposal, and DOT can use this discretion to meet the demand for evolution in the rules governing foreign ownership and control in civil aviation. As Part III suggests, DOT has recently neglected this demand. In order to place in context the now troubled role of DOT in shaping the rules that govern foreign takeovers, three matters must be discussed first: the scope of the agency's functions and mandate; the landscape of civil aviation; and the motivations behind foreign efforts to buy into American aviation.

A. *The Functions and Mandate of the U.S. Department of Transportation*

DOT is the administrative agency that enforces the citizenship and control requirements for airline takeovers.[9] DOT's formal aviation-related functions include the following: negotiation of international air transportation rights; selection of American air carriers to serve international markets in which capacity is limited by treaty; oversight of international rates and fares; maintenance of essential air service to small communities; and consumer affairs.[10] The Federal Aviation Administration ("FAA"), contained within DOT,[11] is responsible for technical and infrastructure aspects of aviation, including: air safety promotion;

---

[7] Title IV of the Act—"Air Carrier Regulation"—covers certain aspects of foreign involvement in U.S. aviation. *See* Appendix A, *infra*.

[8] For the sake of simplicity, the transactions that receive attention in this Article will be called 'takeovers.' However, it should be noted that these transactions (e.g., the purchase of part of Continental Airlines by Scandinavian Air Systems ("SAS"), *see infra* notes 112–115 and accompanying text), were not true corporate takeovers because they did not involve a change of corporate control. That distinction is central to consideration of the permissible boundaries of such takeovers; full foreign ownership and control of a U.S. airline is illegal, although partial takeovers of some U.S. airlines by foreign entities has been allowed.

[9] DOT and the Federal Aviation Administration ("FAA"—an entity within DOT) inherited jurisdiction over aviation matters (with the exception of aviation antitrust matters, which devolved upon the U.S. Department of Justice) on January 1, 1985. On that day, CAB was sunsetted by 49 U.S.C. § 1551 ("Termination of Civil Aeronautics Board and transfer of certain functions"), and its functions were reallocated by 49 U.S.C. app. § 1551(b)(1)(c) Supp. III (1985).

[10] 49 U.S.C.A. app. §§ 1371–87 (West Supp. 1993).

[11] *See* 49 U.S.C. app. 1655(c)(1) (1988).

control of navigable airspace; prescribing regulations on pilot competency; airworthiness of aircraft; air traffic control; and airport design, construction, maintenance, and control.[12]

The power of the DOT that most concerns foreign entities endeavoring to take over U.S. air carriers is the power to grant certificates to conduct air service.[13] Under the old Civil Aeronautics Board ("CAB") regime, a carrier, foreign or domestic, needed a "certificate of public convenience and necessity." Today, an entrepreneur requires a substantively similar certificate in order to conduct civil aviation. DOT issues certificates only to those corporations or other entities that fulfill a "fit, willing, and able" test of air carrier operations.[14]

After DOT has examined a request to buy into a U.S. air carrier and considered all filings from concerned third parties, the DOT Secretary issues a final order, either approving or denying the request. DOT final orders analyze every proposed transaction on the basis of the "fit, willing, and able" test. Two components of this test are particularly important to a foreign entity desiring to take over a U.S. airline—the "citizenship" and "control" elements.[15] Collectively, these final orders constitute the basic case law through which DOT has enunciated its interpretation of the Act. Informal statements by Secretaries of Transportation (including an important pair of statements by former Secretaries Samuel Skinner and Andrew Card), press releases, and Executive Orders supplement DOT's registered findings. An examination of these sources shows that the DOT has shaped U.S. policy by exercising a great deal of discretion. DOT arguably reached the outer limits of allowable agency discretion when it permitted KLM to buy forty-nine percent of the equity and twenty-five percent of the voting stock in Northwest.[16]

---

[12] 49 U.S.C.A. app. §§ 1348–59 (West Supp. 1993).

[13] *See* Federal Aviation Act § 401 ("Certificate of public convenience and necessity"), 49 U.S.C. § 1371 (1988), reproduced in part in Appendix A, *infra*.

[14] *Id.*

[15] DOT often refers to this second element as "ownership/control." *See infra* part III.A, notes 91–100 and accompanying text, for a detailed discussion of the terms "citizenship" and "control."

[16] This transaction, in which KLM purchased a significant stake in Northwest, marked the most liberal DOT allowance to a foreign entity prior to the BA/USAir transaction. DOT's Consent Order of September 29, 1989 found that the transaction between Northwest and KLM did not contravene the citizenship requirements of the Federal Aviation Act. Acquisition of Northwest Airlines by Wings Holdings, Inc., Consent Order, DOT order 89-9-51, Docket No. 46,371 (Sept. 29, 1989). On January 16, 1991, Northwest petitioned DOT to modify the Consent Order so that KLM could hold up to 49% of the equity of Wings Holdings and in order to loosen various other restrictions.

As the next section will discuss, liberalization in airline ownership rules under the statute is consistent with broader changes in the landscape of civil aviation—changes that have significantly shaped American aviation law. Responding to these changes has been—and should continue to be—a primary goal for DOT.[17]

B. *The Changed Landscape of Civil Aviation*

The central reality controlling actions within the DOT mandate is the changed landscape of civil aviation. At least four factors constitute that changed landscape.

First, many American air carriers have fallen upon hard times. The bankruptcies of Trans World Airlines ("TWA"), Pan American ("Pan Am"), Eastern, Continental, and America West symbolize the ailing state of the U.S. civil aviation industry.[18] Even the survivors of the hardships of the last several years have sustained tremendous losses. The Clinton Administration, spurred by the perceived crisis, has established a fifteen-member

---

Petition of Northwest Airlines, Inc. for Modification of Order 89-9-51, Docket No. 46,371 (Jan. 16, 1991). In an Order Modifying Conditions issued on January 23, 1991, DOT granted the new equity ceiling requested by Northwest. Order Modifying Conditions, DOT Order 91-1-41, Docket No. 46,371 (Jan. 23, 1991).

[17] Of course, an administrative agency is limited in how far it can respond to changes in economic and political landscape. The statutory mandate under which DOT issues its orders and findings limits the extent to which DOT can promote liberalism in airline ownership rules. Beyond the outer limits of the statute, Congress must promote appropriate change. At present it appears unlikely that Congress will revise the statutory mandate on foreign takeovers because a delicate political balance governs the matter. Generally, the Airline Pilots' Association and other affiliated unions oppose liberalizing the ownership restrictions on foreigners. They are joined by the biggest U.S. air carriers, Delta, United, and American. Federal Express—a major carrier, though not of passengers—has also joined the lobbying effort against further liberalization. On the other hand, the Bush Administration usually stood for free trade, as do the smaller, ailing U.S. air carriers that relish the chance to enter into relations with foreign airlines that might prove to be corporate life savers. Though the Clinton Administration has been strongly supportive of the North American Free Trade Agreement ("NAFTA") and other trade liberalizing projects, it is not yet clear whether it will promote this particular aspect of free trade. In the context of the prevailing political uncertainty and interest group conflict, DOT can play a decisive role. This Article argues that in the past DOT has acted consistently with the overall landscape in civil aviation. Today, that landscape is a liberalizing one. Therefore, to retain consistency with its past policy, DOT should promote further liberalization by means appropriate to an administrative agency.

[18] For the purposes of this Article, "aviation industry" refers only to that part of the industry concerned directly with passenger carriage. When the term "aviation industry" or any similar term appears, it denotes passenger-carrying airlines, as distinct from the broader aviation industry (airplane building enterprises, airport food concessionaires, etc.).

commission to formulate a government rescue plan for American air carriers.[19] In addition, the fiscal woes of the carriers have generated calls for private assistance in the form of foreign investment.[20]

Second, troubles in the airline industry have reduced the number of competitors in the U.S. market. If the trend of bankruptcies and consolidation continues, it is possible that only two or three large carriers will remain in the United States, raising concern that anticompetitive conditions may soon prevail in American civil aviation. Again, as with the fiscal woes of the carriers, this factor suggests that the United States should encourage foreign investment in U.S. airlines. Capital from outside investors will provide some breathing space for American carriers beset by chronic losses and poor management; this will give those carriers an opportunity to make needed investments in machinery, training, and the like while revising their unsuccessful management practices.[21] Furthermore, infusing foreign capital into the ailing carriers will sustain competition in the American aviation market by guaranteeing that more than two or three carriers remain in business.

---

[19] *See* Martin Tolchin, *Clinton Considers Measures to Help Troubled Airlines*, N.Y. TIMES, Mar. 22, 1993, at A1.

[20] *The State of the Airline Industry: Hearings Before the Subcomm. on the Department of Transportation and Related Agencies Appropriations*, 103d Cong., 1st Sess. 424–55 (1993) (written testimony of Kenneth M. Mead, Director of Transportation Issues, Resources, Community, and Economic Development Division, U.S. General Accounting Office). This testimony suggests four prongs to a new policy to assist American air carriers: (1) improve airlines' access to capital markets by easing restrictions on foreign investments and control; (2) improve access to the growing international market; (3) reduce barriers to competition; and (4) examine airline pricing practices, especially those of bankrupt carriers.

[21] The proposed Lufthansa/Continental relationship offered an excellent example of a foreign air carrier introducing its management expertise and capital into a U.S. partner in desperate need of both, potentially saving a smaller carrier from bankruptcy. *See* Adam Bryant, *Lufthansa in Bid for Continental*, N.Y. TIMES, Sept. 17, 1992, at D5. A host of other American air carriers desperately need cash and, in many cases, look to corporate suitors as a source. Executives at the ailing TWA were encouraged by the possibility that BA might bring fresh capital to the bankrupt carrier. *See* Brett Pulley, *TWA Confirms Sale Talks With USAir*, WALL ST. J., June 30, 1992, at A4. Northwest Airlines's fiscal troubles led it to investigate merging with a stronger airline. *See* Asra Q. Nomani, *NWA Weighs Sale of Routes, Merger Option*, WALL ST. J., Feb. 11, 1991 at A3. Later, Northwest lobbied aggressively in Minnesota for state financing of two new airplane maintenance facilities; critics charged that this would amount to a government bail-out of a private enterprise. *See* Asra Q. Nomani, *NWA's Northwest Sets $744 Million Funding Accord*, WALL ST. J., Nov. 12, 1991, at A2. Even United Airlines, one of the largest and best-placed carriers, posted a first quarter loss of $92.3 million in May 1992, suggesting the extent of the need for outside capital in the U.S. passenger aviation industry. *Business Briefs: UAL Corp.*, WALL ST. J., May 1, 1992, at B4.

A third feature of the changed landscape is an increasingly global economy. If foreign airlines can tap the capital markets of myriad nations while U.S. airlines are limited to mostly domestic capital, then preventing foreign investment in American aviation will hobble U.S. airlines by pushing the U.S. airline industry to the sidelines of world commerce.[22]

Finally, the traditional justifications for limiting foreign investment in American carriers have expired. United States airlines no longer merit protection as a nascent industry; an industry as big as aviation is no fledgling. Fear of losing military and technological secrets is no longer legitimate; civil aviation parted ways with its military cousin decades ago. The technology of civil aviation is so widely disseminated that a policy of restricting foreign investment to guard trade secrets is useless.[23] Although arguments about nascent industry, national security, and technological secrets made sense to Congress in 1938, they make little sense today.[24]

In light of these four changes, the rules governing civil aeronautics have evolved toward accommodating the opening of aviation markets in general and cross-border ownership of airlines in particular. In recently proposed regulations, the European Community expressly provided for a globally integrated passenger aviation industry. And until recently, the United States had taken the lead in adapting *its* rules to global integration; the liberalization of restrictions on foreign ownership and control of American carriers was advancing with the gradual evolution in civil aviation. As the discussion below[25] will show, however, this liberalizing has stalled and even reversed of late. As DOT seeks an expansive interpretation of the law on foreign takeovers of American air carriers, and thus tests the limits of the statutory mandate, heightened political pressure and legal uncertainty

---

[22] Civil aviation has become a cross-border enterprise not only in terms of raising capital but also in the day-to-day operations of the industry. On the global integration of the passenger aviation industry, see EUGENE SOCHOR, THE POLITICS OF INTERNATIONAL AVIATION 221 (1991) (drawing conclusions about the rise of international alliances between air carriers).

[23] *See, e.g.*, U.S. GENERAL ACCOUNTING OFFICE, GAO/RCED-93-7, AIRLINE COMPETITION: IMPACT OF CHANGING FOREIGN INVESTMENT AND CONTROL LIMITS ON U.S. AIRLINES (1992) (describing the irrelevance of old rationales behind the restrictions on foreign investment).

[24] *See, e.g., To Create a Civil Aeronautics Authority: Hearings Before the Committee on Interstate and Foreign Commerce*, U.S. House of Representatives, 75th Cong., 3d. Sess. (1938).

[25] *See infra* notes 123–124 and accompanying text; notes 140–145 and accompanying text.

threaten to separate American aviation law from the evolution of the international aviation industry. Yet now, more than ever, that evolution has both increased the domestic need for attracting foreign investment and stimulated foreign interest in buying American air carriers.

C. *Why Foreigners Want American Carriers—Ownership as the Shortest Route to Cabotage*

The increasing globalization of commerce has led to new efforts to skirt legal restrictions on foreign aviation within the United States—restrictions that have long characterized the aviation landscape. The result has been an upsurge in foreign (especially European) interest in purchasing stakes in American air carriers.[26] The BA offer to buy a commanding stake in USAir has received the most publicity of the foreign forays into American aviation.

What specific factors have induced foreign airlines to consider acquiring stakes in their American counterparts?

By buying into U.S. airlines, the foreign carriers hope to achieve indirectly an aim that U.S. law has long thwarted directly. That aim is "cabotage," which, in aviation parlance, means regular air carriage from a departure point in a country to a terminus in the same country. Most nations jealously guard the

[26] A number of foreign airlines have investigated acquiring stock in U.S. carriers or entering other business relationships with them. For example, KLM and BA began talks in late 1991 with the object of jointly involving themselves in close cooperation with Northwest (a large portion of which KLM already owned). *See* Brian Coleman & Asra Q. Nomani, *KLM Confirms Talks With British Air Amid Speculation on NWA Involvement*, WALL ST. J., Oct. 21, 1991, at B10. At about the same time, Lufthansa eyed USAir as a potential target. *See* Brian Coleman & Asra Q. Nomani, *Lufthansa Eyes USAir in Effort to Find Partner*, WALL ST. J., Nov. 7, 1991, at C14. Earlier that year, Lufthansa joined with American Airlines to run joint routes between Chicago and points in Germany. *See Lufthansa, American Air Agree to Run Joint Routes*, WALL ST. J., Mar. 1, 1991, at A4E. In 1992, American Airlines also considered an ownership relationship with KLM. *See AMR Said to Hold Talks About a Stake in KLM-Royal Dutch*, WALL ST. J., May 11, 1992, at B8. Continental Airlines received attention from both Air Canada and BA in mid-1992. *See Continental Has 4 Potential Suitors Studying Its Books*, WALL ST. J., June 17, 1992, at A4. Talks and moves toward strategic alliances between Canadian Airlines and American, as well as between Delta, Singapore Airlines, and Swissair, add to this rich array of international relationships among air carriers. *See* G. Pierre Goad & Bridget O'Brian, *American Air, Canadian Airlines to Start Detailed Talks on a "Strategic Alliance,"* WALL ST. J., Mar. 20, 1992, at A4; *Delta Air Lines Ticket Alliance*, WALL ST. J., May 7, 1992, at A4. In short, American air carriers—and access to ownership of them—figure prominently in an important and fast-moving sector of world commerce.

Table 1 Comparative Aviation Market Figures *(for 1991, in millions)*

Region	Aircraft km Performed	Ton-km of Passengers	Ton-km of Freight	Ton-km of Mail
Europe	551,270	53,245	18,380	1,270
Africa	38,320	3,490	1,105	55
Middle East	44,970	4,130	2,245	65
Asia and Pacific	358,970	31,190	17,330	740
North America	759,830	68,950	15,730	2,850
Latin America and Caribbean	87,660	8,040	2,780	35

privilege to conduct such internal civil aviation. The United States—by far the world's largest and potentially most lucrative domestic aviation market under a single sovereignty[27]—restricts cabotage quite unbendingly; no foreign air carrier may provide domestic air service in the United States.[28] But despite these prohibitions, the U.S. market inevitably tantalizes businessmen who face limited opportunity for growth in their native lands.[29]

[27] The U.S.S.R. once held the record for the sheer number of domestic flights, but since the collapse of the Soviet empire, the United States indisputably supports the world's largest internal civil aviation market. *See infra* note 29.

[28] *See* Federal Aviation Act § 1108(b), 49 U.S.C.A. app. § 1508(b) (West Supp. 1993).

[29] Comparative civil aviation statistics for various world regions suggest why foreign entrepreneurs eye the American domestic market with great interest (Table 1).

INTERNATIONAL CIVIL AVIATION ORGANIZATION, DOCUMENT 9180/17: CIVIL AVIATION STATISTICS OF THE WORLD Tab. 1-15, 1-16 (1991). The U.S. market offers other advantages over the European market, beyond size. The U.S. market functions under a single legal jurisdiction governed by federal aviation law, but the European market—hardly insubstantial taken as a whole—is divided into a dozen small sectors, each under a separate national regime. European legal heterogeneity greatly complicates the aviation business.

Furthermore, in Western Europe, where rail is the dominant mode of inter-city transit, the potential for passenger aviation growth is limited. Efficient, dependable, and fast (as well as lavishly subsidized) railroads serve many city pairs in continental Europe that short flights, such as the shuttles that serve the U.S. Northeast Corridor, could serve. Moreover, dedication to short- and medium-haul rail transit constitutes a deep-seated assumption in European policy making. For example, a Commission of the European Communities draft proposal for a new aviation liberalization package states:

> It is clear that the transport function would suffer dramatically if the cost of transport is too high. This might not be critical when other modes of transport may serve but for distances over 1000 km air transport is of particular importance.

COMMISSION OF THE EUROPEAN COMMUNITIES, COM(91) 275 FINAL, COMPLETION OF THE CIVIL AVIATION POLICY IN THE EUROPEAN COMMUNITIES TOWARDS SINGLE MARKET CONDITIONS 3 (Sept. 18, 1991). Strongly implicit in this statement is that air

Many leading European passenger carriers have endeavored, by devices that sidestep cabotage, to win access to the perceived American aviation bonanza. The primary device employed by foreign corporations to tap the U.S. aviation market has been the purchase of major stakes in U.S. airlines. But in order to sidestep cabotage successfully, foreigners must struggle against another set of U.S. aviation restrictions—the ownership and control limits embodied in the Act.[30] Though these restrictions are only partial (unlike the cabotage ban, which is complete), they nonetheless present a substantial obstacle to foreign airlines endeavoring to compete in the U.S. market. Thus, the ownership and control limits are critical for U.S. trade relations in the realm of bilateral aviation treaties.

In addition to sidestepping the cabotage ban, takeovers also allow foreign carriers to coordinate their schedules with U.S. carriers. For example, by purchasing a controlling interest in Northwest Airlines, KLM positioned itself not only to profit from internal United States aviation (thus circumventing cabotage restrictions), but also to enhance its own international routes. Coordinated schedules allow domestic Northwest flights to connect with international KLM flights.[31] The partner airlines thus merge their formerly separate route plans into a more or less seamless network that resembles a single airline.[32] This com-

---

transport is not important enough to merit special concern for routes under 1000 kilometers.

Despite this general observation, rail transit is competitive in certain U.S. transportation markets. For example, in 1991, a marketing battle erupted between the Amtrak Metroliner and the Trump and Delta air shuttles over the Northeast Corridor. *A Land-Air Battle Steams Ahead on Lucrative Northeast Corridor Routes*, WALL ST. J., Oct. 31, 1991, at A1.

[30] *See* 49 U.S.C. app. § 1401 (1988).

[31] In March 1991, Northwest announced that it would send passengers from Minneapolis-St. Paul to Amsterdam by bringing passengers to the East Coast on Northwest jets and then transferring them to KLM jets for the trans-Atlantic leg of the journey. *See Northwest Plans Amsterdam Service in Deal With KLM*, WALL ST. J., Mar. 18, 1991, at B3.

[32] BA and USAir have announced that, if DOT allows BA to purchase its requested stake in USAir, these two airlines will extend further the unifying notion tentatively pioneered by Continental and SAS (and continued by Northwest and KLM). In the words of Nancy Vaughan, an official of USAir, "Customers will know that they're on one airline from Altoona, Pa., to Bombay, India." Julie Schmit, *USAir Takes Off With British Air*, USA TODAY, July 22, 1992, at A1. One way of accomplishing this goal is via code-sharing; this means that the domestic and international legs of a journey can be ticketed as a single transaction, with luggage being checked directly through from Altoona to Bombay. Code-sharing has the additional advantage of giving the appearance that the passenger is dealing with only one airline. *See infra* note 48.

bination produces a graphic illustration of the theoretical principle of synergy. The foreign carrier profits from a new asset most easily (and probably only) acquired through purchase of the domestic target—internal U.S. feeder lines to carry passengers to its international departure points. The American carrier profits from its new ability to offer "seamless" service between smaller, outlying cities and overseas hubs; however, this synergy is not equivalent to a foreign airline profiting directly from U.S. domestic routes by replacing domestic carriers (i.e., cabotage) or by earning dividends from ownership of profitable domestic carriers.

In summary, two closely linked but distinct aims comprise the foreign takeover trend in the U.S. passenger aviation market: first, winning *de facto* cabotage via partial ownership; and second, adding more passengers to the foreign carrier's international flights by feeding into those flights the passengers carried domestically on the American target's continental U.S. routes. In the latter situation, the foreign carrier enhances its existing transatlantic service (e.g., KLM obtains a coordinated schedule between its East Coast hubs and Northwest's continental American route network). This exemplifies a rather traditional and conservative strategy of enhancing assets already in hand by investing more in them, and the synergistic benefits to both airlines are obvious. But the former aim, where the foreign carrier profits directly from domestic flights (e.g., KLM shares in the profit from Northwest's lucrative New York/Minneapolis and Detroit/ Chicago routes[33]), may represent a much more aggressive objective, especially given the long history of American restrictions on foreign civil aviation: foreign airlines are endeavoring to open whole new areas of activity by making inroads into the traditionally restricted U.S. domestic market.

II. The Diplomacy of American Ownership Regulation

The diplomacy of American aviation is a diplomacy of the regulation of foreign takeovers. American politics constrains bargaining by U.S. aviation negotiators; with cabotage and many

[33] *Summary of Equity Holdings*, in JANE'S WORLD AIRLINES NWA 1 (A10/01) (Jane's Info. Group 1991). The author thanks the staff of the Army Corps of Engineers Library, Fort Belvoir, Virginia, for providing access to this hard-to-find publication.

other improvements to foreign access virtually foreclosed by domestic interest groups, an allowance for foreign ownership of American air carriers is practically the sole concession that U.S. negotiators can grant to their foreign counterparts. Thus, the administrative interpretation of the law governing ownership is integral to the conduct of American foreign aviation policy. I will argue here that the liberalization of restrictions on foreign ownership of U.S. airlines constitutes the centerpiece of a policy of reciprocity in international aviation. This liberalization, which previously proceeded in pace with the globalization of air commerce, has recently stalled. Although a policy of reciprocity and trade liberalism has been fashioned out of administrative law and bilateral aviation treaties, DOT's 1992 open skies definition belies the agency's hesitation in the promotion of trade liberalism.

A. *Bilateral Aviation Treaties and the Open Skies Definition*

Bilateral civil aviation treaties define privileges, both in traffic between two countries and within their borders.[34] These agreements govern destinations, the number of landing slots,[35] the frequency and routing of flights, code-sharing, and whether cabotage is granted. The competitive position of an air carrier operating outside its home country is largely determined by the terms of bilateral treaties. An extremely restrictive bilateral treaty, such as the one between the United Kingdom and the United States,[36] sharply limits the freedom of one nation's airlines to route flights to destinations inside the other nation's

---

[34] Importantly, federal law provides that U.S. regulations on foreign air commerce shall not interfere with bilateral agreement terms, provided those terms are consistent with American interests and policy:

> § 1517(c) Transportation pursuant to bilateral agreement.
> Nothing in this section shall preclude the transportation of persons (and their personal effects) or property by foreign air carriers if such transportation is provided for under the terms of a bilateral or multilateral air transport agreement between the United States and a foreign government or governments and if such agreement (1) is consistent with the goals for international aviation policy set forth in section 1502(b) of this title and (2) provides for the exchange of rights or benefits of similar magnitude.

49 U.S.C.A. app. § 1517(c) (West Supp. 1993). For the text of section 1502(b) (codifying a 1980 amendment to section 1102 of the Act), see Appendix A, *infra*.

[35] For an informed comment on the importance of access to takeoff and landing slots, see Asra Q. Nomani, *Global Dogfight: World's Major Airlines Scramble to Get Ready for Competitive Battle: American and United, Aided by Hubs Such as Chicago, Seem the Best-Positioned*, WALL ST. J., Jan. 14, 1992, at A1.

[36] *See infra* notes 229–230 and accompanying text.

territory. An important aim of American aviation diplomacy is to achieve favorable bilateral treaties. Since U.S. law restricts foreign ownership of American air carriers, the diplomacy of bilateral treaties is intimately related to U.S. interpretation of its own law of foreign ownership.

The development of an official U.S. open skies definition has led the effort to improve relations in international air commerce. During the spring and summer of 1992, DOT formulated its definition of open skies and promulgated a policy of encouraging full aviation access between the U.S. and its European trade partners. Secretary of Transportation Andrew H. Card, Jr. announced in a speech that "[the United States] will now offer to negotiate open skies agreements with all European countries willing to permit U.S. carriers essentially free access to their markets."[37] This new policy followed from a recent U.S. trend of automatically opening certain American gateway cities to foreign carriers, whereas complex negotiations had once been a prerequisite for granting such rights.[38] Secretary Card stressed the importance of honoring bilateral aviation agreements and stated that "easing restrictions on foreign investment in U.S. airlines" constituted a key part of open aviation.[39]

Soon thereafter, DOT issued an "Order Requesting Comments" on open skies.[40] The Order, served upon all certificated air carriers, all foreign air carriers, and interested U.S. agencies[41] and industry groups,[42] enumerated eleven basic elements that DOT would tentatively include in its open skies definition:

---

[37] Andrew H. Card, Jr., Remarks at the Aero Club Lunch in Washington, D.C. 6 (Mar. 31, 1992) (transcript on file with the *Harvard Journal on Legislation*).

[38] This earlier step of waiving negotiations for access to gateway cities aimed to increase the number of international flights to and from underserved U.S. cities. It is instructive that DOT took that measure—which streamlined the negotiation process—in order to benefit economically disadvantaged areas of the country. This action seems, at least implicitly, to confirm that DOT understands that opening trade as broadly as possible leads to net economic benefits.

[39] Card, *supra* note 37, at 5.

[40] Defining "Open Skies;" Order Requesting Comments, 57 Fed. Reg. 19,323 (Dep't Trans. 1992).

[41] These agencies included: Customs Service; Animal and Plant Health Inspection Service; Department of Justice, Antitrust Division; Department of State, Office of Aviation; Federal Aviation Administration; and Immigration and Naturalization Inspection Service.

[42] These groups included: Airline Pilots Association International; U.S. Airports for Better International Air Service; Airports Association Council International; and the National Air Carrier Association.

(1) Open entry on all routes;[43]
(2) Unrestricted capacity and frequency on all routes;
(3) Unrestricted route and traffic rights, that is, the right to operate service between any point in the United States and any point in the European country, including no restrictions as to intermediate and beyond points, change of gauge, routing flexibility, coterminalization, or the right to carry Fifth Freedom traffic;[44]
(4) Double-disapproval pricing[45] in Third and Fourth Freedom[46] markets, and price leadership in third country markets to the extent that the Third and Fourth Freedom carriers in those markets have it;
(5) Liberal charter arrangements (the least restrictive charter regulations of the two governments would apply, regardless of the origin of the flight);
(6) Liberal cargo regime (criteria as comprehensive as those defined for the combination carriers);[47]
(7) Conversion and remittance arrangements (carriers would be able to convert earnings and remit in hard currency promptly and without restriction);
(8) Open code-sharing opportunities;[48]
(9) Self-handling provisions (right of a carrier to perform/control its airport functions going to support its operations);
(10) Procompetitive provisions on commercial opportunities, user charges, fair competition and intermodal rights; and

---

[43] Under traditional, restrictive bilateral aviation agreements, routes are assigned to a government-designated airline or airlines. Under an open entry provision, route access is not restricted to a limited list of officially appointed carriers.

[44] Bilateral aviation agreements typically disallow service beyond the signatories' territories, except to destinations expressly designated. "Fifth Freedom" refers to the "privilege to fly into the territory of the grantor state for the purpose of picking up, or putting down, traffic destined for, or coming from, third states." SOCHOR, *supra* note 22, at 265.

[45] "Double-disapproval" pricing means that, in order for a given fare to be struck down as anticompetitive, both signatories must reject it.

[46] The "Third Freedom" is the "privilege to fly into the territory of the grantor state and discharge traffic from the flag state of the carrier." SOCHOR, *supra* note 22, at 265. The "Fourth Freedom" is the "privilege to fly into the territory of the grantor state and carry back traffic destined for the flag state of the carrier." *Id.*

[47] This provision calls for equally liberal treatment of flights carrying mixed cargo/passenger loads ("combination carriers") and flights solely carrying cargo.

[48] "Code-sharing" refers to the practice of providing passage under the name of a single airline and through the ticketing and airfare arrangements of that airline, even though passengers must switch airlines *en route*. Without code-sharing, a passenger who needed to switch airlines to go from point A to point C through point B would require separate ticketing, billing, scheduling, and baggage check-in. Open code-sharing makes it easier for airlines to form strategic operating alliances. For example, if BA and USAir wish to appear as if they are a single airline, they must be allowed to code-share.

> (11) Explicit commitment for nondiscriminatory[49] operation of and access for computer reservation systems.[50]

DOT subsequently issued a mild qualification to these eleven elements, noting that case-by-case analysis would be needed to ensure protection of American public interests.[51]

On August 5, 1992, DOT issued its Final Order defining open skies [hereinafter "Open Skies Order"].[52] The Order and its formulation are instructive in analyzing the agency's view of its own mandate and goals. Comments and replies from forty-five parties helped to shape the Final Order. Addressing those respondents who expressed general, blanket opposition to the open skies initiative, DOT re-emphasized its commitment to liberal airline trade agreements:

> We are frankly and firmly committed to freer trade in civil aviation services, and our commitment is grounded, in large part, on our experience with both the market-oriented and the restrictive approaches that govern many of our current bilateral aviation relationships. We have seen much larger dividends in those markets which allow greater scope for airline price and service initiatives.[53]

The agency added that the lack of an exact correspondence between concessions given and received must not impede nego-

---

[49] Great controversy has arisen between airlines and countries concerning the computer reservation systems ("CRS") that travel and airline agents employ to find and reserve flights for passengers. Concerned parties have accused others of using these systems to discriminate against rival airlines and hinder competition. For example, rivals might construct a CRS to exclude competitors' flight reservation and schedule data, or might arrange a system so that competitor information is the last to appear on the screen. For an example of federal intervention in a case of alleged CRS abuse, see U.S. DEPARTMENT OF JUSTICE, RELEASE 89-191, JUSTICE OPPOSES MERGER OF AIRLINE COMPUTER RESERVATION SYSTEMS (June 22, 1989). For a discussion of the anticompetitive impact of these systems, see U.S. DEPARTMENT OF JUSTICE, 1985 REPORT OF THE DEPARTMENT OF JUSTICE TO CONGRESS ON THE AIRLINE COMPUTER RESERVATION SYSTEM INDUSTRY (Dec. 20, 1985), *cited in Airline Computer Reservation System Industry, 1988: Hearing Before the Subcomm. on Aviation of the House Comm. on Public Works and Transportation*, 100th Cong., 2d Sess. 227 (1988).

[50] Defining "Open Skies;" Order Requesting Comments, *supra* note 40.

[51]

> Also, we would naturally expect that our determination in any given case of whether the open-skies criteria had been met would always be subject to public interest considerations, including factors in an individual case that could seriously affect the ability of U.S. carriers to realize the benefits of an open-skies agreement, such as access to key airports.

Defining "Open Skies;" Final Order, DOT Order 92-8-13, Docket No. 48,130, at 12 (Aug. 5, 1992) [hereinafter "Open Skies Order"].

[52] *Id.*

[53] *Id.* at 3.

tiations for free trade: "Indeed, if we were to embark on negotiation initiatives only where we could anticipate precisely equal economic benefits we would have been deterred from some of the most successful agreements we have achieved in the last decade."[54] Further developing its view of the diplomatic process, DOT re-emphasized the central importance of bilateral negotiation and of case-by-case consideration:

> Some commenters have raised the question of whether each and every element must be effective immediately and simultaneously for an open-skies relationship to exist. We accept that in some contexts the phasing in of certain aspects of our definition might not be inconsistent with the overall notion of an open-skies regime. However, *rather than make any specific determination here, we regard this issue as one that we can better assess on a case-by-case basis in the context of the negotiating process.*[55]

DOT's emphasis on the negotiation process, along with its statement that it will phase in open skies measures in response to circumstances, affirmed DOT's broad discretion over the law governing aviation diplomacy and confirmed that free trade is, at least in principle, central to DOT policy.

The final open skies definition is virtually identical to the text originally proposed in April 1992. The Final Order did address additions proposed by interest groups; in particular, outside respondents had suggested that liberal ownership/control and cabotage elements be incorporated within the open skies definition. By expressly excluding these elements from the Final Order definition and by providing justifications for that exclusion, DOT emphasized the sensitivity of the ownership/control and cabotage issues.[56] Exclusion of those elements, coupled with the attention they received in the official statement accompanying the Order, identified them as crucial unresolved questions in U.S. aviation policy.

DOT offered the following explanation for its exclusionary decision:

> Some of the proposals go to matters governed by statute.

---

[54] *Id.*

[55] *Id.* at 13 (emphasis supplied).

[56] In part IV.C, *infra* notes 140–145 and accompanying text, I will argue that this exclusion marks a retreat from a long line of cases that previously pushed the law toward a liberal interpretation of the Act.

> Consequently, they could be implemented only by legislative change and are outside the scope of this proceeding. *Certain other matters* would involve elements of policy, and therefore, would be within our discretion to tailor to specific situations. However, we feel that the best way to proceed is to address these matters on a case-by-case basis. In this way, these concerns can be considered in the context of each bilateral exchange.[57]

In refraining from specifying the statutory constraints upon its actions, DOT implied that it reserves some discretion concerning the two unincorporated proposals, cabotage and ownership/control. DOT then made explicit its suggestion that the agency reserves broad discretion: it stated that "certain other matters" fall within a "policy" category and that all matters in that category are matters of agency discretion. DOT left unspecified what these "certain other matters" are. Having reserved flexibility for itself, DOT restated the importance of the bilateral negotiating process—a process that, by diplomatic rule and necessity, is kept from public view and is the area of greatest agency discretion. Thus, although the exclusion of cabotage and ownership/control from the open skies definition may appear at first to foreclose those issues, the open-ended nature of the Final Order, combined with ample signals from DOT that it views itself as retaining substantial discretion over those issues, reserves for the future the matter of what DOT may or may not offer at the negotiating table. In short, the exclusion of cabotage and ownership/control elements does not eliminate these issues as DOT bargaining chips when the agency pursues liberal bilateral aviation terms. Nonetheless, as I will discuss in Part IV of this Article,[58] the omission of these elements from the Final Order merits serious concern for free trade proponents.

It is instructive to examine the steps that European Community ("EC") nations have taken toward their own open skies definition. The integration of aviation constitutes a major part of European plans for economic unity. Therefore, EC Member States have moved toward creating a multilateral open skies agreement among themselves. The Second Liberalization Package of the EC established Third, Fourth and some Fifth Freedom Traffic rights among all airports of the Member States. Fifth

---

[57] Open Skies Order, *supra* note 51, at 12 (emphasis supplied).
[58] *See infra* notes 125–126 and accompanying text.

Freedom is being implemented in phases, the first of which began on January 1, 1993.[59] A "Proposal for a Council Regulation (EEC) on access for air carriers to intra-Community air routes" contains numerous other liberalization terms:

(1) Abolition of "capacity restrictions between Member States by 1 January 1993;"[60]
(2) Initiation of cabotage rights, as Regulation (EEC) No. 2343/90 confirmed that these form an "integral part of the internal market;"[61]
(3) "[P]has[ing] out restrictions concerning multiple designation, fifth-freedom and cabotage rights;"[62]
(4) Implementation of "non-discriminatory rules for the distribution of air traffic between airports within the same airport system;"[63] and
(5) Provision for negotiating with non-Community countries on traffic rights between Community airports.[64]

Perhaps most significantly, the EC plans to implement full cabotage among its members in phases, starting on January 1, 1993 and reaching full scope by early 1997. The EC intends a four-year "transition period" to soften the impact of cabotage on less competitive European carriers.[65]

B. *Reciprocity in Aviation Rights—the Swedish and Dutch Cases*

The bilateral aviation treaties with Sweden and the Netherlands are the most liberal that the United States has signed. Indeed, in a unique twist in the world of aviation relations, these

---

[59] COMMISSION OF THE EUROPEAN COMMUNITIES, *supra* note 29, at 12.
[60] *Id.* at 40.
[61] *Id.*
[62] *Id.* at 41.
[63] *Id.*
[64] *Id.* at 42. The fifth term carries great import for future U.S. aviation diplomacy: it will redefine Europe as a single negotiating entity, thus replacing a dozen comparatively small countries that would otherwise bargain individually. This term, if implemented, will alter the nature and balance of U.S. open skies diplomacy, because cabotage within the EC will constitute a much more valuable bargaining chip than cabotage within any one EC state.
[65] Les Gouvernements des Etats membres et la Commission des Communautés européennes, Transports Aeriens: Troisième Paquet de Liberalisation 6 (June 24, 1992) (on file with the *Harvard Journal on Legislation*).

two nations have actually granted cabotage to American air carriers, even though cabotage arguably constitutes the most prized and jealously guarded privilege in passenger aviation.

The bilateral aviation agreement between Sweden and the United States went into effect shortly after it was signed in Washington, D.C. in December 1944. Last amended in August 1983, it remains a liberal and highly permissive agreement.[66] A very permissive treaty with the Netherlands was recently replaced by one more liberal still.[67]

Diplomatic negotiators are well aware that they rarely receive something for nothing.[68] America's two most liberal bilateral aviation treaties, however, do not clearly reveal what the United States has given to Sweden and the Netherlands in return for the open access granted to American carriers. But before exploring what Sweden and the Netherlands have gained by signing liberal bilateral aviation treaties, it is necessary to discuss what the United States has gained.

Cabotage does not amount to much of a prize inside the Netherlands, a country covering a mere 37,330 square kilometers (14,413 square miles)—barely one-third the size of Ohio. Cabotage within Sweden is not much more enticing, given the comparatively low aviation activity in that large and sparsely populated country. Sweden's 8.6 million inhabitants are concentrated disproportionately in the southern parts of its 449,964 square

---

[66] Agreement Relating to Air Transport Services, Dec. 16, 1944, U.S.-Swed., 58 Stat. 1466, *amended by* Agreement Amending the Agreement of Dec. 16, 1944, August 2, 1983, U.S.-Swed., T.I.A.S. No. 10,788.

[67] Agreement Between the United States and the Netherlands Amending the Air Transport Agreement of Apr. 3, 1957 (as amended), Oct. 14, 1992, U.S.-Neth., Dep't State No. 93-93, available in Westlaw, 1993 WL 283579 (Treaty). This treaty entered into force on May 11, 1993.

[68] The "goals for international aviation policy" as enunciated in the Act, especially 49 U.S.C. app. § 1502(b)(8), demonstrate that Congress intended reciprocity to hold a significant place in U.S. aviation diplomacy. *See* Appendix A, *infra*. For a variety of discussions about reciprocity in negotiations, see, *e.g.*, Christopher Dupon & Guy-Oliver Faure, *The Negotiation Process*, in INTERNATIONAL NEGOTIATION: ANALYSIS, APPROACHES, ISSUES 40, 46 (Victor A. Kremenyuk ed., 1991) ("*Concessions*, a major component of process dynamics, result from the belief that they will elicit reciprocal concessions from the other party.") (emphasis in original); FRED CHARLES IKLÉ, HOW NATIONS NEGOTIATE 104 (photo. reprint 1982) (1964) ("A standard recommendation for facilitating agreement is that parties should make concessions. The rule that a concession by the opponent ought to be reciprocated with a concession by oneself combines the virtue of flexibility with the norm of reciprocity."); GEORGE C. HOMANS, SOCIAL BEHAVIOR: ITS ELEMENTARY FORMS 62 (Robert K. Merton ed., 1961) ("The open secret of human exchange is to give the other man behavior that is more valuable to him than it is costly to you and to get from him behavior that is more valuable to you than it is costly to him.").

Table 2 Total Domestic Aviation in Sweden (*all figures in thousands and include domestic legs of international flights*)

Aircraft kilometers performed	53,450
Passenger kilometers performed	3,173,451
Seat kilometers available	4,861,907

kilometers (173,732 square miles), a region slightly smaller than California. Compared to the almost non-existent Dutch internal aviation market, Swedish internal aviation might constitute an appreciable market (Table 2).[69] On the other hand, Swedish internal aviation hardly compares with that of the United States (Table 3).[70]

However, one must not forget that access to a country's aviation market consists of more than access to its internal market. If the bulk of a country's aviation consists of terminating or departing international flights, then access to flight routes, landing slots, code-sharing, and airport resources constitutes a precious asset. Crowded European conditions contribute an additional premium to this asset of access.[71] Thus, cabotage rights do not always add up to the biggest prize a nation has to offer (Table 4).[72]

While some American trade partners possess sizeable internal airline markets, Sweden and Holland do not.[73] Instead, their primary rewards to the United States are liberal allowances for international flights.

The American reward to Sweden and the Netherlands is not so direct. Rather, it is fashioned out of administrative interpretations of the law that governs foreign ownership of American air carriers. As discussed above,[74] a foreign airline desiring a

[69] INTERNATIONAL CIVIL AVIATION ORGANIZATION, DIGEST OF STATISTICS NO. 358 (1991); CENTRAL INTELLIGENCE AGENCY, THE WORLD FACTBOOK 1992 (1992).

[70] *Id.*

[71] In Europe, airspace is very crowded, and airports even more so. The cost of expanding European airports generally exceeds comparable costs in even the most densely populated areas of the United States. Since the supply of landing slots at important European aviation hubs is inelastic, winning access to these slots carries great rewards.

[72] INTERNATIONAL CIVIL AVIATION ORGANIZATION, *supra* note 69; CENTRAL INTELLIGENCE AGENCY, *supra* note 69.

[73] *Id.*

[74] *See supra* notes 26–30 and accompanying text.

Table 3 Domestic Aviation in the United States (*all figures in thousands and exclude domestic legs of international flights*)

Aircraft kilometers performed	467,814
Passenger kilometers performed	20,446,847
Seat kilometers available	36,162,840

share of the U.S. domestic passenger aviation market faces a major obstacle in the U.S. refusal to grant cabotage privileges to foreigners. Cabotage within the United States certainly has never been put on the American bargaining table, and it seems unlikely that this refusal will change in the near future.[75] Therefore, foreign airlines must resort to ownership of domestic U.S. air carriers in order to tap into the continental American market, which leads foreign airlines and their governments to become greatly concerned with American regulation of ownership.[76]

It is no coincidence that the major airlines of Sweden and the Netherlands—the two countries that have signed over the most generous aviation privileges to the United States—have been treated generously by DOT in their campaigns to own stakes in U.S. carriers. In short, the U.S. has repaid its best aviation partners with liberal interpretations of American ownership restrictions.

Continental Airlines (an American air carrier) captured the interest of SAS, the Swedish flag carrier, in 1990, and SAS became the first major European airline to purchase a large stake in a U.S. carrier when it bought 16.8% of the nonvoting shares of Continental. In this transaction, DOT set an ownership standard that, compared to earlier precedents, was quite permissive. A review of the U.S. bilateral aviation treaty with Sweden reveals a correspondingly liberal and permissive regime.

Similarly, the Netherlands's bilateral treaty with the United States contains the most liberal provisions of any of America's major bilateral treaties; the Netherlands went so far as to con-

[75] Diverse segments of the American aviation industry, including the Air Line Pilots Association, Delta Airlines, American Airlines, and the Airline Flight Attendants Union have strongly opposed granting cabotage rights to foreign airlines. Political inertia probably works in favor of continuing this restrictive policy.

[76] For an example of foreign criticism of the U.S. ownership regime, see Bryant, *supra* note 3 (describing British threats to limit U.S. access to Heathrow Airport if DOT does not loosen American ownership limits); Martin Fletcher, *Bush to Reject BA Links with USAir*, THE TIMES (London), Dec. 17, 1992, § 2, at 1.

Table 4 Comparisons Between Domestic and International Aviation in Selected Major United States Trading Partners (*in thousands*)

		Aircraft km Peformed	Passenger km Performed	Seat km Available
Sweden	Int'l	49,356	4,656,736	6,941,442
	Dom.	53,450	3,173,907	4,861,907
France	Int'l	210,604	26,537,073	38,493,674
	Dom.	72,309	13,324,196	18,577,681
United Kingdom	Int'l	327,409	54,322,501	78,994,516
	Dom.	58,633	4,020,001	6,326,694
The Netherlands	Int'l	139,198	23,466,440	33,974,206
	Dom.	643	9,673	26,299
Germany	Int'l	265,071	31,205,901	46,911,255
	Dom.	42,140	2,800,419	4,676,830
Japan	Int'l	133,524	23,922,366	37,654,404
	Dom.	41,849	3,657,720	5,657,273

cede the elusive prize of cabotage. Of course, as discussed above,[77] the Netherlands's concession of cabotage privileges to American carriers must be seen mainly as a symbolic gesture; the treaty provides its rewards to the United States in other forms, such as sweeping access to landing slots and routing. The United States reciprocated by interpreting ownership restrictions liberally for the Dutch carrier KLM, which was allowed to buy a sizeable stake in Northwest Airlines.[78]

C. *The Open Skies Definition and the Federal Aviation Act: Conflicting Positions on Reciprocity*

It is not a stated policy of the United States to link generous ownership allowances under the Act with liberal terms in bilateral aviation treaties. In fact, DOT, like many administrative agencies, maintains a conservative posture with regard to statu-

[77] *See supra*, text accompanying notes 69–73.

[78] Also among the rewards bestowed upon the Netherlands by the United States was a quick grant to KLM of more liberal access rights to U.S. airports. Application of Koninklijke Luchtvaart Maatschappij d/b/a KLM Royal Dutch Airlines for an amended foreign air carrier permit under section 402 of the Federal Aviation Act of 1958, as amended, Order Issuing Foreign Air Carrier Permit, DOT Order 93-7-14, Docket No. 48,493 (June 24, 1993).

tory construction, claiming that it only executes the will of Congress (as expressed in the black letter of the Act) to limit foreign ownership of U.S. air carriers. According to this restrictive view of administrative discretion, DOT cannot offer concessions on foreign ownership; if the statute sets out the rules more or less inflexibly, DOT may only ascertain whether a foreign airline petitioning to purchase a U.S. carrier meets certain criteria enunciated unequivocally by Congress.[79] Thus, DOT officially conceives of its function as pure fact-finding, and argues that a link between American ownership rules and foreign bilateral aviation treaty concessions is impossible so long as Congress does not change the law.[80]

Of course, most statutes do (and practically must)[81] leave discretion with the agencies charged with their execution, and the Act is no exception. The statute itself suggests that DOT possesses generous discretion in deciding how to execute American aviation law, and the case law that has developed around DOT's supervision of foreign ownership transactions demonstrates the flexibility implied generally by the statute.

The first indicator of this flexibility is the statute itself. The Act declares ten goals for American international aviation policy,[82] each of which reflects a strong theoretical commitment to free trade and economic liberalism.[83] On its face, the open skies definition of 1992 closely reflects these priorities.[84] Thus, the Final Order defining open skies might appear progressive.

However, the open skies definition actually represents a retreat from the most liberal plausible reading of section 1502(b). Subparts (2) through (7) of section 1502(b) point toward a regime largely free of restrictions. The United States could best achieve the statute's stated goals of allowing air charter service with "the

---

[79] Recall that DOT's Final Order defining open skies enunciates DOT policy and that, as a matter of DOT policy, cabotage and ownership do not figure into bilateral negotiations. *See supra* notes 56–58 and accompanying text.

[80] For the same reasons, neither can the Department of State (which cooperates in bilateral aviation negotiations) take advantage of this linkage.

[81] At least two factors necessitate agency discretion. First, unlike a legislature, an administrative agency possesses expertise in its substantive field. Second, vagueness in statutory drafting represents one of the essential tools of legislative compromise; the law-making process is indispensably served by leaving to agencies certain decisions that Congress itself cannot make.

[82] 49 U.S.C. app. § 1502(b)(1)–(10) (1988), codifying section 1102 of the Act, which is reproduced in Appendix A, *infra*.

[83] Note that the Act allows DOT to seek reciprocal rights from its trading partners. 49 U.S.C.A. app. § 1517(c) (West Supp. 1993).

[84] *See supra* notes 43–51 and accompanying text.

fewest possible restrictions," eliminating "operational and marketing restrictions," and "integrat[ing] . . . domestic and international air transportation" by allowing foreigners cabotage and ownership rights. Yet, the open skies definition emphatically excludes cabotage and ownership/control terms, however much section 1502(b) would seem to recommend them.

One way to analyze this exclusion is to envision section 1502(b) as reflecting two countervailing concerns. On one side of the balance are the goals of trade liberalism; Congress desires a liberal, free-market global aviation environment. This demands a sweeping open skies arrangement—i.e., the most liberal set of bilateral aviation treaties possible. On the other side of the balance, however, Congress wants to retain the autonomy to respond flexibly to unfair treatment by our trade partners—in short, to make reciprocity an integral part of aviation diplomacy. Thus, subparts (1), (8), and (9) of section 1502(b) are necessary to ensure that DOT does not parcel out rights to foreign countries without winning commensurate rights in return. Congress states unequivocally that it expects DOT to pursue rights in return for granting rights ("permanent linkage between rights gained and rights given away").[85] The statute calls upon DOT to create "opportunities for carriers of foreign countries *to increase their access to United States points* if exchanged for benefits of similar magnitude for United States carriers."[86] This clause strongly suggests that Congress intends to allow DOT the discretion to offer increased rights to foreign entities as a bargaining chip for greater American rights abroad.[87] In conjunction

---

[85] 49 U.S.C. app. § 1502(b)(8) (1988).

[86] *Id.* (emphasis supplied).

[87] It should be noted, however, that multilateral international agreements limit the granting of cabotage. The Convention on International Civil Aviation confirms the right of any signatory to withhold cabotage rights and forbids granting exclusive cabotage rights to any single state:

> Article 7
> *Cabotage*
> Each contracting State shall have the right to refuse permission to the aircraft of other contracting States to take on in its territory passengers, mail and cargo carried for remuneration or hire and destined for another point within its territory. Each contracting State undertakes not to enter into any arrangements which specifically grant any such privilege on an exclusive basis to any other State or an airline of any other State, and not to obtain any such exclusive privilege from any other State.

Convention on International Civil Aviation, art. 7, 61 Stat. 1180, 1182 (Dec. 7, 1944). This provision should interest those U.S. citizens who oppose the BA/USAir takeover. Thus far, opposition to the takeover has rested entirely upon the *policy* argument that the United States should not grant such a sweeping concession to the United Kingdom

with subparts (1) and (9), subpart (8) of section 1502(b) appears to invite DOT to exercise broad discretion in deciding which aviation rights to extend to foreign entities. Of course, ownership allowances must be included among these rights. But in its final form, the Open Skies Order excluded provisions substantively identical to those in the statute.

The Open Skies Order thus differs from past exercises of agency discretion in two respects. First, the Order was written to apply across all cases (i.e., it did not arise in the limited context of a single case). Second, by excluding cabotage and ownership/control, it has stayed or even reversed the trend toward trade liberalization embodied in prior case precedent and arguably even the statute.

This Article now turns to an analysis of this administrative precedent.

III. The Administrative Case Law

A review of the administrative case law illustrates two facts: DOT has exercised significant discretion in allowing foreigners to purchase interests in U.S. airlines; and the administrative trend has been toward increasingly liberal allowances.

In particular, DOT has the duty and the discretion to define two terms crucial to any foreign entity entertaining purchase of a U.S. airline under the Act: the terms "control" and "citizenship." What follows is a brief analysis of the administrative law precedent that has developed around definition of these terms.[88]

---

without receiving commensurate concessions in return. *Legal* argument has been limited to protesting that ceding control of a U.S. carrier to a foreigner violates section 101(16) of the Act. Opponents might do well to adopt the following argument: allowing the takeover would be equivalent to granting cabotage to the United Kingdom, in violation of Article Seven of the Convention on International Civil Aviation. See *infra* part V, notes 154–215 and accompanying text, for a detailed discussion of the U.S. interests that oppose foreign airline takeovers.

[88] For additional discussions of the precedent underlying the foreign ownership issue, see Jeffrey Donner Brown, *Foreign Investment in U.S. Airlines: What Limits Should be Placed on Foreign Ownership of U.S. Carriers?*, 41 Syracuse L. Rev. 1269 (1990) (summarizing proposals to change the current law); John T. Stewart, Jr., *United States Citizenship Requirements of the Federal Aviation Act—A Misty Moor of Legalisms or the Rampart of Protectionism?*, 55 J. Air L. & Com. 685 (1990) (discussing the legislative history behind the citizenship requirement); James E. Gjerset, *Crippling United States Airlines: Archaic Interpretations of the Federal Aviation Act's Restriction on Foreign Capital Investments*, 7 Am. U. J. Int'l L. & Pol'y 173 (1991) (discussing congressional intent behind the Act and proposing revisions to relax citizenship requirements).

I will argue that the cases were leading the American ownership regime along a course of liberalization, in response to the evolving state of civil aviation, but that the most recent developments in U.S. policy have departed from this course.

The cases show that DOT exercises substantial power over the definitions of citizenship and control under the present statute. Moreover, the case law shows that DOT has been moving sporadically toward an ever more liberal regime of foreign ownership. DOT's substantial definition-making discretion further implies that, contrary to outward appearances and official disclaimers, the agency does indeed hold a bargaining chip when it negotiates bilateral aviation terms with foreign sovereigns. The correlation between liberal bilateral treaties and generous American ownership concessions strongly suggests that DOT has in fact placed ownership regulation on the negotiating table in return for (or as a reward for) access to foreign airspace, landing slots, and other open-skies rewards.[89] In its power to define the two most crucial terms concerning airline ownership—citizenship and control—DOT possesses an asset that it can hold out to foreign sovereign powers, whose airlines seek permission to purchase stakes in U.S. carriers. Thus, though Congress is hardly ready to extend cabotage to foreign airlines, it nonetheless gave DOT substantial opportunity to reward foreign powers that grant the United States liberal rights under bilateral aviation treaties.[90] The cases illustrate that DOT has, in the past, employed its

---

[89] The case that linkage does in fact constitute U.S. policy received great support in the January 1991 statement of Transportation Secretary Samuel K. Skinner to the National Press Club:

> I am pleased to announce that the Department of Transportation is issuing an order today which will allow KLM to maintain a significant investment in Northwest Airlines. It is appropriate that this decision involves an investment by a carrier whose government has entered into a very liberal aviation relationship with the United States. Our decision clarifies, and in significant ways relaxes, the restrictions on foreign investment that are the product of past interpretations.

Samuel K. Skinner, Remarks to the National Press Club 5 (Jan. 23, 1991) (on file with the *Harvard Journal on Legislation*). Notably, Secretary Skinner was reluctant to describe DOT activity as "law-making." Instead, he categorized the momentous change in DOT precedent (marked by the Northwest/KLM takeover) as a mere change in "interpretation."

[90] Although access to ownership of U.S. airlines constitutes a valuable concession that DOT can extend, it is not the only one. Power to grant access to major gateway cities amounts to another significant bargaining chip in bilateral aviation diplomacy. European countries generally have very few major airports that receive international traffic (and thus very few places to grant entry to foreigners), but the United States has over a dozen major cities that receive significant amounts of international traffic. *See* Nomani, *supra* note 35.

discretion to liberalize the airline ownership regime; however, recent DOT actions have threatened to impede or even reverse the trend toward liberalism. As agency action has approached the limits of the statutory mandate, free trade progress has been contained.

A. *Defining the Terms—"Citizenship" and "Control"*

Although defining the terms "citizenship" and "control" is central to international aviation law, it also lies at the heart of other, diverse areas of regulation. Two areas in particular—government contracting under the Public Works Employment Act of 1976 and licensing by the Federal Communications Commission—exemplify how citizenship and control requirements can pose obstacles in structuring business transactions.

The Public Works Employment Act of 1976 endeavors to rectify past mistreatment of minority citizens by breaking cycles of continuing discrimination in public works contracting. The statute provides that

> [e]xcept to the extent that the Secretary determines otherwise, no grant shall be made under this Act for any local public works project unless the applicant gives satisfactory assurance to the Secretary that at least 10 per centum of the amount of each grant shall be expended for minority business enterprises. For purposes of this paragraph, the term "minority business enterprise" means a business at least 50 per centum of which is owned by minority group members or, in case of a publicly owned business, at least 51 per centum of the stock of which is owned by minority group members. For the purposes of the preceding sentence, minority group members are citizens of the United States who are Negroes, Spanish-speaking, Orientals, Indians, Eskimos, and Aleuts.[91]

The Supreme Court upheld the constitutionality of the above provisions in *Fullilove v. Klutznick*.[92] A New York case, *Shore Air-Conditioning Co. v. New York City Health & Hospitals Corporation*,[93] addressed the issue of what constitutes a *bona fide* "minority enterprise" under this statute. In both cases, as in

---

[91] 42 U.S.C.A. § 6705(f)(2) (West 1983).

[92] Fullilove v. Klutznick, 448 U.S. 448 (1980).

[93] Shore Air-Conditioning Co. v. New York City Health & Hospitals Corp., 93 Misc.2d 15, 401 N.Y.S.2d 948 (1977).

cases arising under the aviation regulations of foreign ownership, defining enterprise control is a central consideration.

Transactions that involve changed ownership and foreign investment lead to unique regulatory questions outside the aviation field as well. For example, section 310 of the Federal Communications Act ("FCA")[94] restricts ownership of operating licenses in the wire and radio communications industry. In particular, the statute provides that no radio station license shall be granted to or held by

> (3) any corporation of which any officer or director is an alien or of which more than one-fifth of the capital stock is owned of record or voted by aliens or their representatives or by a foreign government or representative thereof or by any corporation organized under the laws of a foreign country; [or]
> (4) any corporation directly or indirectly controlled by any other corporation of which any officer or more than one-fourth of the directors are aliens, or of which more than one-fourth the capital stock is owned of record or voted by aliens, their representatives, or by a foreign government or representative thereof, or by any corporation organized under the laws of a foreign country, if the [Federal Communications] Commission finds that the public interest will be served by the refusal or revocation of such license.[95]

Furthermore, section 222(d) of the FCA limits alien ownership of capital stock in a radio or other communications enterprise to one-fifth.[96] These regulations of communications closely resemble section 101(16) of the Act, which defines "citizen of the United States" and thereby sets forth the definitional limits on ownership of U.S. air carriers. As with the "fit, willing, and able" requirement under the Act, a "public convenience, interest and necessity" showing is necessary under the FCA.[97]

Tension over interpretation of the term "control" in the context of the Act has mirrored that in the context of the FCA. The Federal Communications Commission ("FCC") has chosen to define control quite broadly, and although this choice has been challenged by the industry, the courts have supported the FCC. An early D.C. Circuit case held that the term "control" is not

---

[94] 47 U.S.C.A. § 310 (West 1991).
[95] 47 U.S.C.A. § 310(b)(3)–(4) (West 1991).
[96] 47 U.S.C.A. § 222(d) (West 1991).
[97] An early case, Colonial Broadcasters, Inc. v. Federal Communications Comm'n, 105 F.2d 781, 783 (D.C. Cir. 1939), held that agency discretion under the "public convenience and necessity" test passed constitutional muster.

necessarily limited to "legal control" in a formal sense, but may also extend to include "actual control" through equity holding or other relationships.[98] Subsequent FCC holdings reiterated that formal legal control is not the only influence that may render an entity ineligible for licensing.[99] The Supreme Court has upheld the FCC's citizenship and control limitations as constitutional.[100]

Citizenship and control definitions are central to the rule making of myriad administrative agencies. Like the FCC and other agencies, DOT has been compelled to develop case law devoted to defining these two terms. What follows is a summary of this case law as it relates to the foreign takeover question.

B. *From* Daetwyler *to Northwest/KLM*

Analysis of the administrative law governing foreign takeovers of U.S. airlines must begin with an early landmark decision by CAB. The *Daetwyler* case established clear standards for determining citizenship and control of a corporation under the Act:

> We believe that the intent of Congress in enacting Section 101(13) [now 101(16)] was to insure that air carriers issued licenses by the United States as U.S. air carriers would be owned and controlled by citizens of the United States. To adopt the examiners' supposition that a construction to the effect that either ownership *or* control alternatively might suffice would defeat the congressional purpose of the citizenship provision.[101]

There was no dispute in *Daetwyler* that seventy-five percent of the stock of the company in question, Interamerican, was nominally owned by U.S. citizens and that two-thirds of its board of directors and other managing officers were U.S. citizens.[102] Interamerican appeared to meet the minimum standards of section 101(16). The dispute concerned whether Mr. Daetwyler, a Swiss citizen, actually controlled the air carrier.

---

[98] Lorain Journal Co. v. Federal Communications Comm'n, 351 F.2d 824 (D.C. Cir. 1965), *cert. denied*, 383 U.S. 967 (1966).

[99] *See, e.g.*, *Re* Arthur A. Cirilli, 3 F.C.C.2d 893 (1966); *Re* 1400 Corp., 11 F.C.C.2d 321 (1968); *Re* Data Transmission Co., 44 F.C.C.2d 935 (1974).

[100] Federal Communications Comm'n v. National Citizens Comm. for Broadcasting, 436 U.S. 775 (1978).

[101] Willye Peter Daetwyler d/b/a Interamerican Airfreight Co., Foreign Permit, 58 C.A.B. 118, 120–21 (1971) (emphasis in original).

[102] *Id.* at 119.

CAB decided that, in fact, Interamerican was dominated by Mr. Daetwyler:

> The Daetwyler control is reflected by the fact that the corporation is the direct successor to the Daetwyler forwarding company and created wholly at the instance of Daetwyler. Daetwyler has, in fact, retained 25 percent stock ownership and represents one-third of the board of directors of the corporation, the maximum control technically permitted under the literal terms of section 101(13) [now 101(16)]. Moreover, the corporation will continue to do business as part of [a] system of Daetwyler-controlled companies. In addition, there not only exist the close personal relationships between Daetwyler and his attorney and employees who constitute the remaining stockholders and officers and directors but also, with the exception of Daetwyler's longtime attorney in both corporate and personal matters, the U.S. citizen stockholders, officers, and directors are employees of other admittedly Daetwyler-controlled corporations, with interests which could be expected at the very least to represent a substantial inducement upon such employees to follow the wishes of Daetwyler.[103]

Although Mr. Daetwyler had structured his enterprises so that they satisfied the letter of the statute, CAB rejected this formalistic adherence to the law, ruling that the enterprise also had to conform to the spirit of the statute as defined by the administrative agency. In a statement that would prove important in subsequent control cases, CAB held that

> where an applicant has arranged its affairs so as to meet the bare minimum requirements set forth in the Act, it is the Board's view that the transaction must be closely scrutinized and that the applicant bears the burden of establishing that the substance of the transaction is such as to be in accordance with the policy, as well as the literal terms of the specific statutory requirements.[104]

With the burden of proof thus allocated to the applicant's disadvantage, CAB concluded that Mr. Daetwyler and his business did not qualify to receive a license as an American air carrier.[105]

---

[103] *Id.* at 120.

[104] *Id.* at 121.

[105] For subsequent related cases citing *Daetwyler*, see, *e.g.*, Premiere Airlines, DOT Order 82-5-11, Docket Nos. 38,965, 39,158, at 4 (May 5, 1992); Challenge Air Cargo, DOT Order 91-4-32, Docket No. 46,489, at 2 (July 15, 1993); Pride Air, DOT Order 87-5-59, Docket No. 42,139, at 2 (May 21, 1987); Key Airlines, DOT Order 84-4-83, at 10-11 (Apr. 23, 1984); Intercontinental Airways, DOT Order 81-9-64, Docket No. 39,948, at 6 (Sept. 11, 1981); Northwest-Wings, DOT Order 89-9-51, Docket No.

In *Intera Arctic Air Services*, DOT reasserted the *Daetwyler* finding that merely fulfilling the letter of the control/ownership statute does not render a certification applicant immune from scrutiny:

> [F]oreign influence may be concentrated or diffuse. It need not be identified with any particular nationality. It need not be shown to have sinister intent. It need not be continually exercisable on a day-to-day basis. If persons other than U.S. citizens, individually or collectively, can significantly influence the affairs of [the carrier], it is not a U.S. citizen.[106]

In *Page Avjet Corporation*,[107] CAB adopted even more stringent definitions of citizenship and control than it had in *Daetwyler*. In *Page Avjet*, although all voting stock and over seventy-five percent of the nonvoting stock lay in the hands of U.S. citizens, CAB nevertheless held that the corporation was subject to foreign control. "We have found control to embrace every form of control and to include negative as well as positive influence," CAB concluded. "We have recognized that a dominating influence may be exercised in ways other than through a vote."[108] The nonvoting foreign shareholders, while few in number, held the power to veto major company decisions, including any decisions pertaining to company consolidation, merger, acquisition, or liquidation.[109]

From this stringent baseline, however, DOT began progressing toward liberalism in ownership rules. In 1990, moderating slightly the tough standards enforced in *Page Avjet* and other progeny of *Daetwyler*, DOT allowed SAS to hold 16.8% of the holding company of Continental Airlines and to keep three of its officers on the holding company's 15-member board of directors.[110] The voting power of the SAS shares would amount to 18.4%.[111] The DOT order clearly set out a two-part test for

---

46,371, at 4 (Sept. 29, 1989); Hutchinson Auto and Air Transport, DOT Order 91-8-15, Docket No. 47,690, at 11 (Aug. 30, 1991).

[106]Intera Arctic Services, Inc., DOT Order 87-8-43, Docket No. 44,723 (Aug. 18, 1987).

[107]Page Avjet Corporation, DOT Orders 83-7-5, 82-8-41, Docket No. 40,905 (July 1, 1983).

[108]*Id.* at 3.

[109]*Id.* at 4. CAB proceeded to give Page Avjet Corporation 30 days to submit a plan of reorganization that comported with CAB's interpretation of the Act. Failing such submission, the company was ordered to cease operations within 60 days. *Id.* at 5.

[110]Acquisition of Stock in Continental Airlines Holdings, Inc., by Scandinavian Airlines System, DOT Order 90-9-15, Docket No. 47,167 (Sept. 12, 1990).

[111]*Id.*

determining citizenship and control, developed from DOT's interpretation of section 101(16) of the Act. The first prong of this test requires that the president and two-thirds of the board of directors and other managing officers must have U.S. citizenship, and that at least seventy-five percent of the outstanding voting stock must reside in U.S. hands. Under the second prong, the carrier must be controlled by U.S. citizens.[112]

Continental and SAS satisfied the test by implementing several guarantees that SAS would not acquire the leverage to control Continental. Three special provisions addressed concerns that SAS might control the Continental Holdings, Inc. board of directors:

(1) Neither SAS nor its nominees would be allowed to propose a slate of directors or to solicit proxies for the election of directors different from those proposed by the board;[113]

(2) SAS agreed that its three designated directors would excuse themselves from participating in any matter which might have a "direct and predictable" effect (a) on actual or potential competition with Continental, or (b) on bilateral or multilateral aviation negotiations to which SAS and/or the countries of Denmark, Norway, or Sweden are a party;[114] and

(3) Although SAS would hold 100% of one class of preferred stock in the Continental holding company and the affirmative vote of the majority of the holders of that class of stock would be required to approve certain transactions, SAS agreed to cast its votes on such transactions in accordance with the recommendations of the board of directors of the holding company.[115]

These precautions against *de facto* foreign control satisfied DOT; therefore, the agency granted SAS and Continental the right to expand their relationship. In developing a way around the *Daetwyler* dilemma, DOT had begun to adapt its interpretation of the law to suit better an increasingly international market.

The most recent precedent in the line of cases prior to the BA/USAir transaction arose out of a purchase of Northwest Airlines stock by KLM. In this case, DOT assumed its most

---

[112] *Id.* at 4.
[113] *Id.* at 3.
[114] *Id.*
[115] *Id.*

liberal stance. The transaction involved KLM, the Dutch carrier, and Wings Holdings, Inc., a holding company for Northwest Airlines. The original Fitness Submission by Wings was made on September 22, 1989. Under its terms, KLM would hold less than five percent of the voting stock of Northwest (through Wings Holdings), but the Dutch carrier would own fifty-seven percent of Northwest's total equity. Also, KLM would gain the right to name one of the twelve members of Northwest's board of directors and to appoint a special three-person committee to advise on the general management of Northwest's financial affairs.

DOT responded to this Fitness Submission with a Consent Order on September 29, 1989.[116] The Consent Order directed as follows: that the parties reduce KLM's equity interest in Northwest to a sum not in excess of twenty-five percent of total equity;[117] that KLM dissolve its special advisory committee;[118] and that certain restrictions (similar to those stipulated in the Continental/SAS transaction) be put in place over KLM's appointees to the Northwest board of directors.[119]

Operating for some time under the Consent Order and under various extensions of time for KLM divestiture, KLM and Northwest filed a petition with DOT on January 15, 1991, requesting four amended terms: (1) termination of the requirement to reduce KLM's total equity to twenty-five percent; (2) permission for KLM to hold forty-nine percent of the equity in Northwest, including 10.544% of the voting interest; (3) permission for KLM to designate three members of the Wings board of directors; and (4) removal of certain financial reporting conditions contained in the original order. DOT granted the request to allow KLM to increase its equity to forty-nine percent and to appoint three members of Wings' board.[120] DOT reiterated, however, that a twenty-five percent limit still remained when foreign ownership of voting stock was concerned.[121]

Despite its limits, the KLM/Northwest decision marked a new liberalism in interpreting the restrictions on foreign takeovers of U.S. air carriers. In contrast with *Daetwyler* and its successors,

---

[116] Acquisition of Northwest Airlines, Inc. by Wings Holdings, Inc., Consent Order, *supra* note 16.

[117] *Id.* at 2.

[118] *Id.* at 5.

[119] *Id.* at 6.

[120] Acquisition of Northwest Airlines, Inc. by Wings Holdings, Inc., Order Modifying Conditions, DOT Order 91-1-41, Docket No. 46,371 (Jan. 23, 1991).

[121] *Id.* at 5.

the KLM/Northwest and SAS/Continental precedents marked a turn toward free trade in air commerce. Unfortunately, further liberalization would prove much more difficult to achieve.[122]

C. *The Beginnings of Reaction—The Northwest/KLM Interpretation Meets Protest*

Perhaps signaling that DOT was on the verge of overstepping the permissible bounds of the Federal Aviation Act, the Northwest/KLM transaction generated petitions in opposition from several major U.S. airlines and from the Air Line Pilots Association. Among the airlines, Delta and America West specifically cited reciprocity in bilateral aviation treaties as a concern surrounding the transaction. America West suggested the following:

> As the Department [of Transportation] knows all too well from numerous bilateral negotiations that have been adversely influenced by a dominant air carrier in a particular foreign country that is not eager to compete, most foreign air carriers are not aggressive competitors (particularly with respect to price)—not an inconsequential consideration when improved competition is one of the expressed goals of those advocating revisions to the foreign investment laws. [T]he existing bilateral agreements between the U.S. and other countries represent a careful balancing of benefits between the two countries. Simply permitting a carrier in a foreign country to "buy" its way into the lucrative domestic U.S. market by a unilateral change in U.S. law will surely undermine the integrity of those agreements.[123]

Delta also expressed concern over the KLM/Northwest transaction. "[G]rave implications [could arise] because foreign carriers could be allowed virtually unlimited access to U.S. markets while U.S. carriers would be prevented from enjoying free access to [foreign] markets."[124]

---

[122] Although section 401 of the Act may be interpreted to allow a foreign entity to own up to 49% equity and 25% control of a U.S. airline, further extension of that standard (in favor of foreign ownership) would be of dubious legality. It appears that further allowances for foreign ownership could be derived only from Congress. As will be discussed *infra*, part VI.B, notes 237–246 and accompanying text, one member of Congress has proposed revising or replacing the statute.

[123] Answer of America West Airlines, Inc. to Petition of Northwest Airlines, Inc. for Modification of Order 89-9-51, Docket No. 46,371 (Jan. 18, 1991).

[124] Answer and Objection of Delta Air Lines, Inc., to Petition of Northwest Airlines, Inc. for Modification of Order 89-9-51, Docket No. 46,371 (Jan. 18, 1991).

The protests of Delta, America West, and others notwithstanding, the KLM/Northwest decision marked a turning point in DOT rule making on foreign takeovers of domestic air carriers. The decision defined much of the legal debate that would later surround citizenship and control in the BA/USAir transaction—opponents of the USAir takeover would fight to distinguish the USAir case from the Northwest/KLM precedent. However, broader issues of public policy would become relevant to the BA/USAir takeover effort.

The sensitivity of the citizenship and control issues and the conflicting influences of the concerned parties have produced a pronounced indecision over the trade liberalism of DOT's aviation policy. This indecision stems from the fact that the administrative case law has reached the outer limits of trade liberalism as permitted by the governing statute; agency officials now struggle against an outdated law.

IV. At the Limits of the Federal Aviation Act: Agency Uncertainty and the Open Skies Order

The administrative law governing ownership and control of domestic air carriers suffers from flux and uncertainty. Unofficial statements of agency officials provide an important source by which to gauge policy direction.[125] Though they do not enjoy the status of law, these statements serve as a barometer for agency direction and ethos. In particular, statements by the two Secretaries of Transportation who served under President Bush—Samuel K. Skinner and Andrew H. Card—reveal contradictions and uncertainty in DOT's stance toward foreign takeovers.

The Open Skies Order of August 1992, on the other hand, is an official policy statement that carries much more legal weight than the Secretaries' pronouncements. Yet, upon careful examination, the Open Skies Order is no more reassuring a source of direction than the often contradictory unofficial statements of the DOT chiefs. In fact, the Open Skies Order, rather than advancing the goal of aviation liberalism, may lead to a reversal of past gains. This retrenchment at DOT appears particularly

---

[125] *See* Jerry L. Mashaw & David L. Harfst, *Inside the National Highway Traffic Safety Administration: Legal Determinants of Bureaucratic Organization and Performance*, 57 U. Chi. L. Rev. 443, 443–45 (1990).

inappropriate when compared to the progress toward free aviation—at least *in principle*—that the European states have made.[126]

A. *The Skinner and Card Statements*

Addresses by the Secretaries of Transportation and other DOT officials often reveal the underlying justification behind DOT policies and hint at future directions. Although these informal statements have generally paralleled the case law trend toward liberalization, lately the commitment to increased liberalization has wavered. In some instances, DOT officials have even suggested that the trend toward liberalization will go no further. Remarks at airline industry conferences by Bush Administration Transportation Secretaries Samuel K. Skinner (1987–1991) and Andrew H. Card (1991–1993) are particularly noteworthy.

Transportation Secretary Samuel K. Skinner issued a watershed statement in recent U.S. aviation history at the National Press Club on January 23, 1991. In his comments, the Secretary stressed the need to maintain leverage against unfair restrictions on U.S. aviation abroad:

> [Security and trade protection] requirements also represent a necessary response to the highly regulated nature of international aviation, in which landing rights are parceled out through narrowly drawn bilateral agreements. If we are to maintain the leverage we need to open more foreign markets to U.S. carriers, we have to maintain oversight of the extent to which foreign carriers enjoy access to *our* market through investments.[127]

Although this statement should not be interpreted to suggest that Skinner was summoning a new era of protectionism and a reversal of past liberalization, his mention of maintaining "leverage" and responding to "narrowly drawn" aviation agreements does reveal caution. In the same speech, however, the Secretary emphasized that a new era of globalized aviation demands a more liberalized U.S. approach to international aviation diplomacy:

---

[126] Clearly, the European states also have deeply entrenched trade protections in practice. The point, however, is that the Open Skies Order—an important exposition of free trade aspirations—falls short of articulating the liberalization ideal. Draft versions of its counterpart in Western Europe appear, at least implicitly, to more closely approach that ideal. *See supra*, notes 59–65 and accompanying text.

[127] Skinner, *supra* note 89 (emphasis in original).

> [I]t is essential in a time when international aviation is changing so much, that we guard against going beyond the restrictions necessary to achieve these important objectives. We have therefore reexamined our application of the "control" test to ensure that it's fully consistent with today's realities.
>
> It is likely that *we will see, over time, a genuine globalization of airline companies*, much as we have seen in so many other industries. That means more than merely landing in as many countries as possible. It means establishing a genuine, fully competitive presence in all of the world's most important aviation markets.
>
> The first steps toward that development can be seen today in the important linkages that airlines from different countries are forging with each other in a quest for more effective marketing tools. One means of establishing such linkages is the medium of investment.
>
> It is essential that our international aviation policy acknowledge this trend and respond to it effectively. *U.S. carriers must stay ahead of the pack in the move toward globalization.*
>
> We have concluded that one of the important steps we can take in this connection is to create an environment more receptive to foreign investment in the U.S. airline industry. To that end, we have decided to refine our interpretation of the foreign ownership statute. We will no longer consider total foreign *equity investment*, by itself, to be an indicator of foreign *control* if it is less than 50 percent of total equity. And we will not consider debt financing obtained from foreign sources as a potential means of control, provided the loan agreement confers no extraordinary rights on the foreign lender.
>
> Regarding foreign directors and officers, there is no need to formulate more stringent limits as we have done in the past. Instead, we will allow foreign membership on airline boards to reflect their voting stock ownership—as long as it satisfies the statutory ceiling, and as long as representatives of foreign airlines do not take part in decisions affecting competition. Our recent order relating to the transaction between Continental and Scandinavian Air was consistent with this approach.[128]

This statement indicates that the terms of the Northwest/KLM transaction had come to represent the new and accepted rule for agency interpretation of the Act. But at the same time, Secretary Skinner also offered reassurances that security and trade fairness

---

[128] *Id.* at 4–5.

would remain significant concerns in evaluating proposed foreign takeovers of U.S. carriers. Skinner's remarks, while revealing his apparent wish to promote liberalization, also sought to reassure the opposition interests and to acknowledge the limits of DOT discretion.

Skinner's hedging on the issue of liberalization was even more pronounced in his return to the older, puzzlingly conservative view that DOT wanted no part in any legislative evolution of aviation law:

> In announcing these changes, I want to emphasize that I am not now calling for any change in the statute, since I am not persuaded yet that such changes are warranted, and because statutory amendments require careful consideration and debate within the Congress. Thus, existing ceilings on foreign ownership of voting stock and citizenship requirements for officers and directors will have to be complied with.[129]

This uncertainty stands in marked contrast to DOT's one-time steady (albeit cautious) commitment to liberalizing civil aviation. From *Daetwyler* to the Northwest/KLM case, DOT pushed the interpretation of the statute to ever more liberal reaches, thereby recognizing the evolution of international aviation commerce and diplomacy. The logical, if not necessary, next step would be to push for an interpretation even more liberal than the Northwest/KLM Order. But now, at the frontier of the existing statute, DOT has begun to waver in its promotion of free trade.

Secretary Skinner reiterated many of the above points in his address to the Second Annual Airline Industry Conference in Washington, D.C. Skinner's talk, entitled "Global Competitiveness: New Thinking for the 1990's," emphasized the need to invite as much capital as possible into the U.S. aviation industry. In addition, Skinner suggested that DOT play a role in fostering change in the legislative mandate:

> We have also identified airlines' inability to raise sufficient capital as a possible barrier to vigorous competition. So we've refined our interpretation of the foreign control test to permit increased foreign investment, and *have advocated legislative changes to allow non-U.S. citizens to own up to 49 percent of an airline's voting stock.*[130]

---

[129] *Id.* at 5.

[130] Samuel K. Skinner, "Global Competitiveness: New Thinking for the 1990's," Address before S.H. & E. Second Annual Airline Industry Conference 4 (Nov. 21, 1991) (emphasis supplied) (transcript on file with the *Harvard Journal on Legislation*).

Instead of eschewing agency involvement in the legislative process (DOT's former policy), DOT's highest officer declared that the agency should promote liberalization via statute, as well as in the administrative law. Such comments, coupled with the recent case law, strongly suggest that the underlying conviction of DOT is to favor liberalization and conform American aviation policy to the changing global business environment—although DOT, operating at the frontier of its organic statute, is still beset by uncertainty.

The post-Northwest/KLM uncertainty continued under Skinner's successor, Andrew H. Card. The new Secretary's statements continued to suggest that although DOT officials favored a policy of liberalization, they were tempered by a reluctance to push beyond the then-new liberal interpretation of the Act. Continuing the pattern of wavering set by his predecessor, Secretary Card discouraged DOT promotion of legislative amendments to aviation policy. At an address to the Aero Club Lunch on March 31, 1992, Secretary Card vowed a new and aggressive open skies initiative:

> The Bush Administration has already taken some important steps to advance the cause of more open international aviation markets. The steps range from easing restrictions on foreign investment in U.S. airlines—*within the limits of the current law*—to the creation of a program that opens new U.S. gateway cities to foreign carriers without insisting on prior negotiations.
>
> Yet more needs to be done. We need to stimulate greater interest in creating an even more market-oriented international aviation environment. Today, we are prepared to do just that—to take another step in the direction of open markets.
>
> . . . .
>
> Today, I am announcing that we will now offer to negotiate open skies agreements with all European countries willing to permit U.S. carriers essentially free access to their markets.[131]

Thus, an avowedly aggressive policy of free trade still stopped short of progressing beyond what had already been achieved. Uncertain over the responses from industry and other interests, DOT qualified its commitment to trade liberalization by assuring that it would remain disengaged from the legislative process.

---

[131] Card, *supra* note 37, at 5–6 (emphasis supplied).

Secretary Card also addressed whether DOT should link private ownership cases, such as the BA/USAir transaction, to treaties on a national level. The Secretary stated, "I have not allowed for a formal linkage of a private matter and a government to government negotiation."[132] He went on to remark, however, that the BA/USAir transaction was more than a business deal between private parties—it constituted "a watershed transaction."[133] In connection to that transaction, the Secretary suggested that in return for American approval of the BA/USAir deal, the United Kingdom should accord the United States more permissive terms in the bilateral aviation treaty between the two countries.[134]

To summarize, the statements of Secretaries Skinner and Card reveal an evolution toward liberalization in American aviation policy, but they also highlight the officials' hesitance to acknowledge their role in promoting that evolution. This hesitance reflects the uncertainty that has beset the DOT as it pushes up against the outer limits of its statutorily allowed discretion.

B. *The Role of Private Interests in Aviation Diplomacy*

A policy of linkage is latent in past U.S. action and was apparently on the minds of Bush Administration officials. The contradictory signals that Secretary Card and other DOT officials sent on the issue of linkage, however, underscore a central tension in U.S. aviation diplomacy. On the one hand, overt linkage between aviation treaty terms and ownership allowances would enable the United States to parlay control of the profitable U.S. aviation market into freer worldwide trade on a country-by-country basis. Additionally, it would clarify an area of the law that has been increasingly clouded by uncertain agency action. On the other hand, linkage is disfavored by politically powerful lobbies that oppose open American skies and foreign ownership of U.S. airlines.[135] In addition to this political opposition, there is also a hesitance to link purportedly private matters to national diplomatic initiatives. But in at least two respects, this private-versus-public impediment to an open declaration of linkage is unjustified.

---

[132] Agis Salpukas, *Official Discerns a Lever in the British-USAir Bid*, N.Y. TIMES, Sept. 17, 1992, at D5.

[133] *Id.*

[134] *Id.*

[135] *See* part III.C, *supra* notes 123–124 and accompanying text.

First, an analysis of foreign airline industry practice casts serious doubt on the cogency of the public-private dichotomy. Lufthansa, Air France, and a dozen other major carriers around the world operate as quasi-governmental entities, and some are full-fledged state corporations.[136] Even some airlines that are ostensibly fully privatized enjoy a degree of government protection that casts doubt upon the purity of their "privateness." In view of this worldwide perspective on civil aviation, DOT's emphasis on separating "public" and "private" law is not well-founded.

Second, it is a fiction to maintain that individual private interests have—or should have—no role in determining U.S. policy. A review of the KLM and SAS transactions makes it difficult to believe that DOT's decisions on these "private" transactions were not influenced by the generous treaty terms that the Netherlands and Sweden extended to the United States. It is a well-known fact that the leaders of important industries contribute, through lobbying, consulting, and even personal contact, to shaping national policy.[137]

Furthermore, the proposition that private business relations are relevant to policy determinations finds support in the Act itself. Section 1102(a), pertaining to international agreements, provides that DOT will not hamper any U.S. carrier from fulfilling obligations that it acquired through contract with a foreign country, so long as such contract does not compromise the public interest.[138] This provision indicates that the United States recognizes and will allow the enforcement of private law between U.S. air carriers and foreign governments. Therefore, the law does not

---

[136] Until quite recently, the state-run airline represented the norm, rather than the exception. Until less than four years ago, Air Canada was state-run. Japan Airlines was privatized in 1987, and the Australian government has only recently sought to liquidate its near-100% interest in Quantas. Air France constitutes a domestic state monopoly. A state holding company (Instituto per la Ricostruzione Industriale) owns Alitalia, and Iberia is the flag carrier of Spain. Thai Airways International and Saudi Arabian Airlines (two major air carriers outside the leading industrial countries) are owned by their respective governments. And Aeroflot—still the world's largest airline (at least in terms of the sheer size of its fleet)—remains a government entity, if no longer a monopoly over its homeland's airspace. *See Snapshots of the World's Major Airlines*, WALL ST. J., Jan. 14, 1992, at A8.

[137] For example, a memorandum entitled "Opposing the British Airways/USAir Merger," prepared by a prominent lobbying firm, states explicitly that the major U.S. air carriers intend to influence local, state, and federal officials. See *infra* notes 161–167 and accompanying text for a more detailed discussion of this memorandum.

[138] Section 1102(a) of the Act is codified at 49 U.S.C. app. § 1502(a) (1988). *See also* Appendix A, *infra*.

support Secretary Card's position that private and government matters are and should be kept distinct.[139]

Furthermore, these critiques of the current U.S. policy are undermined by positive considerations that support private-public linkages. Linking a private foreign corporation's petition to DOT with its home government's bilateral aviation treaty can be valuable. For example, by making it clear to Lufthansa that a future bid for a stake in Continental is dependent upon the openness of the German government to U.S. aviation, DOT in effect drafts Lufthansa as an American lobbying agent in Germany. Thus, the linkage proposed in Part II is realized (albeit in veiled and piecemeal ways). If the linkage were clarified, broadly applied, and enhanced through statutory revision, foreign corporations could become useful instruments of our aviation diplomacy. The United States would no longer have to depend solely upon *quid pro quo* arrangements between the formal representatives of each country in diplomatic meetings; the United States would also have a number of foreign airlines working on its behalf. A foreign airline seeking access to the U.S. market via ownership would bring pressure to bear on its government, and such pressure—coming from a native source rather than solely from U.S. diplomats—would exert influences that negotiations in a vacuum might otherwise lack.

Though DOT has largely refrained from openly involving private interests in aviation policy, the agency has exhibited uncertainty in this matter, much as it has regarding the liberalization of American aviation law in general. In his March 1992 speech at the Aero Club, Secretary Card invited the assistance of the aviation industry in defining open skies. But by the late spring of 1992, the message conveyed by DOT was uncertain and even contradictory in two respects.

First, although DOT had announced that it would press aggressively to further the evolution of the rules governing U.S. aviation, particularly the rules governing foreign access and ownership, the agency also indicated that it would stop short of pressing to change the statute that clearly limits such evolution.

---

[139] Section 1102(c) of the Act further corroborates the proposition that there is no legal restraint prohibiting the U.S. government from conducting its aviation diplomacy with a heightened awareness of the transactions, status, and relationships of individual U.S. air carriers. In fact, § 1102(c)—"Consultation with affected groups"—suggests that coordinating aviation diplomacy closely with individual American aviation interests is even legally mandated. *See* Appendix A, *infra.*

In short, the agency promised to be the champion of liberalization, but would not fight a restrictive legislative mandate.

Second, while DOT stated that linkage between private transactions and diplomacy was inappropriate, the agency simultaneously invited private interests to help shape a major diplomatic trade initiative. From this background of contradiction and uncertainty emerged the Open Skies Order of August 1992.

C. *The Open Skies Order in Context—Liberalization or Reversal?*

The fruit of the open skies initiative announced by Secretary Card at his Aero Club speech was the DOT Final Order defining open skies.[140] Although this order appears to promote the liberalization of U.S. aviation rules, its definition of open skies establishes boundaries that arguably represent a reversal in aviation free trade.

The Open Skies Order clearly enunciates what hitherto could only be ferretted out of contradictory pronouncements by DOT Secretaries: the United States does not formally link ownership access in America to bilateral treaty permissiveness over a trading partner's territory.[141] Though Secretaries Skinner and Card clouded the issue of linkage, the Final Order makes it clear that DOT does not overtly take this approach.

The Final Order defining open skies consists of eleven points.[142] Although public comments solicited during the drafting period suggested that cabotage and ownership/control terms be added to these eleven points, both of these terms were expressly rejected by DOT.[143] Applying the canon *expressio unius est exclusio alterius*,[144] it would appear that the Final Order therefore forecloses the linkage between aviation access and ownership/control that this Article proposes. In comparison to past DOT actions, the Open Skies Order arguably narrows DOT's man-

---

[140] The open skies definition is discussed *supra*, notes 43–58 and accompanying text, in relation to bilateral diplomacy.

[141] The elements of the open skies definition as enunciated by the Final Order are substantively identical to those in the DOT Order Requesting Comments. *See supra*, notes 43–50 and accompanying text.

[142] *Id.*

[143] Defining "Open Skies;" Final Order, *supra* note 51.

[144] "The expression of one thing is the exclusion of the other."

date and thereby, ironically, narrows the possibilities for open skies.

Prior to the Open Skies Order, there had been no clear demarcation of DOT discretion outside the text of the Act. Although the administrative case law constrained discretion, DOT had proven itself ready to quite freely liberalize the rules on ownership in cases from *Daetwyler* to the Northwest/KLM transaction. The case law did not bind the agency to a single set of interpretations, and was settled only in the sense that it was moving in a liberal direction. Due to this flexibility, DOT was able to obtain more open bilateral treaty terms with important trade partners. Although the terms of the Federal Aviation Act clearly precluded some transactions,[145] DOT's discretion was broad enough to allow it to put ownership and control on the bargaining table in return for liberalization of bilateral aviation treaties—as exemplified by DOT's treatment of Dutch and Scandinavian carriers in light of U.S. bilateral treaties with Sweden and the Netherlands. That this linkage was executed *sub rosa* and without formal announcement does not veil its effects. Therefore, if liberalism in trade is a desirable policy, DOT's efforts to open closed markets through linkage were successful—up to a point.

As argued above, however, DOT's liberalizing hand lost its steadiness as the agency approached what it believed was the outer limit of its statutory discretion. For example, the Bush Administration's Transportation Secretaries took contradictory positions on the question of amending the Act.

The Open Skies Order gave an official gloss to the conservative policy on linkage and formally narrowed DOT's discretion to negotiate ownership and control in exchange for access to foreign airspace. By implicitly foreclosing linkage, the Open Skies Order runs against the grain of past success in American bilateral bargaining.

D. *European Proposals and the Implications of the Open Skies Order for Linkage*

With leading transportation officials at times promoting the evolution of U.S. aviation law and diplomacy, why has the

[145] For example, the purchase of more than 25% of the voting stock of an American carrier by a foreign entity would clearly violate the statute. *See supra* note 6.

United States refrained from linking overtly foreign ownership allowances with permissive provisions in bilateral treaties? More pointedly, why did DOT, at least by implication, exclude linkage from the Open Skies Order of August 1992?

U.S. behavior seems even more puzzling when contrasted with the comparatively clear European position on linkage. Recent draft proposals for new aviation regulations in the EC have, at least implicitly, moved in the direction of linkage, and they certainly avoid foreclosing it. For example, the EC has drafted a proposed regulation on operating licenses that addresses the important question of non-European access to the European aviation market and that bears substantial resemblance to a proposed U.S. law on foreign takeovers.[146] This proposed EC regulation sets limits on non-domestic ownership and control:

> OPERATING LICENCE
> Article 4
>
> 1. No undertaking shall be granted an operating licence by a Member State unless air transport is its main occupation and the registered office and principal place of business is located in that Member State.
>
> 2. The undertaking must be owned and continue to be owned directly or through a majority shareholding by Member States and/or nationals of Member States. It must at all times be effectively controlled by such States or nationals. The majority of the board must be representatives of such States or nationals
>
> 3. Notwithstanding paragraph 2 air carriers which have already been recognized in Annex I to Council Regulations (EEC) Nos 2343/90 and 294/91 retain their rights under this and associated Regulations as long as they meet the other obligations in the present Regulation.[147]

Although the EC amplifies and clarifies these restrictive provisions in an addendum to Article Four, it also provides for compromise, if such compromise would secure liberalized aviation treaties from foreign powers:

> Due to basic characteristics of the international aviation system, requirements on Community ownership, control and location are required. An air carrier must be owned and

---

[146] See *infra*, notes 237–246 and accompanying text, for a discussion of the Clinger bill; *see supra* notes 59–65 and accompanying text for an enumeration of the EC's proposed air carrier regulations.

[147] COMMISSION OF THE EUROPEAN COMMUNITIES, *supra* note 29, at 23.

> effectively controlled by a majority of Community nationals and the majority of the board must consist of such nationals.
>
> It is, however, clear that the ownership limitations may hinder a normal business development in the interests of Community air carriers. It is, therefore, desirable to introduce a possibility to conclude more liberal agreements with third countries on a mutually beneficial basis, without prejudice to international commitments.[148]

The EC thus acknowledges that its ownership regulations are flexible, especially where aviation agreements with the EC's trading partners are concerned. The EC proposal distinguishes itself from the U.S. open skies definition in that the U.S. definition seems to foreclose diplomatic linkage while the EC document officially recognizes a role for "mutually beneficial" aviation policy.

An additional indication of the EC's approach is its acknowledgement that a problem exists and that a solution is necessary:

> The Commission always considered the lack of a coherent policy in the area of issuing (or confirming) operating licenses for air carriers as a major outstanding policy question in the area of market access and market entry. It has therefore, in close cooperation with experts from Member States and interested parties . . . thoroughly examined this issue.[149]

The EC demonstrates an awareness that an effective airline ownership policy requires coherence in the rules that govern operating licences, and that these rules largely determine access to investment in the internal market. The EC thus recognizes that a clear policy on licensing naturally facilitates the framing of a clear policy on foreign investment.

In contrast, such observations have not guided DOT. I shall now consider the factors that have shaped this unsatisfactory American policy.

E. *Constraining Influences on Linkage in U.S. Aviation Policy*

In the BA/USAir transaction, DOT received filings in opposition from numerous U.S. airlines, unions, and other industry interests, including Delta, United, and American Airlines.[150]

---

[148] *Id.* at 9.

[149] *Id.* at 8.

[150] These motions are discussed in part V, *infra* notes 156–195 and accompanying text.

These opposition motions by the major airlines called for linkage between investment access in the United States and favorable treatment of American carriers overseas. Despite this support for linkage, DOT shied away from enunciating such a position. DOT's behavior is best understood in the full diplomatic and political context that frames the agency's action and constrains the trend towards linkage.

One constraining factor has been the nature of the bilateral negotiating process. The content of aviation treaty negotiations is not a matter of public record and thus is not susceptible to public scrutiny. Secrecy at the negotiating table makes it possible—and perhaps even necessary—for DOT to refrain from acknowledging the linkages that do and, perhaps, should occur. Although bilateral aviation negotiations occur behind closed doors, one exception to confidentiality has arisen as a matter of practice. The Air Transport Association of America ("ATA") sends one auditor to each session to prepare reports on the talks. This auditor serves as a representative of the airline industry as well as related (albeit not always allied) interest groups (such as the Airline Pilots Association). These reports are confidential; for example, the cover of one such report states in clear terms that its contents are not for public dissemination:

> Objection is hereby made under the provisions of Section 1104 of the Federal Aviation Act of 1958 (49 U.S.C. 1504), as amended, by Public Law 95-504, to the public disclosure of any information contained in the attached documents. The grounds for this claim are:
> A. Such information, if disclosed, will prejudice the formulation and presentation of positions of the United States in international negotiations since the contents of the documents contain comments and views of U.S. officials in the course of negotiations, and,
> B. The release of certain traffic and operating data or plans of U.S. air carriers contained herein will adversely affect the competitive position of those carriers.[151]

The ATA further claims that its notes are of a sensitive political nature.[152] Thus, the protocols of negotiation influence DOT's posture on ownership and aviation treaties.

Of equal or greater import is the traditional and current pos-

---

[151] Notes of Lisa Ray, ATA Representative on U.S. Delegation, Negotiations with Germany in Washington, D.C., 1–2 (Mar. 12–13, 1992) (on file with the *Harvard Journal on Legislation*).

[152] *Id.* at 2.

ture of interest groups in the U.S. aviation industry. Recently, some U.S. airlines have suggested that linkage between foreign ownership and U.S. aviation rights overseas should become U.S. policy. However, no airline has aggressively espoused the principle of linkage; mere suggestions contained within an opposition motion fall short of the type of aggressive lobbying commonly needed to force an agency's hand. A long-term view demonstrates that U.S. aviation interests have consistently and aggressively opposed foreign takeovers of U.S. carriers. Moreover, these interests have evinced even starker opposition to cabotage.

The opposition to cabotage has been described by an officer of one lobbying group as an utterly non-negotiable position. Unions, major airlines and other industry groups speak nearly in chorus on the matter of cabotage—they adamantly oppose allowing foreign airlines to conduct domestic flights.[153] The recent Open Skies Order demonstrates the influence of these powerful lobbies: DOT effectively foreclosed the possibility of cabotage for foreign airlines as well as the possibility of linkage. Thus, the influence of domestic aviation interests constitutes a formidable political barrier to the trend toward trade liberalization and free aviation. This barrier and others have caused DOT policy making to waver uncertainly.

V. The Battle of British Airways: On the Frontier of Liberalization

Between July 1992 and March 1993, BA attempted to purchase a sizable stake in USAir and merge many operations of the two airlines. Over the course of nine months, BA floated two

---

[153] For example, consider this statement by the AFL-CIO:

> These actions could destroy the ability of some U.S. airlines to compete and cause them to be replaced by foreign air carriers on both international and domestic routes. Furthermore, they should raise serious questions of safety and security in the minds of lawmakers who are concerned with protecting the travelling public.
>
> U.S. aviation workers are as competitive as any in the world, but they must be permitted to compete on a level playing field. The AFL-CIO calls on Congress and the Administration to adopt a rational and strategic policy designed to address the global realities of commercial aviation and ensure that this important U.S. industry not succumb to the predatory practices of foreign airlines and their national governments.

AFL-CIO Executive Council, Foreign Control of U.S. Aviation Industry, Bal Harbour, Florida (Feb. 22, 1991) (on file with the *Harvard Journal on Legislation*).

proposals. The first, involving a $750 million investment in USAir, met staunch and effective opposition from USAir's domestic competitors and other interests. When, during the final days of the Bush Administration, it became clear that the transaction would not be approved, BA withdrew the proposal, only to revise its plans almost immediately. The result was a modified proposal, designed to assuage American anxieties over ceding control of an airline to a foreign competitor. The second proposal passed muster with a new DOT Secretary.

The terms of the deal made it clear that BA intended to use the USAir deal as a first step in a continuing effort to loosen American restrictions on foreign ownership. But contrary to the views of some commentators, the Clinton Administration's decision to allow the BA/USAir transaction does not amount to a surrender of American privileges without the prospect of a reciprocal reward. Rather, by allowing the first step of the BA/USAir deal, the United States has in fact improved its bargaining power over the United Kingdom and now more than ever before appears closer to, and more capable of, following a rational policy of linkage. This move by the Clinton Administration might restore liberal direction to this area of trade policy.

A. *The First British Airways Proposal*

In the summer of 1992, BA began the most ambitious effort by a foreign entity to acquire a domestic airline in American aviation history. On July 21, 1992, BA announced its intention to purchase twenty-one percent of the voting shares and forty-four percent of the equity of USAir—a major American air carrier.[154] The proposal would have amounted to a $750 million British stake in the American air carrier. Moreover, the deal would have given BA a substantial voice in the management of their prospective U.S. acquisition.

Opposition to the BA/USAir takeover far exceeded opposition to previous foreign takeover bids. In part, this heightened opposition was related to the timing of the takeover bid, which ar-

---

[154] "British Airways PLC has agreed to invest $750 million in USAir in a move that would form the world's largest airline alliance. The arrangement, which needs to be cleared by U.K. and U.S. regulators, calls for British Air to obtain 21% of the voting stock and a 44% equity stake in USAir Group, Inc." Brett Pulley, *USAir Will Get $750 Million from British Air*, WALL ST. J., July 22, 1992, at A3.

rived amidst a great deal of foreign interest in American air carriers.[155] The already edgy American aviation establishment felt threatened by BA's proposed takeover of USAir—the most ambitious foreign ownership venture to date.

Pursuing a strategy common among private groups that attempt to stop government action, the opponents of the BA/USAir deal registered their views by filing motions with the responsible government agency and also by directing activity at parties outside the agency. In response to the attempted foreign takeover of USAir, opponents waged a full-fledged political and public relations campaign against all foreign takeovers of U.S. airlines.

In particular, organized labor mounted a concentrated attack on the BA/USAir deal. Typical of the efforts launched by unions was the position paper of the Allied Pilots Association ("APA").[156] The position paper argued that BA would effectively assume control of the American target under the proposed terms of the transaction.[157] If the original deal had succeeded, seventy-five percent of the new USAir board of directors would have consisted of Americans; but since a "super majority" of eighty percent would have been required for any major policy to pass, the APA argued that such an arrangement would have given the British minority effective control of USAir. Thus, the transaction would violate the U.S. law barring foreigners from controlling U.S. air carriers.[158] The APA also argued that the transaction failed to meet DOT's implicit reciprocity test: although the United Kingdom's bilateral treaty with the United States is conspicuously illiberal, BA nevertheless would have received a generous ownership allowance under the terms of the proposed

---

[155] See *supra*, notes 110–122 and accompanying text, for details of the SAS/Continental and KLM/Northwest deals. Less than two months after the announcement of the proposed BA deal, Germany's Lufthansa began to investigate taking a sizeable stake in Continental Airlines, and submitted a $400 million bid for Continental, "among the most troubled carriers in the United States," which had operated under Chapter 11 protection since December 1990. Undoubtedly this added to the impression that the American aviation preserve was under siege. Bryant, *supra* note 21. Air Canada, AeroMexico and SAS joined Lufthansa a short time later to consider another deal involving taking a stake in Continental.

[156] Allied Pilots Association, British Airways-USAir Merger . . . Is it Free Trade or a Giveaway of the U.S. Market and Jobs? (1992) (on file with the *Harvard Journal on Legislation*). The APA represents some 11,000 pilots at American Airlines and other carriers.

[157] *Id.* at 3. This argument was a major point of disagreement. The issue of whether the BA/USAir deal results in a beneficial "partnership" or an illegal takeover lies at the center of the debate over foreign ownership.

[158] *Id.*

transaction. The APA report cited the protectionist regulatory policies of the United Kingdom, which restrict U.S. airlines in their U.K. operations far more than U.K. airlines are limited in their U.S. operations.[159] Finally, on both legal and policy grounds, the paper criticized the deal for effectively granting BA cabotage through the "back door" of ownership.[160]

In an even more impressive effort, American, Delta, United, Federal Express, and United Parcel Service retained Powell Tate, a Washington, D.C.-based public affairs firm, to initiate an intensive lobbying campaign against the BA/USAir transaction. Powell Tate produced a memorandum on the proposed transaction, entitled "An Issue of National Policy, Competition, and Fairness: Opposing the British Airways/USAir Merger."[161] The memorandum framed two major "strategic alternatives" for the airlines opposing the transaction:

> The development of an effective campaign against the BA/USAir deal will rest first and foremost on the definition of a firm objective: will it be used to influence the outcome of talks on a new bilateral commercial aviation accord with Britain, or will it work to block the transaction as a matter of national policy? The selection of themes, allies, and approaches—the core components of any public affairs campaign—will depend on which direction is chosen.[162]

The memorandum concluded that a strategy of using the USAir deal for leverage in obtaining a better bilateral treaty with the United Kingdom would represent a "short-term holding action" only, and that a lasting victory would require halting the deal altogether.[163]

The memorandum noted that, although the Bush Administration firmly supported openness to foreign investment "within the limits of the law," the deal's opponents could block it on legal grounds alone.[164] "The difference between formal control and effective control," the memorandum argued, "is a nuance ripe for exploitation."[165] Although the nuts and bolts of the battle would have to be fought over narrow legal questions of the

---

[159] *Id.* at 1.
[160] *Id.* at 3.
[161] Powell Tate, An Issue of National Policy, Competition and Fairness: Opposing the British Airways/USAir Merger (undated) (on file with the *Harvard Journal on Legislation*).
[162] *Id.* at 2.
[163] *Id.* at 7.
[164] *Id.* at 3.
[165] *Id.*

definition of control, Powell Tate believed that the debate should be framed in terms of broad policy concerns, namely trade equity and national economic vitality. By framing the deal as a "*de facto* merger," opponents would emphasize that the proposed transaction amounted "essentially [to] a surrender of U.S. cabotage rights."[166]

The memorandum enumerated interest groups and political institutions that could be approached and cultivated in order to derail the transaction. These included DOT, the Department of Justice, Congress, the Bush Administration, the Clinton presidential campaign, U.K. and EC regulatory authorities, USAir shareholders, USAir and other unions, "national opinion leaders," special interests such as travel agent organizations and consumer groups, and local political allies in cities that serve as air route hubs for carriers likely to suffer from the BA/USAir combination.[167]

In furtherance of the strategy outlined in the Powell Tate memorandum, the companies opposed to the transaction retained Weil, Gotshal & Manges, a prominent New York law firm, to analyze the legal question of whether the transaction would cede control of USAir to the British. Not surprisingly, the attorneys concluded that "if the transactions contemplated by the Investment Agreement were consummated, British Airways would control USAir."[168] They analyzed the BA investment proposal in detail and identified six elements that, taken together, would give BA effective control over USAir:

(1) "Size of Investment in USAir; Voting Interest in USAir." The memorandum emphasized that BA would hold a voting interest of twenty-one percent, more than twice the size of the next largest shareholder. If an entity stands out as the largest shareholder, it is strongly presumed that the entity will gain effective control of the company.[169]

(2) "Assured Representation on USAir Board; Possibility of Increased Representation; Approval of Nominees to the USAir Board."

---

[166] *Id.* at 6.

[167] *Id.* at 4–5, 7–9.

[168] Ira M. Millstein & Robert L. Messineo, Weil, Gotshal & Manges, Memorandum Re: Control of USAir Group Inc. by British Airways Plc 1 (Sept. 21, 1992) (on file with the *Harvard Journal on Legislation*).

[169] *Id.* at 2.

The memorandum argued that BA would come to control over fifty-six percent of the USAir board through complex provisions in the Investment Agreement between the two companies.[170]

(3) "Veto Power over Fundamental USAir Corporate Events."
Major strategic decisions would be subject to a BA veto.[171]

(4) "Effective Control (through Veto Power) of Essential USAir Operating Decisions."
Significant capital or financing decisions could also be blocked by a BA veto.[172]

(5) "Control over USAir's Management."
Even routine management decisions could fall upon BA.[173]

(6) "Control (through the Integration Principles) over Strategic Direction and Operations of USAir."
"Integration principles" (presumably the devices that would lend BA and USAir the appearance of a single airline) would leave USAir and BA under effective, unified BA management.[174]

Political and public relations strategies supplemented the legal arguments opposing the deal; examples included a letter sent to every member of the U.S. Congress[175] and a publication summarizing the opponents' views of the transaction.[176] The highly political tone of these communications suggests that perhaps DOT was more susceptible to political pressure than to legal arguments, although the strictly legal battle at the agency contributed significantly to the demise of the first BA deal.

BA and USAir signed an "Investment Agreement" on July 21, 1992, under which BA would invest $750 million in USAir in

---

[170] *Id.*

[171] *Id.* at 3.

[172] *Id.*

[173] *Id.* at 3–4.

[174] *Id.* at 2–4.

[175] Letter from American Airlines, Inc., to Members of the U.S. Congress (Aug. 28, 1992) (on file with the *Harvard Journal on Legislation*). This letter is particularly interesting because it hedges the policy and the legal strategy against one another:

> Even if British Airways and USAir restructure the transaction to comply with today's law, approval should not be granted unless the British Government completely liberalizes the U.S.-U.K. aviation agreement—which is one of the most restrictive in the world—and takes action to assure that U.S. carriers have the same rights as British Airways to fly beyond the U.K. to cities around the world.

Id. at 1.

[176] American Airlines, Delta Airlines, Federal Express, United Airlines, United Parcel Service & Powell Tate, An Issue of National Policy, Competition and Fairness: The Case Against the British Airways Takeover of USAir (Aug. 21, 1992) (on file with the *Harvard Journal on Legislation*).

exchange for a large equity position as well as voting and other rights. Shortly thereafter, the companies filed with DOT for approval of the proposed transaction.

The first petition in opposition to the BA/USAir deal came from Delta Air Lines two months later.[177] The Delta petition requested that DOT do the following:

> 1) institute and docket a public proceeding to review the British Airways/USAir transaction and to issue a formal decision;
>
> 2) require British Airways and USAir to file in the Docket all documents and information relating to the proposed transaction in accordance with an Evidence Request attached by Delta;
>
> 3) disapprove the transaction on the grounds that the Investment Agreement (between USAir and British Airways), if consummated, would violate the citizenship, cabotage, and route transfer requirements of the Federal Aviation Act . . . ; and
>
> 4) disapprove the transaction as a matter of policy based upon the illiberality of the bilateral agreement and the substantial lack of bilateral reciprocity between the United States and the United Kingdom.[178]

Delta reiterated many of the arguments against consummation of the proposed transaction with BA. Of most immediate concern was Delta's call for a public hearing and for the production of evidence by BA and USAir. Delta argued that the weightiness and precedential nature of the proposed transaction—which the DOT Secretary himself termed a "watershed"—were of sweeping public import, and demanded open proceedings with a public record.[179] In particular, Delta contended that the case would turn on the minutiae of the relationship proposed between the two airlines—fine points of the Investment Agreement that were mostly still unpublicized.[180]

DOT denied Delta's petition requesting a public proceeding to review the proposed transaction;[181] Delta filed a Petition for Re-

---

[177] Petition of Delta Air Lines, Inc., In the Matter of the Acquisition by British Airways Plc of USAir Group, Inc., Docket No. 48,358 (Sept. 17, 1992).

[178] *Id.* at 1–2.

[179] *Id.* at 4–5.

[180] *Id.* at 5–6.

[181] Notice of Dismissal, DOT Order 92-9-35, Docket No. 48,358 (Sept. 18, 1992).

consideration of that order a few days later.[182] DOT quickly responded to the Petition for Reconsideration by issuing an Order on Reconsideration the day after Delta's filing.[183] That very brief Order granted reconsideration of the first Delta petition and simultaneously denied the relief requested in that petition.[184]

Other opponents to the transaction then sent two additional formal petitions and at least one letter to DOT, supporting Delta's original petition and adding some new arguments as well. Carl B. Nelson, Jr., Associate General Counsel for American Airlines, argued in a letter to a DOT official that

> the public interest in this matter simply demands that the Department conduct a public proceeding. The Department has received thousands of letters on this transaction from elected officials, civic organizations, airports, airline employees, and others, and the proposal has been the subject of editorial comment in newspapers and other media throughout the United States. The future course of the air transportation system of the United States will be determined by DOT's decision. In these circumstances, to brush aside the call for a public proceeding because "informal" procedures were used in an obscure cargo case last year[185] is *a clear abuse of discretion.* Moreover, DOT's haste in rejecting Delta's request, thereby precluding answers from interested parties, is an unprecedented departure from the deliberative process that agencies of government are expected to observe.[186]

On the same day, United filed a petition requesting a public hearing on the transaction and opposing the prospective takeover.[187] United distinguished its petition from Delta's, arguing that recent developments required immediate decisions on crucial issues.[188] Central among these developments was a consult-

---

[182] Petition of Delta Airlines, Inc., for Reconsideration of the Notice of Dismissal, DOT Order 92-9-35, Docket No. 48,358 (Sept. 21, 1992).

[183] Order on Reconsideration, DOT Order 92-9-44, Docket No. 48,358 (Sept. 22, 1992).

[184] *Id.*

[185] Here, Mr. Nelson refers to the Petition for an Investigation into the Continuing Fitness of Wrangler Aviation, Inc., DOT Order 91-12-51 (Jan. 10, 1992), the precedent DOT cited as grounds for dismissing Delta's petitions of September 17 and September 21, 1992.

[186] Carl B. Nelson, Jr., Letter to Jeffrey N. Shane, Assistant DOT Secretary for Policy and International Affairs 1 (Sept. 24, 1992) (emphasis supplied) (on file with the *Harvard Journal on Legislation*).

[187] Petition of United Air Lines, Inc., In the Matter of the Acquisition by British Airways, Plc of USAir Group, Inc., Docket No. 48,370 (Sept. 24, 1992).

[188]

> It seems clear that this matter has moved, or is about to move[,] from the information gathering stage to the process of formulating a policy determina-

ation between United Kingdom and United States officials concerning the future of the aviation agreement between the two nations and a public announcement stating that the two governments intended to accelerate the pace of discussions on trade liberalization. United further noted that only cases of military emergency merited such secrecy in an administrative proceeding.[189]

On the same day that the United petition was filed, American Airlines filed a short answer supporting United and reiterating the call for a public procedure.[190]

DOT's subsequent Notice of Dismissal immediately denied United's petition.[191] DOT thereafter received two further submissions in support of the United petition. One of these[192] was accompanied by a letter signed by the directors of eight U.S. airports and municipal aviation bodies.[193] The other, filed by the United Pilots Master Executive Council,[194] attempted to refute an argument often made by those who favor foreign takeovers—that a transaction such as the BA deal brings foreign capital into the United States and thereby fuels the U.S. economy. The Council cited a recent article in the *Aviation Daily* in which BA's chief financial officer had stated that most of the money slated for investment by his company in USAir would be borrowed in the United States.[195]

Although the American press was not unanimous in its opinion of the proposed transaction, it did offer the public optimistic

---

> tion. United submits that if the Government delays much further a decision to institute a proceeding for public participation and comment, it will open itself to unwanted and unnecessary criticism of the integrity of the process.

Id. at 2–3.

[189] Section 1001 of the Act states that "[e]very . . . official act of [DOT] shall be entered of record, and the proceedings thereof shall be open to the public upon request of any interested party, unless . . . secrecy is requisite on grounds of national defense." 49 U.S.C.A. app. § 1481 (Supp. 1993).

[190] Answer of American Airlines, Inc., To Petition of United Air Lines, Inc., For Establishment of Public Procedures, Docket No. 48,370 (Sept. 24, 1992).

[191] Acquisition by British Airways, Plc, of USAir Group, Inc., Notice of Dismissal, DOT Order 92-9-58, Docket No. 48,370 (Sept. 25, 1992).

[192] Answer of Certain U.S. Airports to Petition of United Air Lines, Inc., Docket No. 48,370 (Sept. 25, 1992).

[193] Letter to DOT Secretary Andrew H. Card, Jr. (Sept. 17, 1992) (on file with the *Harvard Journal on Legislation*). The letter was signed by directors of the following entities: Dallas-Fort Worth International Airport; Detroit Metropolitan Wayne County Airport; San Jose International Airport; Metropolitan Nashville Airport; Greater Orlando Aviation Authority; New Orleans International Airport; Port of Portland, Oregon; and the Connecticut Department of Transportation.

[194] Answer of United Pilots Master Executive Council to Petition of United Air Lines, Inc. for Establishment of Public Procedures, Docket No. 48,370 (Sept. 25, 1992).

[195] *Id.* at 4.

reports on the prospective BA/USAir relationship. One article in the *New York Times* painted the deal as one that would improve American air carrier standards, infuse outside cash into an ailing domestic industry, and serve as a lever for opening British markets.[196] The *Times* also gave extended attention to the salubrious effects of the "seamless travel" that would result from the BA/USAir agreement and other similar foreign takeovers. Focusing on the KLM/Northwest relationship, one article celebrated air carrier mergers as heralding an era of increased international air service for small U.S. cities.[197]

An editorial that appeared in the *New York Times* on October 21, 1992 was typical of media support for the transaction.[198] This editorial suggested that the United States could use the ownership issue to extract open skies terms from trade partners. This suggestion fell very much in line with the trend of U.S. aviation diplomacy.

Support from the press and the Administration notwithstanding, by late October it appeared that the BA/USAir takeover was in jeopardy. Even an official visit by Secretary Card to London did not smooth the way for the deal. Talks between Card and his U.K. counterparts ended fruitlessly on October 22, and a continuation of the talks was put off until November 9—very close to the December 24 deadline for the BA/USAir deal.[199] Another discouraging sign appeared a week later, when Lufthansa announced that it would not invest in Continental.[200] Such was the atmosphere when DOT commenced its formal consideration of the BA/USAir takeover proposal.[201]

Despite these unfavorable circumstances, at least one major advance in American aviation liberalization occurred shortly thereafter. On November 16, DOT gave preliminary approval to a special integrated relationship between Northwest and KLM.

---

[196] Peter Passell, *Economic Scene: Play Now, Fly Later*, N.Y. TIMES, Oct. 8, 1992, at D2.

[197] Agis Salpukas, *The Big Foreign Push to Buy Into U.S. Airlines*, N.Y. TIMES, Oct. 11, 1992, at F11.

[198] *Open Skies, Friendlier Skies*, N.Y. TIMES, Oct. 21, 1992, at A22.

[199] *A Snag for British Air-USAir Talks*, N.Y. TIMES, Oct. 24, 1992, at 37.

[200] Agis Salpukas, *Lufthansa Ends Plan for Continental*, N.Y. TIMES, Nov. 3, 1992, at D1. Lufthansa had expressed interest in purchasing a $400 million interest in Continental, in cooperation with investor Marvin Davis. The purchase would have consisted of a $100 million investment in Continental and $300 million worth of Continental's corporate bonds.

[201] *U.S. Begins Formal Look at USAir-British Air Deal*, WALL ST. J., Nov. 6, 1992, at A4.

KLM already had purchased a $400 million stake in Northwest; as approved, the arrangement would allow code-sharing, consolidation of redundant operations such as ticketing and baggage handling, and the implementation of identical aircraft exteriors, uniforms, and stationery. The result of the new integration would be much-touted "seamless service" between domestic American feeder lines and international routes. Not surprisingly, the Administration's approval of the Northwest/KLM deal elicited angry replies from Delta, United, and other American carriers.[202] Reiterating prior statements, the major carriers urged that no further concessions be made to European carriers, unless European governments reciprocated by liberalizing their own aviation rules.[203] Undoubtedly, the pending BA/USAir proposal was the main object of these remarks.

If liberalization of American access to the United Kingdom market was the one U.K. concession that might have saved the BA/USAir deal, then the death knell of that deal sounded on November 18, 1992. On that day, Sir Colin Marshall, the chief executive officer of BA, stated that the United Kingdom would not put access for American air carriers to Heathrow Airport on the bargaining table.[204] Sir Colin's comments effectively excluded the one potential concession that the Bush Administration had cited as a justification for allowing the British to purchase a large stake in USAir. Complaints about limited access to Heathrow had formed a leitmotif of American opposition to the BA/USAir proposal; if the United Kingdom had liberalized its aviation arrangements with American carriers—especially by increasing U.S. access to the most important British airport—the deal would have been acceptable to the Bush Administration. Instead, the major American air carriers issued a new round of official condemnations of the proposal.[205]

In what one U.K. commentator described as "an eleventh-hour call to the Americans,"[206] British Prime Minister John Major

---

[202] Agis Salpukas, *U.S. Backs Merger of Northwest Air and Dutch Carrier*, N.Y. TIMES, Nov. 17, 1992, at A1.

[203] *Id.*

[204] Bridget O'Brian, *British Air Chief Says U.S. Insistence Threatens Proposal*, WALL ST. J., Nov. 19, 1992, at B10. BA receives most favored status at Heathrow Airport, one of the most important European hubs, and air carriers of other flags suffer near exclusion from the choice slots there.

[205] *AMR Urges U.S. to Reject British Airways-USAir Deal*, N.Y. TIMES, Dec. 3, 1992, at D4.

[206] Colin Narbrough, *Major To Press for BA Deal*, THE TIMES (London), Dec. 19, 1992, at Business News 18.

personally urged President Bush to approve the BA/USAir proposal. The Prime Minister issued his appeal while meeting with the President at Camp David on December 19, just five days before the deadline for the proposed takeover. By this time, the opposing sides in the struggle over foreign ownership had each retained influential Republican party insiders as lobbyists and had stepped up their public relations campaigns.[207] In spite of great efforts by the British government, BA, and USAir, American opponents of the takeover were poised to prevail.

On December 22, 1992, officers of BA in London announced their abandonment of the proposed takeover.[208]

B. *Diplomacy, Interest Groups, and the Statute—Behind the Collapse of the First British Airways Deal*

The U.K. government and BA responded to the failure of the first bid to acquire a major stake in USAir by asserting that the United States was seeking "unwarranted and unilateral concessions" from the United Kingdom.[209] It is a fundamental premise of this Article, however, that the U.S. tactics were neither unwarranted nor unilateral. The United States has made—and should continue to make—the acquisition of concessions from foreign governments to American air carriers the cornerstone of American ownership policy. Access to the American aviation market serves as a back door to cabotage, which constitutes a rich reward to foreign carriers and demands significant reciprocation from foreign governments.

Although the U.S. airlines align themselves against foreign interests, they nevertheless reach a conclusion parallel to the one expressed by the British: granting liberal ownership allowances in a transaction such as the BA/USAir deal, without achieving commensurate liberalization by the United Kingdom, would amount to the unilateral surrender of American economic rights. Not suprisingly, major U.S. airlines represent a visible source of

---

[207] Edmund L. Andrews, *U.S. is Said to Oppose USAir Deal*, N.Y. TIMES, Dec. 21, 1992, at D1, D7. American Airlines, United, Delta, and Federal Express retained William E. Brock, United States trade representative and Secretary of Labor under President Reagan. BA and USAir hired Howard H. Baker, Jr., a former U.S. Senator and Reagan's Chief of Staff.

[208] Richard W. Stevenson, *British Air Halts Plan to Purchase Big Stake in USAir*, N.Y. TIMES, Dec. 23, 1992, at D4.

[209] *Id.* at D2.

opposition to foreign takeovers. But the airlines have not completely rejected foreign investment; rather, the theme of reciprocity has pervaded recent opposition motions.[210] American aviation interests propose that a country that stands to benefit from a generous ability to buy into a U.S. airline should be pressured to open its skies to U.S. aviation. This proposal is tantamount to asking for linkage between ownership review and bilateral aviation treaties.

If some industry representatives now openly call for this type of diplomatic linkage—and if such linkage has in fact occurred in past cases[211]—then what has prevented DOT from officially acknowledging linkage as its policy?

One explanation for DOT's hesitation lies in the exigencies of diplomacy. It is crucial to remain flexible in bilateral negotiations; flexibility facilitates tailoring a response to the peculiarities of each situation. By declaring itself locked into a single set of terms, a government loses credibility when it proceeds to alter those terms on a case-by-case basis. By refraining from formalizing a general rule, a government avoids the appearance of inconsistency. Variations among America's existing bilateral treaties suggest that flexibility at the bargaining table has been of significant value.

Another possible factor preventing DOT from openly enunciating a policy of linkage in foreign takeovers is that the agency is shackled by the Federal Aviation Act, the statute that governs its operations. But this argument is more like a typical *pro forma* administrative disclaimer than a real policy-driving force. DOT says that it has no discretion over ownership restrictions and thus no real bargaining power when its officers join foreign diplomats in the negotiating room. This argument has some merit; absent an executive order or clear signal of Congressional intent directing DOT to engage in linkage, DOT could be criticized for overstepping its bounds. But even if DOT lacks the discretion to formalize linkage, higher authorities—such as Congress or even the President—could have formalized what has proven to be fruitful strategy in aviation diplomacy. Their failure to act has been one of the most intractable factors preventing the U.S. government from formalizing a policy of linkage.[212]

---

[210] *See, e.g.*, Allied Pilots Association, *supra* note 156.

[211] *See supra* notes 66–78 (discussing the Swedish and Dutch examples).

[212] "Formalization" could mean at least two things. First, linkage could receive

Given that U.S. trade unions adamantly oppose domestic cabotage rights for foreign airlines, it should come as no surprise that they also oppose foreign ownership of U.S. air carriers, which can operate as a surrogate for cabotage.[213] In fact, trade unions are openly hostile to foreign ownership. The unions are much more reluctant than the airlines to consider reciprocal concessions between the United States and its trading partners. To the unions, the cabotage ban and the statutory foreign ownership allowance constitute absolute limitations. The unions' hard position against liberal ownership allowances could help to explain the U.S. government's reluctance openly to declare a policy of linkage; if DOT appeared to be playing fast and loose with values that some powerfully organized constituencies hold dear, the political heat might become dangerously intense.[214] Moreover, that level of political gaming may well lie beyond the ambit of a federal agency.[215]

---

official gloss via inclusion in the open skies definition. This option was discussed in part IV.C, *supra* notes 140–145 and accompanying text. The second option is to codify linkage via a statutory amendment to the Act. Part VI.B, *infra* notes 237–246 and accompanying text, will address this alternative in detail.

[213] Section 1102(c) of the Act, *see* Appendix A, *infra*, provides specific directions for DOT regarding the accommodation of special interests when dealing with parties. Compare this section, particularly the last phrase, to Secretary Card's statement to the press that DOT does not link private transactions to diplomatic negotiations. Salpukas, *supra* note 132 and accompanying text (quoting Secretary Card).

[214] The airlines, unlike the unions, have some reason to entertain liberalized ownership. With many major U.S. carriers bankrupt or near to it, the prospect of capital infusions from abroad seems attractive. Continental and TWA operate under Chapter 11 protection; one-time airline giants Pan Am and Eastern recently fell. It is not surprising that net losses in recent years have left Delta, USAir, United, and American anxious about their futures. Despite brighter results in recent quarters, heavy debt from past bleak seasons burdens even the most sound American airlines.

The need for cash infusion supports arguments in favor of a proposed bill to broaden aviation rights for foreigners. Representative Clinger, upon introducing a bill to revise the rules on foreign ownership of U.S. airlines, said: "It is clear that our Nation's airline industry is in trouble. Today I am introducing legislation to try to help. This Bill will improve the airline industry's access to capital by liberalizing the rules on foreign investment." 137 CONG. REC. E393 (daily ed. Feb. 4, 1991) (statement of Rep. Clinger).

Rather than recognizing the benefits of foreign ownership, the unions envision a nightmare scenario in which foreign entities fly foreign aircraft with foreign crews on U.S. routes. By blocking any changes in law or policy that could make that scenario more plausible, the unions may believe that they are protecting members from potential job losses.

[215] Union opposition to free trade is a long-standing tradition, and lately it has complicated negotations regarding NAFTA. Interestingly, the Bush Administration aggressively championed NAFTA, even though this meant confronting powerful labor interests. But in the context of bilateral aviation treaties and foreign airline ownership restrictions, the Bush Administration took a different approach.

For a recent discussion of the problem of administrative agencies grown too strong,

C. *The Second British Airways Proposal—Searching for a More Palatable Formula*

Although a federal agency abuses its discretion by operating outside the boundaries of the law, it is equally impermissible for an agency to disallow a transaction that no law bars. BA based its second proposal for a partial takeover of USAir upon this principle.

The first takeover proposal had at least two significant flaws. First, although the British investment in USAir would have been within the percentage limitations on foreign ownership established by the Act,[216] the proposal hit a sensitive nerve in the domestic aviation industry and in labor circles. The fact that the proposed takeover *neared* the limits of the law was enough to motivate important American interests to work hard to block the transaction.

The second significant factor that derailed the first BA deal was that it raised a serious legal question: would BA exercise control over USAir? The statute defines the limit on foreign ownership of voting stock as twenty-five percent. It is well known, however, that a single large shareholder can exercise effective control over a corporation by accumulating less than a quarter of the voting stock.[217] The first deal would have given BA the power to select four of the 16 members of the USAir board of directors. The USAir corporate charter requires eighty percent board approval to implement many important decisions. Eighty percent of 16 is 12.8; hence, four directors can veto any decision requiring a majority of thirteen directors.[218] Arguably, foreign control of four members of this board would amount to foreign control of USAir.[219]

---

see Saikrishna Bangalore Prakash, *Hail to the Chief Administrator: The Framers and the President's Administrative Powers*, 102 YALE L. J. 991 (1993).

[216] Recall that the Act limits foreign equity to 49% and voting stock to 25%. *See* 49 U.S.C.A. app. § 1301(16), § 1508 (West Supp. 1993). The first BA proposal would have resulted in foreign ownership of 44% of USAir's equity and 21% of its stock.

[217] *See generally* 18A AM. JUR. 2D *Corporations* § 167 (1985) ("[The corporate form] lends itself conveniently to a structure where management may be centralized in just a few shareholders."). For an illustration of how one shareholder can exercise control seemingly disproportionate to his holdings as a percentage of all outstanding shares, see Levine v. Smith, 591 A.2d 194 (Del. 1991) (describing Ross Perot's involvement with General Motors).

[218] Such a veto power is a common device. Bylaw provisions often establish "special voting and quorum requirements for directors' action." *See* 18A AM. JUR. 2D *Corporations* § 177 (1985).

[219] Professors John C. Coffee of Columbia University Law School and Joel Seligman

On January 21, 1993, BA unveiled a second plan by which it would acquire a significant stake in USAir. Without seeking government approval, BA purchased a new series of USAir shares for $300 million. The shares were convertible to common stock and would immediately convey 19.9% of the voting interest in USAir to BA. Upon shareholder approval of the agreement, that interest would rise to 21.8%.[220] Thus, the initial stake was more than one percent lower than the stake proposed in the first takeover effort. By making a less ambitious opening move, BA hoped to defuse some of the opposition that had contributed to the demise of the first proposal.

Of greater importance was the revised composition of the board of directors. Under the new proposal, BA would control only three, not four, of the USAir directors, so that BA would not hold the veto power it would have acquired under the first proposal. Thus, BA averted the second obstacle that had hampered the earlier deal—trepidation over foreign control of the board.

Although the second proposal appears to be much more modest than the original deal, this new transaction constitutes only the first stage of a three step plan by BA to increase its control of USAir.[221] A second step would allow BA to purchase $200 million of preferred stock in USAir through an option, exercisable within three years of the first step. Then, within two more years, BA would be able to buy $250 million worth of additional preferred stock. A BA spokesman stated that the end result would be a 32.4% voting interest in British hands.[222] The abandoned July 1992 proposal had not generated any administrative precedent; BA withdrew its offer before DOT issued a final order on the proposal. Thus the agency neither expressly foreclosed nor approved the terms of the July 1992 proposal.

---

of the University of Michigan Law School have emphasized that one need not hold a simple majority of the voting stock of a corporation in order to gain control of that corporation. Jeanne B. Pinder, *Will British Airways Control USAir?*, N.Y. TIMES, Mar. 17, 1993, at D1.

[220] Jeanne B. Pinder, *British Air Buys 20% of USAir*, N.Y. TIMES, Jan. 22, 1993, at D1.

[221] The fact that the new deal is scaled back should not conceal its significance. Under the code-sharing arrangement of the new proposal, USAir can now offer direct service to London, via Philadelphia and New York, from 38 cities in 19 states and Puerto Rico. Application of British Airways PLC for an Exemption; Application of USAir for a Statement of Authorization to Code-Share; Application of USAir for a Statement of Authorization for a Wet Lease, DOT Order 93-3-17, Docket Nos. 48,634, 48,640 (Mar. 15, 1993).

[222] Jeanne B. Pinder, *British Air Buys 20% of USAir*, N.Y. TIMES, Jan. 22, 1993, at D1.

But the stated purpose behind the second, revised proposal—eventually to acquire 32.4% of a domestic air carrier—would clearly run afoul of past rulings of DOT and, moreover, the Federal Aviation Act itself.

Without granting permission to go beyond the initial $300 million stake, DOT announced its approval of the second BA proposal on March 15, 1993.[223] In a press conference, Frederico F. Peña, the Secretary of Transportation for the new Clinton Administration, explained that BA's ability further to increase its stake in USAir would depend upon the institution of a more liberal bilateral aviation treaty with the United Kingdom.[224] Thus, it appeared that the Clinton Administration would continue the Bush Administration's policy of pressing for open skies by offering ownership allowances to the major carriers of important trade partners. Peña expressed his expectation that approval of the first step in the BA takeover would pressure the United Kingdom to continue negotiations on open skies, and might even lead to increased opening of U.K. airports to American carriers.[225] The Clinton Administration brought at least one important change to aviation policy: Secretary Peña clearly stated that the United States will use the promise of ownership in domestic carriers as a lever to pry open protected foreign aviation markets.[226]

BA won approval of this new proposal primarly because, by modifying the terms of its original proposal, it mollified DOT's concerns over ceding control of USAir to a foreign entity. But a report issued by the U.S. General Accounting Office ("GAO") on January 8, 1993, shortly before BA announced its second proposal, may have contributed to the comparative ease with which the proposal was consummated.[227] The GAO report reviewed the rationale behind the limits on foreign ownership of domestic air carriers. It concluded that limits on foreign ownership are rooted in obsolete concerns over national security and trade protection. By officially voicing the consensus that the governing statute has become woefully outdated, the GAO report

---

[223] Application of British Airways PLC, *supra* note 221.

[224] Martin Tolchin, *U.S. Gives Britain a Warning in Approving USAir Deal*, N.Y. TIMES, Mar. 16, 1993, at D1.

[225] *Id.*

[226] *Id.*

[227] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 23.

helped to set a new tone for administrative decisions under the Act—including the decision to approve the BA/USAir proposal.

VI. Liberalizing for Leverage: Free Trade Potential in the Trend Toward Open Skies

The Federal Aviation Act of 1958 still stands; and despite contextual changes favoring a more liberal U.S. foreign takeover policy, BA's avowed goal of acquiring an excessive stake in USAir (i.e., one that clearly exceeds the statutory limit) ensures that some of the nagging questions surrounding application of this statute will not be resolved anytime soon. When the BA/USAir deal is understood as a mere first step, it becomes clear that U.S. foreign takeover policy remains in tension with the evolving landscape of civil aviation.

The trend toward globalized air carriers continues, offering potential gains if the proper policy toward foreign takeovers is implemented. However, absent a clearly enunciated policy to link access to the U.S. aviation market with bilateral aviation treaty terms, the United States will lose an opportunity to open foreign skies to U.S. airlines.

Even if granting cabotage to foreigners remains out of the question, the United States government—specifically, DOT—possesses a valuable diplomatic bargaining chip in its power to award access to ownership. Two chief questions face policy makers deciding how to use this advantage. First, how far may DOT proceed toward liberalizing ownership access within the confines of the statute as it currently exists? Second, would an amended statute better serve national interests? Political considerations will help to determine the importance of, and answers to, these questions. For example, does the intransigence of interest groups, especially trade unions, present an insurmountable obstacle to amending the statute? If legislation such as the Clinger bill were likely to prevail in Congress, then a statutory amendment might be the best route. On the other hand, if policy makers decide that Congressional action is unlikely in the face of strong opposition interests, then administrative law will determine the future of foreign ownership.

Regardless of whether the statute is amended, current law supports at least some liberalization of the statutory restrictions

on foreign ownership via reciprocal treatment of trade partners. Adoption of the state banking law model holds great potential for improving America's position in international air commerce, and might also prevent repetition of past failures in the strategic use of trade policy.

A. *Liberalization and Its Benefits Within the Present Statutory Confines*

The BA/USAir deal will offer an excellent chance for DOT to use its discretion to further the diplomatic and free trade interests of the United States. Indeed, some industry leaders such as Robert Crandall, Chairman of American Airlines, have suggested that BA should receive full DOT permission for its acquisition of USAir if the United Kingdom agrees to a liberalized bilateral aviation treaty with the United States.[228] Similar conditions should apply to any ownership proposal arising from a foreign air carrier's interest in a U.S. carrier.

Germany, the United Kingdom, and other nations have maintained extremely restrictive bilateral treaties with the United States. The bilateral treaty between the United Kingdom and the United States was signed on July 23, 1977 at Bermuda, where a previous U.S./U.K. bilateral treaty had been signed in 1957.[229] This so-called "Bermuda II" agreement was most recently amended on May 25, 1989.[230]

The U.S./U.K. aviation treaty includes the following features:

(1) An absence of traffic rights beyond the United Kingdom;
(2) Capacity-control regulations;
(3) Limits on the number of U.S. carriers allowed to serve the United Kingdom from each U.S. gateway city;

---

[228] Robert Crandall said the following of the BA/USAir deal: "The U.S. government must insist that U.S. carriers be granted equal access to . . . beyond routes to other countries [from the United Kingdom] before it grants such access to the enormous U.S. market. [Otherwise, it] would not be long before the U.S. aviation industry would be controlled by foreign flag carriers, [a]nd that is both anti-consumer and anti-competitive." American Airlines Press Release 2 (Sept. 15, 1992) (on file with the *Harvard Journal on Legislation*).

[229] Agreement Concerning Air Services With Annexes and Exchange of Letters, Jul. 23, 1977, U.S.-U.K., 28 U.S.T. 5367.

[230] Agreement Concerning Reciprocal Recognition of Airline Fitness and Citizenship Determinations, May 25, 1989, U.S.-U.K., available in Westlaw, 1993 WL 428728 (Treaty); Agreement Amending the Agreement of July 23, 1977, As Amended (TIAS

(4) Prohibitions on service to major U.K. cities;
(5) Built-in designations of U.S. air carriers to Heathrow landing slots (the most valuable in Europe) that eliminate U.S. autonomy in designating carriers; and
(6) An absence of code-sharing rights.

The bilateral treaty with Germany was signed on July 7, 1955.[231] The most recent amendment was signed on April 25, 1989 and entered into force on August 6, 1992.[232] Typical of the terms that bilateral aviation treaties contain, the U.S./Germany treaty designates a list of allowable destinations, a list of "beyond points" to which flights continuing out of the signatories' territories are limited, and terms under which an airline may transfer passengers to other airlines.[233] The bilateral agreements with Spain, France, and Japan contain similar restrictions.[234]

The United Kingdom's stingy allowances to U.S. airlines leave much to be bargained for. Pressure from carriers in Britain, Germany, and other countries may lead those governments to agree to more permissive bilateral aviation treaties. If the United States makes it clear that it will grant favorable treatment of ownership proposals only when the carriers' governments reciprocate with trade concessions to U.S carriers, then BA would have incentive to pressure its government.[235] A clear statement of the link between ownership allowances and treaty terms would therefore generate tangible benefits in the form of favorable treaties and regulations.[236] However, if the United States

---

8641), Concerning Air Services, May 25, 1989, U.S.-U.K., available in Westlaw, 1993 WL 428729 (Treaty).

[231] Air transport agreement, July 7, 1955, U.S.-F.R.G., 7 U.S.T. 527.

[232] Apr. 25, 1989, U.S.-F.R.G., 30 U.S.T. 7323.

[233] Germany and the United States recently implemented a new aviation agreement. Agreement Between the United States and Germany Amending and Implementing the Interim Agreement Concerning Air Services of Nov. 6, 1992, Apr. 27, 1993, U.S.-F.R.G., available in Westlaw, 1993 WL 245244 (Treaty).

[234] AMERICAN AIRLINES, INDUSTRY BRIEF 7–12 (Sept. 1992) (on file with the *Harvard Journal on Legislation*).

[235] Part IV.B, *supra* notes 137–139 and accompanying text, proposes how private interests abroad could be enlisted to lobby their governments for trade liberalization in air commerce.

[236] However, the three major U.S. airlines—American, Delta, and United—have warned that the value of the American concession (i.e., access to ownership of a U.S. domestic carrier) may so far exceed any reciprocal right won in return as to render the deal unwise. Access to the huge American passenger aviation market through indirect cabotage is such a bonanza, the U.S. airlines argue, that even open access to landing slots and flight routes to Europe are paltry rewards in comparison.

does not demand reciprocal treatment, the bargaining chip of ownership regulation will be wasted.

B. *Liberalizing through Legislation*

Of course, even if the executive branch decides to pursue diplomatic links between foreign ownership and bilateral treaties, the extant statutory mandate still imposes strict limits on what negotiators can offer. If it were possible to overcome the political influence of the interest groups that adamantly oppose liberalized trade, then a new law might succeed in increasing the scope of American bargaining power.

One Congressman, William F. Clinger, Jr. (R-Pa.), advanced a bill in 1991 to replace the ownership restriction provisions with a more liberal rule.[237] Though the Clinger bill has been largely ignored in Congress, it offers a starting point for considering changes to the existing law.[238]

In a statement to his fellow legislators, Representative Clinger explained the need for fresh capital to keep the ailing U.S. air carrier industry on its feet.[239] Citing the "globalization" of industry, Clinger told the House of Representatives that his bill would help American civil aviation adapt to the changed world market.[240] He emphasized to his colleagues that the bill would set a forty-nine percent limit on ownership (increased from the twenty-five percent limit provided by the statute in its existing form), yet would still give the DOT Secretary the discretion to assure that "airlines do not fall into the hands of our enemies."[241] The bill proposes a reciprocal regime as follows:

---

[237] H.R. 782, 102d Cong., 1st Sess. (1991) (the "Clinger bill"). The purpose of the bill was described as follows: "[T]o amend the Federal Aviation Act of 1958 to authorize the Secretary of Transportation to reduce under certain circumstances the percentage of voting interests of air carriers which are required to be owned or controlled by persons who are citizens of the United States." *Id.* at 1. The text of the bill is reproduced in Appendix B, *infra.*

[238] Though the initiative represented by the Clinger bill has been ignored, Congress has moved, at least tentatively, to address the crisis in the nation's aviation industry. In 1992, Congress created a National Commission to Ensure a Strong Competitive Airline Industry. Airport and Airway Safety, Capacity, Noise Improvement, and Intermodal Transportation Act of 1992, Pub. L. No. 102-581, § 204, 106 Stat. 4872, 4891 (1992), 49 U.S.C.A. app. § 1371 (note) (West Supp. 1993). Foreign investment in U.S. air carriers is among the topics that the Commission studies. Perhaps, then, the Commission will encourage Congress to give serious thought to Representative Clinger's proposal.

[239] 137 CONG. REC. E393 (daily ed. Feb. 4, 1991) (statement of Rep. Clinger).

[240] *Id.*

[241] *Id.*

> [A] person who is not a citizen of the United States may purchase voting interests of . . . an air carrier without regard to whether or not such purchase would result in . . . failing to meet such voting interest requirements of section 101(16) if . . .
>
> (1) the air service agreement between the United States and the foreign country of which such person is a citizen is a procompetitive agreement which, at a minimum, allows air carriers to provide air service from any point in the United States to any significant air service point in the foreign country;
>
> . . . .
>
> (3) the laws and regulations of the foreign country would permit a citizen of the United States to acquire, under similar terms and conditions, the same percentage of voting interests of [an air carrier] in the foreign country . . . .[242]

The Clinger bill attempts to strike a balance between the competing concerns of trade openness and trade equity while simultaneously providing for national security. If passed, the Clinger bill would amend Title IV of the Act and grant the DOT Secretary the case-by-case discretion to ensure that the balance outlined by the bill is attained properly. This discretion would include the power to override the objective provisions of the rule if the Secretary deemed it "consistent with the national security interests of the United States [and] otherwise in the public interest."[243] As an additional limit upon foreign influence in aviation, the bill would require that at least two-thirds of the seats on every U.S. airline's board of directors be occupied by Americans.[244]

The Clinger bill extends the scope of the ownership allowance (as defined in the Northwest/KLM transaction)[245] to encompass not only equity stock, but also voting stock, and therefore control. Where the law presently limits foreign ownership of voting stock to twenty-five percent, the Clinger bill would allow up to forty-nine percent.

More importantly, the bill proposes to codify the diplomatic linkages that have already been demonstrated by DOT action, if not DOT words. The bill provides that the United States shall award liberal access to American aviation only to those countries that reciprocate this privilege. Passing the Clinger bill would

---

[242] Clinger bill, *supra* note 237, at 2–3.
[243] *Id.* at 3.
[244] *Id.* at 2.
[245] *See supra*, notes 116–121 and accompanying text.

thus add statutory certainty in a realm that has been subject only to confused, arbitrary agency behavior. This is not to say that DOT's behavior has been irrational; reciprocation of generous treatment abroad via linkage to ownership access here makes sense and has promoted American interests. The bill wisely follows a lead set by DOT by translating prudent (but ill-expressed) agency policy into a clear Congressional mandate.

Adopting the bill would acknowledge and promote the evolution of aviation law. Because DOT has adhered to the maxim that an agency cannot make law but merely executes the legislative fiat, it has not made clear and open statements of its design to accomodate the changing global economy. A new statute could help cure DOT's hesitation and incoherence.

If ownership allowances are to be parlayed into diplomatic leverage, the agency must retain case-by-case discretion; the Clinger bill would let DOT determine how generously to award takeover allowances. The revised statute would guide DOT's discretion by formally authorizing a linkage between ownership and bilateral aviation treaty terms. Thus, if a foreign entity petitioned DOT to be allowed to purchase a stake in a U.S. airline, the generosity of the petitioner's home country toward U.S. aviation would calibrate DOT's final order on the request. For example, a purchaser from Sweden, a country that has negotiated a permissive and generous treaty with the United States, would receive the broadest foreign ownership allowance under the statute (forty-nine percent under Clinger's proposed law). But a purchaser from the United Kingdom, which has a highly restrictive treaty, would receive discriminatory treatment and might even be denied the opportunity to buy any significant stake.

A proper redrafting of the statute, such as Clinger has advanced, would be the best way to obtain openness and clarity on the subject of linkage. The United States would then gain the necessary leverage to implement a systematic policy of winning greater access to foreign aviation markets for U.S. flag carriers.[246]

---

[246] That the proposal granting foreign access to the U.S. market should be accompanied by changes in bilateral treaty terms does not imply that free trade is always a zero-sum game. If fostering increased competition within the home market—a benefit to consumers—was the principal aim of economic policy, and obtaining new markets overseas was only a secondary goal, then the United States would do best to remove *all* hindrances to its aviation market, regardless of whether foreign countries reciprocated. Greater competition could result, generating even more economic benefit beyond increased access to foreign markets.

C. *State Law Precedent for a Reciprocal Aviation Statute*

State banking law has generated a regime of reciprocal liberalization that has proven highly successful in expanding interstate banking. These experiences lend support to proposals (such as the Clinger bill) for revising the Federal Aviation Act.

It is axiomatic that the U.S. banking industry is heavily regulated.[247] A particularly complex aspect of bank regulation controls bank ownership across state lines. Historically, interstate banking has been severely restricted; for instance, the McFadden Act of 1927 barred banks from establishing branches outside their home states.[248] As a result of the McFadden Act's limitations on branching, bank holding companies became a common device. A bank holding company would own subsidiary banks; since the subsidiaries were full banks rather than branches, they escaped the ban imposed by the McFadden Act.[249] The Bank Holding Company Act of 1956, however, put an end to this use of bank holding companies.[250] This act further limited interstate banking by eliminating the option of *de facto* branching under the cover of holding companies.

At the same time, however, the Douglas Amendment put new power in the hands of the states. Although the Douglas Amendment prohibited bank holding companies from acquiring banks in states beyond the holding company's home state, it included an exemption allowing a holding company to acquire a bank if the statutory laws of the target bank's state specifically authorized the acquisition.[251] Thus, while the Douglas Amendment established a federal default rule that severely restricted interstate banking, it left states with the option of replacing the default rule with a more liberal regime of their choice.

California's response is a model for state innovation in federal banking law. Section 3756 of the California Financial Code governs acquisitions of California banks by out-of-state bank holding companies. Approval of a proposed acquisition depends upon two findings by California banking authorities:

---

[247] "[W]e have a confusing, balkanized . . . structure of regulation, with three or more regulators involved in many transactions . . . . American banking law is undoubtedly quirky, and many oddities can be explained only on historical grounds." JONATHAN R. MACEY & GEOFFREY P. MILLER, BANKING LAW AND REGULATION 2 (1992).

[248] *Id.* at 400–01.

[249] *Id.* at 26.

[250] This act in its present form is codified at 12 U.S.C. § 1841 (1988).

[251] 12 U.S.C. § 1842(d) (1988). *See* MACEY & MILLER, *supra* note 247, at 295–96.

> (1) That there is substantial reciprocity between this state and the state or states in which the operations of the foreign bank holding company and any parent foreign bank holding company of the foreign bank holding company are principally conducted.
>
> (2) That the transaction will not have an adverse effect on the public convenience or advantage in this state.[252]

This statute closely parallels the foreign airline takeover legislation proposed by Representative Clinger: a clear-cut reciprocity clause forms the centerpiece of the rule, but it is tempered and qualified by a "public convenience" clause.[253]

The structure of the reciprocity statutes in California and other states has important implications for agency decision making and burdens of proof. It appears that if a transaction were approved under the California reciprocity clauses, the burden of proof on the California banking superintendent would be light. On the other hand, were the superintendent to withhold approval of a proposed transaction by invoking the qualifying clauses, then a heavier burden of proof would logically inhere; the superintendent would have to explain why the clear-cut and mechanistic reciprocity clause was being overridden. The same would be true of decision making by a DOT Secretary under the statute proposed in the Clinger bill. In both situations, the agencies would probably hesitate before invoking the qualifying clauses without strong justification. At the same time, the qualifying clauses would remain available to override the mechanistic reciprocity clauses in cases where allowing foreign or out-of-state ownership would contravene public policy.

The form of the banking statutes, if applied to the aviation context, would allow for a balanced regime of reciprocity in foreign ownership. A new aviation law would encourage approval of transactions in clear-cut cases while allowing administrative latitude when strong arguments against allowing ownership are advanced. For the statute to succeed, courts reviewing DOT decisions would have to demand that a heavy burden of

---

[252] CAL. FIN. CODE, § 3756(b)(1)–(2) (West 1989 & Supp. 1993).

[253] *Cf.* the Clinger bill, *supra* note 237, at §§ 420(a)(1)–(5). The California statute closely parallels similar banking statutes in a number of states. *See, e.g.*, ARIZ. REV. STAT. ANN. § 6-627 (1989 & Supp. 1993); IND. CODE ANN. § 28-2-16-16 (West 1987 & Supp. 1993); ILL. ANN. STAT. ch. 205, ¶ 10/3.07 (Smith-Hurd 1993); MO. ANN. STAT. § 362.600 (Vernon 1968 & Supp. 1993); NEB. REV. STAT. § 8-902.02 (1991); N.J. REV. STAT. § 17:9A-370 (1984 & Supp. 1993); 7 PA. CONS. STAT. ANN. § 117 (1967 & Supp. 1993); VT. STAT. ANN. tit. 8, § 1054 (1984 & Supp. 1992).

proof be imposed upon the administrator who invoked the qualifying clause. This allocation would guard against casual use of the qualifying clause and thus preserve the usefulness of the reciprocity clause.

D. *Strategic Liberalization and Avoiding Trade Warfare*

This Article proposes that strategic awards of ownership access and other rights should become the principal device in an American campaign to achieve global free trade in aviation. At the very least, opening the American aviation market to all competitors would convey benefits in the form of increased market discipline, even if no reciprocal opportunities were granted abroad. But universal open skies would serve American interests more effectively than would indiscriminate increases in access to the American market without use of the opening to obtain reciprocal rewards from U.S. trade partners. Furthermore, if this U.S. strategic policy succeeded in opening the aviation markets of other countries, then the United States could take some credit for improving civil aviation worldwide.

One serious counterargument to this theory is the risk that strategic liberalization of restrictions on foreign activity in U.S. aviation will degenerate into trade warfare and thereby harm, rather than promote, long-term American interests. But my proposal to achieve leverage through liberalizing U.S. aviation restrictions can be distinguished in two chief respects from prior failed efforts to use trade strategically. First, liberalization for leverage does not require erecting barriers where none previously existed—in fact, no barriers are needed at all. My proposal relies solely upon the removal of barriers which have been in place since the aviation industry was born. Second, because the process of open skies diplomacy is bilateral and the United States has more to offer than other foreign countries, special advantages for the United States are guaranteed; the United States has a more powerful bargaining position than its trade partners.

In the past, when the United States has endeavored to use trade as a lever to open restricted foreign markets, efforts have yielded few benefits and have resulted in increased trade protectionism or even trade warfare. The Hawley-Smoot Tariff Act of 1932 provides an extreme example of how the strategic use of

trade restrictions can fail. The act imposed fifty-three percent (and later fifty-nine percent) duties on imports in order to stimulate Depression-era industries and send a message to protectionist European states. Most American trading partners retaliated, resulting in a steady shrinkage in world trade.[254] The Reciprocal Trade Agreements Act, another trade law, was passed in 1934; and although its restrictions were considerably more elastic than those of the Hawley-Smoot Tariff Act, American trade policy remained protectionist through World War II.[255]

Recognizing the folly of trade warfare, the major industrial states entered into the General Agreement on Tariffs and Trade (GATT) following the war. GATT provided for successive rounds of trade negotiations, continuing through the present time. For the 25 years following World War II, the United States took the lead in trade liberalization under the umbrella of GATT. However, the advent of "voluntary" restraints began a process of reversal. From January 1969 through December 1974, the United States imposed a series of such restraints on imports of steel and other goods. The Tokyo Round of GATT negotiations in 1974 was characterized by a new reluctance to lower tariff and other trade barriers further than had already been done during the previous twenty years of GATT,[256] and the Carter Administration implemented new restrictions.[257]

The Reagan Administration commenced a new trade policy, based on reciprocal treatment on a bilateral and product-by-product basis. Nonetheless, it has been argued that the Reagan reciprocal trade policy resulted in increased tariffs, despite the Administration's professed of free trade philosophy.[258] The failure of reciprocity in the 1970s and the 1980s was that each effort to pry open foreign markets with threats and promises resulted in further accretions to the rules limiting trade. One scholar, Victor A. Canto, commenting on the trade policy of the last twenty years, said that "if a nation failed to meet the reciprocity conditions, special restrictions and/or tariffs conceivably could be

---

[254] VICTOR A. CANTO, THE DETERMINANTS AND CONSEQUENCES OF TRADE RESTRICTIONS IN THE U.S. ECONOMY 2 (1986).

[255] *Id.* at 2–3.

[256] *Id.* at 8–9.

[257] *Id.* at 10. Imports as varied as Russian anhydrous ammonia and Japanese light truck chassis were subject to these restrictions.

[258] *See* Peter Morici, *The Implications for the Future of U.S. Trade Policy*, in MAKING TRADE WORK: THE CANADA-U.S. AGREEMENT 134–35 (Peter Morici ed., 1987); Canto, *supra* note 254, at 14–15.

imposed on its exports to the United States . . . . In short, the Trade Act of 1974 [and ensuing policy] permitted any industry, when threatened by imports, to seek and obtain protection."[259]

The reciprocal trade policy in other industries during the last two decades has differed greatly from the aviation commerce policy proposed in this Article. While unrestricted trade in other industries may once have been the norm, federal law has sharply limited significant foreign investment and participation in American aviation since the very inception of civil aviation. Thus, sharp trade restrictions have always been a prominent feature of the civil aviation landscape. Reducing the barriers against foreign ownership of American air carriers would constitute a sizeable and unprecedented reward to foreign air carriers and investors. The policy of reciprocal trade in aviation would not involve the imposition of new barriers as punishment for restrictive behavior by foreign states. Rather, the only U.S. actions necessary to execute the policy proposed in this Article would be to lower barriers or delay their reduction.

With steel and other industries, openness as a starting point led to retaliatory trade policy when restrictions were imposed; the starting point in civil aviation, by contrast, is already close to complete restriction. Thus, reciprocal trade policy in aviation will consist of promises, not threats,[260] and does not pose the same danger of degenerating into trade warfare.

A second distinguishing feature of reciprocity in civil aviation is the degree of American bargaining power. In the bilateral context that frames aviation diplomacy, the United States enjoys the advantage of having a far larger market than that of any single trade partner.[261] Thus, any trade partner would realize huge benefits by awarding American air carriers access to its markets; according to the theory of linkage, the generous nation would then gain access to the vast continental American market for its air carriers. As an added incentive, the foreign carriers that obtain this benefit would enjoy a competitive advantage over the carriers based in restrictive foreign countries. Thus a "dom-

---

[259] Canto, *supra* note 254, at 24.

[260] For a thoughtful commentary on the flaws in retaliatory trade policy, see Leonard Weiss, *Reciprocity*, in MANAGING TRADE RELATIONS IN THE 1980'S: ISSUES INVOLVED IN THE GATT MINISTERIAL MEETING OF 1982 165–231 (Seymour J. Rubin & Thomas R. Graham eds., 1984). "The reciprocity bills place an exaggerated emphasis on the utility of retaliatory restrictions as a means to obtain greater access in foreign markets." *Id.* at 166.

[261] See *supra* note 29 and accompanying chart for comparative data.

ino effect" could occur: the advantages that liberal countries (like the Netherlands and Sweden) gain would impose burdens upon the carriers of more restrictive countries (such as the United Kingdom and Germany). These restrictive countries would then be encouraged to drop their barriers to American aviation. This process constitutes the second distinguishing advantage of reciprocity in civil aviation.

VII. Conclusion

Air commerce has become a multinational industry, and airlines throughout the world have sought to forge ties to their counterparts in foreign countries. The recent interest among European airlines in entering the American market exemplifies this trend; furthermore, with the EC moving to open the skies among its members, American carriers will begin to view the internal European market as a potential area for expansion.

However, the globalization of air commerce stands in marked contrast to American aviation law. Restrictions that have been in place since the inception of civil aviation in the 1920s continue to prevent U.S. airlines from participating fully in the development of international air commerce.

The DOT seems to have implicitly recognized that the liberalization of ownership restrictions is consistent with the present state of the aviation industry. The administrative case law, as well as certain pronouncements by leading DOT officials, indicate a willingness to effect this liberalization, but the law imposes limits. While the Federal Aviation Act leaves DOT with a substantial degree of discretion in executing foreign ownership policy, it still reflects the earliest aviation laws by imposing clear limits on foreign ownership of American carriers—limits which DOT must respect.

Within its mandate, the DOT has reacted sensibly to the trend toward globalization of air commerce; most notably, it has allowed several watershed transactions in which European air carriers have conducted partial takeovers of their American counterparts via ownership arrangements. But DOT has exercised its discretion as far as the statutory mandate allows. Interest groups opposed to liberalizing air commerce have become vigilant in their efforts to restrain DOT initiatives. Such groups have mounted a full-fledged political and public relations campaign

against the latest transaction, rather than being satisfied with mere public comments as they have been in the past. This heightened activism has coincided with the exhaustion of liberalization possibilities under the present statutory regime. Both the black letter law and the pressures of the polity have contributed to slowed progress toward free trade in air commerce.

Assuming that globalization of air carriers is not a transient phenomenon, it is imperative that the United States continue the liberalization of its foreign ownership regime. This Article has explored the possibilities for doing so, both under the current Act and under a proposed amendment. Opposition to granting foreign carriers increased access to the U.S. market (whether by cabotage or ownership) suggests that DOT will not succeed in advancing trade liberalism much further under the current statute. Therefore, an amendment to the Act is necessary.

The DOT must now decide whether to continue to promote liberalization. If it chooses this path, the agency, as an important voice in the debate over the foreign ownership regime, must take an active role in encouraging legislative change. Although an administrative agency is bound to act in accordance with the will of Congress, this necessity does not diminish the agency's role—perhaps unique to the American system of administrative law—in forcing the evolution of legal rules, especially when case-by-case adjudication and rule making will no longer suffice. To participate in the political debate would be consistent with DOT's recent role in adjusting the U.S. foreign ownership regime to the pace of globalized air commerce. Even if the forum for debate moves from DOT to the U.S. Congress, DOT should remain an integral part of the free trade process.

APPENDIX A

TITLES IV, V, AND XI OF THE FEDERAL AVIATION ACT—
"Air Carrier Regulation," "Nationality and
Ownership of Aircraft," and "Miscellaneous"
(in pertinent part)

TITLE IV—AIR CARRIER REGULATION

CERTIFICATE OF PUBLIC CONVENIENCE AND NECESSITY

CERTIFICATE REQUIRED
Sec. 401. (a) No air carrier shall engage in any air transportation unless there is in force a certificate issued by the Board authorizing such air carrier to engage in such transportation.

APPLICATION FOR CERTIFICATE
(b) Application for a certificate shall be made in writing to the Board and shall be so verified, shall be in such form and contain such information, and shall be accompanied by such proof of service upon such interested persons, as the Board shall by regulation require.

NOTICE OF APPLICATION
(c) Upon the filing of any such application, the Board shall give due notice thereof to the public by posting a notice of such application in the office of the secretary of the Board and to such other persons as the Board may by regulation determine. Any interested person may file with the Board a protest or memorandum of opposition to or in support of the issuance of a certificate. Such application shall be set for public hearing, and the Board shall dispose of such application as speedily as possible.

ISSUANCE OF CERTIFICATE
(d) (1) The Board shall issue a certificate authorizing the whole or any part of the transportation covered by the application, if it finds that the applicant is fit, willing, and able to perform such transportation properly, and to conform to the provisions of this Act and the rules, regulations, and requirements of the Board hereunder, and that such transportation is required by the public

convenience and necessity; otherwise such application shall be denied.

. . . .

TERMS AND CONDITIONS OF CERTIFICATE
(e) Each certificate issued under this section shall specify the terminal points and intermediate points, if any, between which the air carrier is authorized to engage in air transportation and the service to be rendered; and there shall be attached to the exercise of the privileges granted by the certificate, or amendment thereto, such reasonable terms, conditions, and limitations as the public interest may require. A certificate issued under this section to engage in foreign air transportation shall, insofar as the operation is to take place without the United States, designate the terminal and intermediate points only insofar as the Board shall deem practicable, and otherwise shall designate only the general route or routes to be followed. Any air carrier holding a certificate for foreign air transportation shall be authorized to handle and transport mail . . . .

PERMITS TO FOREIGN AIR CARRIERS

PERMIT REQUIRED
Sec. 402. (a) No foreign air carrier shall engage in foreign air transportation unless there is in force a permit issued by the Board authorizing such carrier so to engage.

ISSUANCE OF PERMIT
(b) The Board is empowered to issue such a permit if it finds that such carrier is fit, willing, and able properly to perform such air transportation and to conform to the provisions of this Act and the rules, regulations, and requirements of the Board hereunder, and that such transportation will be in the public interest.

APPLICATION FOR PERMIT
(c) Application for a permit shall be made in writing to the Board, shall be so verified, shall be in such form and contain such information, and shall be accompanied by such proof of service upon such interested persons, as the Board shall by regulation require.

NOTICE OF APPLICATION

(d) Upon the filing of an application for a permit the Board shall give due notice thereof to the public by posting a notice of such application in the office of the secretary of the Board and to such other persons as the Board may by regulation determine. Any interested person may file with the Board a protest or memorandum of opposition to or in support of the issuance of a permit. Such application shall be set for public hearing and the Board shall dispose of such application as speedily as possible.

. . . .

CONSOLIDATION, MERGER, AND ACQUISITION OF CONTROL

ACTS PROHIBITED

Sec. 408. (a) It shall be unlawful unless approved by order of the Board as provided in this section—

. . . .

(4) For any foreign air carrier or person controlling a foreign air carrier to acquire control, in any manner whatsoever, of any citizen of the United States engaged in any phase of aeronautics;

. . . .

POWER OF BOARD

(b) Any person seeking approval of a consolidation, merger, purchase, lease, operating contract, or acquisition of control, specified in subsection (a) of this section, shall present an application to the Board, and thereupon the Board shall notify the persons involved in the consolidation, merger, purchase, lease, operating contract, or acquisition of control, and other persons known to have a substantial interest in the proceeding, of the time and place of a public hearing. Unless, after such hearing, the Board finds that the consolidation, merger, purchase, lease, operating contract, or acquisition of control will not be consistent with the public interest or that the conditions of this section will not be fulfilled, it shall by order approve such consolidation, merger, purchase, lease, operating contract, or acquisition of control, upon such terms and conditions as it shall find to be just and reasonable and with such modifications as it may prescribe: *Provided*, That the Board shall not approve any consolidation,

merger, purchase, lease, operating contract, or acquisition of control which would result in creating a monopoly or monopolies and thereby restrain competition or jeopardize another air carrier not a party to the consolidation, merger, purchase, lease, operating contract, or acquisition of control: *Provided further*, That if the applicant is a carrier other than an air carrier, or a person controlled by a carrier other than an air carrier or affiliated therewith within the meaning of section 5(8) of the Interstate Commerce Act, as amended, such applicant shall for the purposes of this section be considered an air carrier and the Board shall not enter such an order of approval unless it finds that the transaction proposed will promote the public interest by enabling such carrier other than an air carrier to use aircraft to public advantage in its operation and will not restrain competition.

. . . .

FORM OF CONTROL
Sec. 413. For the purposes of this title, whenever reference is made to control, it is immaterial whether such control is direct or indirect.

TITLE V—NATIONALITY AND OWNERSHIP OF AIRCRAFT

REGISTRATION OF AIRCRAFT NATIONALITY

REGISTRATION REQUIRED
Sec. 501. (a) It shall be unlawful for any person to operate or navigate any aircraft eligible for registration if such aircraft is not registered by its owner as provided in this section, or (except as provided in section 1108 of this Act) to operate or navigate within the United States any aircraft not eligible for registration . . . .

ELIGIBILITY FOR REGISTRATION
(b) An aircraft shall be eligible for registration if, but only if—
(1) It is owned by a citizen of the United States and it is not registered under the laws of any foreign country; or

(2) It is an aircraft of the Federal Government, or of a State, Territory, or possession of the United States, or the District of Columbia, or of a political subdivision thereof.

. . . .

TITLE XI—MISCELLANEOUS

INTERNATIONAL AGREEMENTS

GOALS FOR INTERNATIONAL AVIATION POLICY
Sec. 1102.

. . . .

(b) In formulating United States international air transportation policy, the Congress intends that the Secretary of State, the Secretary of Transportation, and the Civil Aeronautics Board shall develop a negotiating policy which emphasizes the greatest degree of competition that is compatible with a well-functioning international air transportation system. This includes, among other things,

(1) the strengthening of the competitive position of United States air carriers to at least assure equality with foreign air carriers, including the attainment of opportunities for United States air carriers to maintain and increase their profitability, in foreign air transportation;

(2) freedom of air carriers and foreign air carriers to offer fares and rates which correspond with consumer demand;

(3) the fewest possible restrictions on charter air transportation;

(4) the maximum degree of multiple and permissive international authority for United States air carriers so that they will be able to respond quickly to shifts in market demand;

(5) the elimination of operational and marketing restrictions to the greatest extent possible;

(6) the integration of domestic and international air transportation;

(7) an increase in the number of nonstop United States gateway cities;

(8) opportunities for carriers of foreign countries to increase their access to United States points if exchanged for benefits of similar magnitude for United States carriers or the traveling public with permanent linkage between rights granted and rights given away;

(9) the elimination of discrimination and unfair competitive practices faced by United States airlines in foreign air transportation, including excessive landing and user fees, unreasonable ground handling requirements, undue restrictions on operations, prohibitions against change of gauge, and similar restrictive practices; and
(10) the promotion, encouragement, and development of civil aeronautics and a viable, privately owned United States air transport industry.

CONSULTATION WITH AFFECTED GROUPS
(c) To assist in developing and implementing such an international aviation negotiating policy, the Secretaries of State and Transportation and the Civil Aeronautics Board shall consult, to the maximum extent practicable, with the Secretary of Commerce, the Secretary of Defense, airport operators, scheduled air carriers, charter air carriers, airline labor, consumer interest groups, travel agents and tour organizers, and other groups, institutions, and government agencies affected by international aviation policy concerning both broad policy goals and individual negotiations.

APPENDIX B

H.R. 782—THE CLINGER AMENDMENT TO
The Federal Aviation Act of 1958

SECTION 1. REDUCTION OF U.S. CITIZENSHIP VOTING INTEREST OWNERSHIP REQUIREMENT

Title IV of the Federal Aviation Act of 1958 (49 U.S.C. App. 1371–1389) is amended by adding at the end the following new section:

"SEC. 420. REDUCTION OF U.S. CITIZENSHIP VOTING INTEREST OWNERSHIP REQUIREMENT.

"(a) GENERAL RULE.—Notwithstanding the requirement of section 101(16) of this Act that at least 75 percent of the voting interests of an air carrier be owned or controlled by persons who are citizens of the United States or of one of its possessions, a person who is not a citizen of the United States may purchase voting interests of a corporation or association which is, or owns or controls, an air carrier without regard to whether or not such purchase would result in the corporation or association failing to meet such voting interest requirement of section 101(16) if the Secretary of Transportation finds that—

"(1) the air service agreement between the United States and the foreign country of which such person is a citizen is a pro-competitive agreement which, at a minimum, allows air carriers to provide air service from any point in the United States to any significant air service point in the foreign country;

"(2) after the purchase, the president, chairman of the board of directors, chief operating officer, and two-thirds or more of the board of directors of the corporation or association which is, or owns or controls, the air carrier would be citizens of the United States;

"(3) the laws and regulations of the foreign country would permit a citizen of the United States to acquire, under similar terms and conditions, the same percentage of voting interests of

a person who provides in the foreign country transportation by aircraft of persons or property as a common carrier for compensation or hire as the percentage of ownership which the person making the purchase would have in the air carrier after the purchase;

"(4) the purchase is consistent with the national security interests of the United States; and

"(5) the purchase is otherwise in the public interest.

"(b) MAXIMUM REDUCED PERCENTAGE.—The 75 percent voting interest requirement of section 101(16) of this Act may not be reduced to less than 51 percent under this section."

. . . .

ARTICLE

MAIL FRAUD AND THE INTANGIBLE RIGHTS DOCTRINE: SOMEONE TO WATCH OVER US

GERALDINE SZOTT MOOHR*

Federal mail fraud prosecutions of state and local officials have generated much controversy. In their efforts to eradicate local political corruption, federal prosecutors have extensively used, and possibly abused, the doctrine of an "intangible right" to the honest services of public officials. After examining the development of the intangible rights doctrine, Ms. Moohr argues that the doctrine runs afoul of federalism, separation of powers, and the First Amendment. She concludes that the current mail fraud statue's incorporation of the intangible rights doctrine makes the statute void for vagueness. Finally, Ms. Moohr proposes new mail fraud legislation that does not infringe on important constitutional principles.

> You must nevertheless bear in mind what I am about to say to you: in the seed of the city of the just, a malignant seed is hidden, in its turn: the certainty and pride of being in the right—and of being more just than many others who call themselves more just than the just.[1]

The reality of American politics begins with political favors, progresses through the rewards of party patronage and venial kickback schemes, and too often ends with stuffed ballot boxes. Although political corruption at the state and local levels has long existed in American politics, everyone agrees in principle that it is a "bad thing." Most of us would prefer to live in a platonically inspired Greek city-state or a utopian village of the colonial republicans, where all citizens participate and the best become public servants and govern according to altruistic ideals of community service. The achievement of this ideal, however, requires that the governed demand good government.

Political corruption is profoundly undemocratic; it interferes with the democratic process by giving an unfair advantage to

---

* Associate, Covington & Burling. B.S., University of Illinois; M.S., Bucknell University; J.D., American University, 1991. The author is especially grateful to Ira Robbins for his substantive suggestions and encouragement, and also thanks Jamin Raskin, whose seminar inspired the Article, and Richard Lavine, Catherine Mulrow, and Roger Sherman for their insightful comments and careful readings of earlier drafts of the manuscript.

[1] ITALO CALVINO, INVISIBLE CITIES 162 (William Weaver trans., 1974).

incumbents and insiders who control party patronage and the government bureaucracy.[2] For the past twenty years, the federal government has tried to secure good government by prosecuting state and local public officials for public corruption. In the two decades between 1970 and 1990, 943 state officials and 3226 local officials were convicted in federal courts of various charges of political corruption.[3] As of December 31, 1990, 1561 state and local officials awaited trial.[4] Comparative figures reflect the increased federal attention to political corruption. In 1970, federal prosecutors indicted 45 federal, state, and local officials; in 1980, the figure was 442; and in 1990, 968.[5] Federal prosecutors often base these prosecutions on charges of mail fraud.[6]

---

[2] Even if one accepts the notion that a certain level of political patronage facilitates party loyalty and promotes an effective two-party system, clearly it should be monitored and controlled. *See* Rutan v. Republican Party of Ill., 497 U.S. 62, 92–115 (Scalia, J., dissenting) (discussing role of patronage in two-party political system), *reh'g denied*, 497 U.S. 1050 (1990).

[3] PUBLIC INTEGRITY SECTION, U.S. DEP'T OF JUSTICE, REPORT TO CONGRESS ON THE ACTIVITIES AND OPERATIONS OF THE PUBLIC INTEGRITY SECTION FOR 1990, at 24–25 (Sept. 1991).

[4] *Id.*

[5] *Id.* The number of indictments of state and local officials has remained fairly constant since 1980. The increase between 1980 and 1990 is due in large part to the increased numbers of federal officials who have been indicted and convicted. *Id.*

[6] *See Mail Fraud: Hearing Before the Subcomm. on Criminal Justice of the House Comm. on the Judiciary*, 100th Cong., 2d Sess. 7 (1988) [hereinafter *Mail Fraud Hearings*] (statement of Rep. Conyers (D-Mich.)) (almost all of 1000 political corruption convictions in 1987 were based at least in part on violations of mail and wire fraud); Michael W. Carey et al., *Federal Prosecution of State and Local Public Officials: The Obstacles to Punishing Breaches of the Public Trust and a Proposal for Reform, Part One*, 94 W. VA. L. REV. 301, 356–67 (1991–92) (listing 78 federal political corruption convictions in West Virginia from 1984 to 1991, of which 13 involved mail fraud); Gregory H. Williams, *Good Government by Prosecutorial Decree: The Use and Abuse of Mail Fraud*, 32 ARIZ. L. REV. 137, 144–45 (1990) (noting that 40% of corruption cases include mail fraud charges).

Prosecutors also rely on the wire fraud statute, 18 U.S.C. § 1343 (1988 & Supp. IV 1992); the Hobbs Act, 18 U.S.C. § 1951 (1988); and the Travel Act, 18 U.S.C. § 1952 (1988 & Supp. IV 1992). These statutes, together with mail fraud, are predicate crimes for criminal offenses under the Racketeer Influenced and Corrupt Organizations Act, 18 U.S.C. §§ 1961–1968 (1988 & Supp. IV 1992) ("RICO"), and therefore allow prosecutors to add RICO charges. A series of mailings can form the required "pattern of racketeering activity" that gives rise to a RICO offense, 18 U.S.C. § 1962 (1988). *See* United States v. Mandel, 591 F.2d 1347, 1352 (4th Cir.) (Maryland governor charged with 15 counts of mail fraud and one RICO offense), *conviction aff'd per curiam in relevant part on reh'g en banc*, 602 F.2d 653 (4th Cir. 1979), *cert. denied*, 445 U.S. 961 (1980). This practice, combined with the broad definition of "enterprise," significantly expands the reach of the RICO statute. *See* Sedima, S.P.R.L. v. Imrex Co., 473 U.S. 479, 502 (1985) (Marshall, J., dissenting) (arguing that, given the "extraordinary expansion" of mail fraud, the implications of making two instances of mail fraud actionable under RICO are "staggering"); Steven B. Duke, *Legality in the Second Circuit*, 49 BROOK. L. REV. 911, 924 & n.81 (1983) (suggesting that an attorney who mailed two dishonest briefs within 10 years might be subject to RICO charges).

A paradigmatic example of a mail fraud prosecution for political corruption involves an elected official who directs state insurance contracts to party supporters and political friends.[7] In this typical case, the state legislature budgets money for the insurance contracts; the insurance is purchased; the premiums charged do not exceed those of other insurers; and the designated companies give value, insurance coverage, for the money. The politician may or may not receive a financial benefit through an ownership interest in one of the designated insurance companies, but in either case the state does not sustain pecuniary harm. Although the favoritism, naked self-interest, and misuse of power are unacceptable, such behavior has not, until recently, been defined as criminal by the federal government. Indeed, some regard these practices as an inevitable application of the adage, "to the victor belong the spoils."

Federal prosecutors bring mail fraud charges in such cases because the statute confers an intangible right to a politician's "honest service" upon the citizenry—even when the state has not lost money or property. The right to honest service is substituted for the right in money or property normally protected by criminal laws that prohibit fraud. When combined with the principles of fiduciary duty, the theory holds a public official liable for depriving the electorate of honest services by failing to disclose material information. Thus, the "intangible rights" doctrine allows prosecutors to apply the statute to a wide range of conduct—some of which is not otherwise illegal—of both elected officials and nonelected political participants.

Whether or not federal prosecutions based on the intangible rights doctrine are justified, they are problematic. Prosecutions of political corruption may erode confidence in the federal prosecutor and, ultimately, in the fairness of criminal enforcement.[8] Mail fraud prosecutions for political corruption that em-

---

[7] The model is taken from the facts of two landmark mail fraud cases. *See* McNally v. United States, 483 U.S. 350 (1987); United States v. Margiotta, 688 F.2d 108 (2d Cir. 1982), *cert. denied*, 461 U.S. 913 (1983).

[8] Examples include the objections of officials who claim that they are targeted because of their race or political affiliation. *See, e.g.*, Ronald Smothers, *Charge Against Mayor Strikes Chord in Birmingham*, N.Y. TIMES, Jan. 20, 1992, at A12 (Mayor Richard Arrington of Birmingham, Alabama charges federal prosecutors with racism after experiencing 15 investigations within three mayoral terms); Bridget O'Brian, *A Shrewd Deal or Gubernatorial Bribe?*, NAT'L L.J., Dec. 23, 1985, at 13 (Governor Edwin Edwards of Louisiana claims his federal prosecution is politically motivated). *See also* Susan Sward & Bill Wallace, *Ethics in Government*, S.F. CHRON., Feb. 4,

ploy the intangible rights doctrine conflict with a fundamental principle of federalism by diminishing state and local autonomy and threatening other constitutional provisions. Furthermore, it is not at all clear that the control of political corruption should be achieved through vague, undefined federal legislation that gives enormous discretion to police and prosecutors.

The perplexing problem is whether political corruption can be prevented and punished without denigrating constitutional protections and individual civil rights. A Hobson's choice is seen to exist between enacting a specific statute that may be circumvented or a vague statute that is subject to selective enforcement.[9] Since the late 1970s, law enforcement officials, the judiciary, and, in their turn, legislators have chosen the latter option. This Article explores the consequences of that choice by examining the effect of the present mail fraud statute on three constitutional values: federalism; the separation of powers; and First Amendment protection for political activity. I conclude that the benefits of federal prosecutions do not outweigh the long-term interest in protecting those constitutional provisions.[10]

In the first part of this Article, I examine the judicial development of the intangible rights doctrine, its eventual demise, and its subsequent legislative reincarnation. At the urging of federal prosecutors, the judiciary developed the doctrine in a series of cases that culminated in *McNally v. United States*.[11] In that case, the Supreme Court ended the long-standing practice of using

---

1991, at A1 (relating series of federal mail fraud prosecutions of local officeholders and their claims of political harassment); Philip Shenon, *U.S. Officials See Sweeping Effort to Combat Municipal Corruption*, N.Y. TIMES, Mar. 30, 1986, at A1 (same). The credibility of such charges supports the argument that the integrity of the federal prosecutor's office may be damaged by federal intervention in state and local politics. *See* Jed S. Rakoff, *The Federal Mail Fraud Statute (Part I)*, 18 DUQ. L. REV. 771, 779 (1980) (commenting that if mail fraud statute is abused or used irrationally, it will cease to command moral force).

[9] *Mail Fraud Hearings*, *supra* note 6, at 80 (statement of R. Coombs). *See also* Rakoff, *supra* note 8, at 779 (stating that problem is to retain effectiveness of mail fraud statute while eliminating potential for abuse); Peter M. Oxman, *The Federal Mail Fraud Statute After* McNally v. United States, *107 S. Ct. 2875 (1987): The Remains of the Intangible Rights Doctrine and Its Proposed Congressional Restoration*, 25 AM. CRIM. L. REV. 743, 786–89 (1988).

[10] The prosecution of politicians for classic fraudulent schemes is not considered here. *See, e.g.*, United States v. LaRouche, 896 F.2d 815 (4th Cir. 1990) (upholding mail fraud conviction where candidate solicited campaign contributions under false pretense that he would repay the "loan" and that money would be used for particular purpose).

[11] 483 U.S. 350 (1987).

mail fraud in cases of political corruption when it ruled that the mail fraud statute was not intended to, and did not, ensure honest services. Congress promptly responded by amending the mail fraud statute, adding the intangible right to honest services to the property rights traditionally protected by criminal law.

In the second part, I analyze three structural constitutional protections that are affected by the intangible rights doctrine and conclude that its use has negative effects on all three provisions. First, the application of the doctrine to political corruption diminishes state and local autonomy and disturbs the Constitution's structural mandate of a federalist system. The relationship between the federal government and the states is one of inherent tension, strained even further by the extensive power of the modern administrative bureaucracy. The balance between national and local government is tilted toward the former when federal authorities acting under ambiguous legislative standards monitor local political conduct. Second, the intangible rights doctrine encourages judicial crime creation, which is prohibited by the constitutional structure requiring the separation of judicial and legislative power. Third, federal prosecutors may use the intangible rights doctrine to infringe First Amendment speech and association rights, effectively discouraging the political participation essential to a democratic two-party system. Furthermore, the impact of the intangible rights doctrine on these three constitutional provisions is exacerbated by the discretion that the doctrine accords to prosecutors.

Next, I evaluate the rationale and application of the vagueness doctrine and conclude that the amended mail fraud statute is facially void for vagueness. The statute is vague because its ambiguous language fails to provide enforcement standards to prosecutors and adequate notice of criminality to the public. It is facially vague because its use undermines the federalist structure, blurs the separation of powers, and chills protected First Amendment political activity.

In the final part of this Article, I propose legislation designed to provide more protection for constitutional provisions while controlling political corruption. The proposed legislation limits prosecutions to situations in which a pecuniary loss occurs, reduces the statute's emphasis on intent, and redefines federal government jurisdiction.

I. MAIL FRAUD AND THE INTANGIBLE RIGHTS DOCTRINE

The original mail fraud statute was enacted in 1872,[12] primarily to address the sale of counterfeit currency through the mail. The statute has evolved to allow the federal government to prosecute defendants for conduct over which it otherwise would have no jurisdiction and only a peripheral interest. The following discussion examines the structure and language of the statute, briefly reviews the pivotal judicial constructions that broadened and then narrowed the reach of the statute, and documents the legislative response.[13]

A. *The Mail Fraud Statute*

To violate the mail fraud statute, a person must (1) devise or intend to devise a scheme to defraud, obtain money or property by false pretenses, or sell or otherwise deal in counterfeit currency; and (2) mail, receive via mail, or cause to be delivered by mail a document for the purpose of executing or attempting such a scheme.[14] Thus, the statute has only two elements: an

---

[12] Act of June 8, 1872, ch. 335, § 301, 17 Stat. 283, 323. The original mail fraud statute was part of comprehensive legislation prohibiting use of the mail to conduct lotteries, distribute obscene literature, and commit fraud. *See* Symposium, *Federalism and the Scope of the Federal Criminal Law*, 26 AM. CRIM. L. REV. 1737, 1738 (1989) [hereinafter Symposium] (noting that law marked the first use by Congress of its postal authority to criminalize conduct that posed no obvious threat to a federal interest).

[13] The history of the mail fraud statute has been recounted by several commentators. *See, e.g.*, John C. Coffee, Jr., *The Metastasis of Mail Fraud: The Continuing Story of the "Evolution" of a White Collar Crime*, 21 AM. CRIM. L. REV. 1, 26–48 (1983) (tracing development of fiduciary fraud in private and public sector); Daniel J. Hurson, *Limiting the Federal Mail Fraud Statute—A Legislative Approach*, 20 AM. CRIM. L. REV. 423, 426–31 (1983) (detailing recent "explosive growth" in mail fraud prosecutions); Donald V. Morano, *The Mail-Fraud Statute: A Procrustean Bed*, 14 J. MARSHALL L. REV. 45, 45–75 (1980) (judicial development of fiduciary fraud); Rakoff, *supra* note 8, at 779–86 (reviewing legislative history of mail fraud statute); Williams, *supra* note 6, at 140–42 (explaining early legislative and judicial history).

[14] The mail fraud statute provides:

> Whoever, having devised or intending to devise any scheme or artifice to defraud, or for obtaining money or property by means of false or fraudulent pretenses, representations, or promises, or to sell, dispose of, loan, exchange, alter, give away, distribute, supply, or furnish or procure for unlawful use any counterfeit or spurious coin, obligation, security, or other article, or anything represented to be or intimated or held out to be such counterfeit or spurious article, for the purpose of executing such scheme or artifice or attempting so to do, places in any post office or authorized depository for mail matter, any matter or thing whatever to be sent or delivered by the Postal Service, or takes or receives therefrom, any such matter or thing, or knowingly causes to be delivered by mail according to the direction thereon, or at the place at which it is directed to be delivered by the person to whom it is addressed, any such

intent to devise a scheme to defraud and a mailing for the purpose of executing the scheme.[15] Through judicial construction, the definitions and consequent scope of mail fraud's two elements developed in an inverse manner.[16] The courts gradually narrowed the "use of the mail" element to a jurisdictional requirement, while they expanded the definition of a "scheme to defraud" to include the loss of intangible rights.[17] This broad definition of fraud was justified as necessary to achieve the statute's fundamental purpose of prohibiting the misuse of the mails to further fraudulent schemes.[18]

The original statute was grounded on the federal interest in protecting the integrity of the mail, and this interest potentially limited the use of the statute.[19] Although the mail fraud statute still requires a minimal link between the scheme and a mailing

---

matter or thing, shall be fined not more than $1,000 or imprisoned not more than five years, or both. If the violation affects a financial institution, such person shall be fined not more than $1,000,000 or imprisoned not more than 30 years, or both.

18 U.S.C. § 1341 (1988 & Supp. IV 1992).

The 1989 amendment, which added the provision regarding a financial institution, did not change the substantive statute. *See* 1989 U.S.C.C.A.N. 86, 196. Likewise, the 1990 amendment increasing the penalty to 30 years if a financial institution is involved did not alter the substance of the statute.

Under the Sentencing Guidelines, an elected official found guilty of depriving citizens of their intangible rights to honest services will serve a minimum sentence of 27 to 33 months, with time added if the monetary loss exceeded $2,000. *See* UNITED STATES SENTENCING GUIDELINES § 2C1.7 (1992).

[15] *See* Pereira v. United States, 347 U.S. 1 (1954); United States v. Young, 232 U.S. 155 (1914); United States v. Mandel, 591 F.2d 1347, 1360 n.9 (4th Cir.) (citing United States v. Brewer, 528 F.2d 492, 494 (4th Cir. 1975)), *conviction aff'd per curiam in relevant part on reh'g en banc*, 602 F.2d 653 (4th Cir. 1979), *cert. denied*, 445 U.S. 961 (1980).

The original statute's third element, an intent to effect a fraud by using the mail, was eliminated in 1909. The original statute covered a "person [who] devised or intend[ed] to devise any scheme or artifice to defraud, or [to] be effected by either opening or intending to open correspondence." Act of June 8, 1872, ch. 335, § 301, 17 Stat. 323. The 1909 amendment eliminated this language. Act of Mar. 4, 1909, ch. 321, § 215, 35 Stat. 1130. *See* Stokes v. United States, 157 U.S. 187 (1895); Morano, *supra* note 13, at 45 n.2; *see generally* Rakoff, *supra* note 8.

[16] *See* Roger J. Miner, *Federal Courts, Federal Crimes, and Federalism*, 10 HARV. J.L. & PUB. POL'Y, 117, 121 (1987) (arguing that judicial interpretation has converted mail fraud into a "vehicle for the prosecution of an almost unlimited number of offenses bearing very little connection to the mails"); Morano, *supra* note 13, at 47 (asserting that courts have construed mail fraud statute to address virtually any conduct involving deception).

[17] *See* Rakoff, *supra* note 8, at 778 (judicial focus on the mail, which emphasized that purpose of statute was to prevent "misuse of the mails," allowed courts to ignore the need to define closely "a scheme to defraud").

[18] *See* Durland v. United States, 161 U.S. 306 (1896); United States v. Margiotta, 688 F.2d 108 (2d Cir. 1982), *cert. denied*, 461 U.S. 913 (1983).

[19] *But see* Rakoff, *supra* note 8, at 819 (arguing that mailing requirement has always served primarily as a basis for federal jurisdiction).

that furthers the scheme to defraud,[20] that link is established if the use of the mail is "reasonably foreseeable" by the defendant.[21] Essentially, the federal interest in the mail is now merely used to establish jurisdiction.

Conversely, courts broadly interpreted the second element, "a scheme to defraud," in order to expand the reach of the statute and, eventually, to formulate a new kind of fraud. This expansion was readily accomplished because the statute's language immediately presented an opportunity for judicial interpretation:[22] "Whoever, having devised or intending to devise any scheme or artifice to defraud, *or* for obtaining money or property by means of false or fraudulent pretenses . . . ."[23] Instead of punishing a scheme to defraud that resulted or would result in obtaining money fraudulently, the courts chose to punish both "devising a scheme to defraud" and "obtaining money by false pretenses."[24] The courts, however, like the legislature, did not define a "scheme to defraud."[25] The lack of a definition invited and resulted in

---

[20] In order to constitute a "mailing" under the statute, the defendant must use the mail to execute a fraud. *See* United States v. Maze, 414 U.S. 395 (1974) (merchant's mailing of credit card sales slips to bank and bank's subsequent mailing of statement to card owner do not constitute mailing because success of fraudulent scheme did not depend on who paid); Parr v. United States, 363 U.S. 370 (1960) (tax bills used to secure funds were not mailings because they did not further scheme, as bills were neither illegal nor fraudulent); Kann v. United States, 323 U.S. 88 (1944) (post-transaction mailings occurred after fraud completed and were therefore not for purpose of executing scheme). *See also* Schmuck v. United States, 489 U.S. 705, 723 (1989) (Scalia, J., dissenting) (arguing that statute prohibits "mail fraud," not "mail and fraud"). *But see* United States v. Sampson, 371 U.S. 75 (1962) (ruling that a mailing subsequent and unnecessary to successful completion of fraudulent scheme but which lulled victim into complacence constituted a mailing under statute).

[21] *See* Pereira v. United States, 347 U.S. 1, 8–9 (1954) (holding that when defendant knows that use of mail will follow in ordinary course of business or can reasonably foresee such use, defendant "causes" mails to be used); United States v. Bryant, 766 F.2d 370 (8th Cir.) (in wire fraud context, statute satisfied if defendant could reasonably foresee that interstate wire would be sent), *cert. denied*, 474 U.S. 1054 (1985).

Recent cases have further eroded the value of the mailing requirement as a possible limitation on the statute's applicability. *See, e.g.*, *Schmuck*, 489 U.S. at 706 (holding that mailing in scheme that was consummated before mailing occurred was within statute if scheme was ongoing and continuous). *See generally* Morano, *supra* note 13, at 50–53 & n.20 (reviewing mailing element); Rakoff, *supra* note 8 (same).

[22] *See* Rakoff, *supra* note 8, at 775–77 (commenting on "interesting" and "idiosyncratic" characteristics of mail fraud statute).

[23] 18 U.S.C. § 1341 (emphasis added).

[24] *See, e.g.*, United States v. Bohonus, 628 F.2d 1167 (9th Cir.), *cert. denied*, 447 U.S. 928 (1980); United States v. McNeive, 536 F.2d 1245, 1248–49 (8th Cir. 1976).

This conclusion was by no means inevitable. One possible construction of the language is that it addresses the complementary objectives of the defendant. "To defraud" encompasses the victim's loss and "to obtain money or property" encompasses the defendant's gain. *See* Morano, *supra* note 13, at 53–54.

[25] *See* United States v. Maze, 414 U.S. 395 (1974); Blachly v. United States, 380 F.2d 665, 671 (5th Cir. 1967) (refusing to define "defraud" because doing so "would tend

"virtually open-ended liability."[26] The phrase, "a scheme to defraud," came to prohibit a plan, that is, to forbid a state of mind, rather than physical conduct.[27] The demise of the mailing element coupled with the emphasis on "scheme to defraud" means that the statute essentially prohibits the actor from intentionally devising a scheme. A mail fraud charge may thus be based solely on intent.[28]

Mail fraud is an inchoate crime; the actor need not cause actual harm by achieving a fraudulent goal to violate the statute.[29] Because it emphasizes guilt rather than consequences, the law has, in the interest of crime prevention, prohibited inchoate crimes and attached criminal culpability at an earlier point in time.[30] In contrast to attempt and conspiracy, which assign culpability for a substantive offense, mail fraud is facially inchoate—the substantive offense itself requires only that the actor scheme to defraud.[31] In addition, because the actor may be

---

to reward subtle and ingenious circumvention"); KATHLEEN F. BRICKEY, CORPORATE AND WHITE COLLAR CRIME 93 (1990) (noting failure to define the term "defraud").

[26] *See* Coffee, *supra* note 13, at 11 (substantive crime of mail fraud consists of little more than an evil scheme); Hurson, *supra* note 13, at 425 (characterizing mail fraud as "element-free liability").

[27] Hurson, *supra* note 13, at 450 (in mail fraud, "criminal liability is defined by the intent of the accused").

[28] One commentator has suggested that if the mailing is indeed only a jurisdictional act, the statute may not define a crime since it would not prohibit overt and reprehensible conduct. *See* Rakoff, *supra* note 8, at 775.

It is axiomatic that in order to incur criminal liability, the mental and physical elements must converge. *See* 4 WILLIAM BLACKSTONE, COMMENTARIES *21 ("[T]o make a complete crime, . . . there must be both a will and an act."); 2 JAMES F. STEPHEN, A HISTORY OF THE CRIMINAL LAW OF ENGLAND 78 (London, MacMillan 1883) (stating that sinful thoughts are never punished because if they were, "all mankind would be criminals").

[29] *See* Durland v. United States, 161 U.S. 306, 315 (1896) (mailings do not have to be effective); United States v. Utz, 886 F.2d 1148, 1151 (9th Cir.) (unsuccessful scheme to defraud remains within purview of § 1341), *cert. denied*, 497 U.S. 1005 (1990); United States v. Bronston, 658 F.2d 920, 927 (2d Cir. 1981) (government did not have to show that victims were in fact defrauded), *cert. denied*, 456 U.S. 915 (1982); United States v. George, 477 F.2d 508, 512 (7th Cir.) ("[I]t was unnecessary for the Government to prove 'that anyone actually be defrauded.'"), *cert. denied*, 414 U.S. 827 (1973); United States v. Regent Office Supply Co., 421 F.2d 1174, 1179–81 (2d Cir. 1970) (prosecutor need not show that intended victim was actually harmed).

[30] *See* Ira P. Robbins, *Double Inchoate Crimes*, 26 HARV. J. ON LEGIS. 1, 5–13 (1989). The crime of attempt punishes behavior when the harm intended by the actor and criminalized in a substantive criminal statute did not occur because law enforcement agents intervened, because the actor failed to complete the act, or because it was impossible to achieve the goal. Conspiracy and solicitation are also inchoate in that each criminalizes an agreement to commit a subsequent, substantive crime. Reckless endangerment statutes, enacted to punish those who put others at risk of harm, are inchoate in that they do not require tangible harm, or even an intent to inflict harm.

[31] There is no general federal statute proscribing attempt, and federal criminal statutes are written to include the attempt to commit the substantive crime as well as

charged upon an intention to devise a scheme, that is, planning to plan, mail fraud is, even on its face, doubly inchoate.[32] The inchoate nature of mail fraud increases the significance of intent and results in the criminalization of the deception, rather than the end to be achieved by the deception.[33]

Another consequence of the inchoate nature of mail fraud is that courts effectively use the mailing element to corroborate the actor's intent to defraud.[34] A third consequence and yet another anomaly is that repeated mailings give rise to additional counts of mail fraud even though the defendant intended only one scheme, or intended to devise only one scheme, to defraud.[35]

The reduction of the mailing element to a jurisdictional factor, the open-ended construction of "scheme to defraud," and the inchoate nature of the crime presented a unique opportunity to courts and prosecutors.

---

the substantive offense. *See* United States v. Everett, 700 F.2d 900, 903 n.7 (3d Cir. 1983); United States v. Manley, 632 F.2d 978, 987 (2d Cir. 1980); United States v. Alvarez, 610 F.2d 1250, 1253 (5th Cir. 1980); United States v. York, 578 F.2d 1036 (5th Cir. 1978).

As a consequence, mail fraud is a specific intent crime requiring that the actor specifically intends to devise a scheme to defraud. *See* United States v. Schwartz, 924 F.2d 410, 420 (2d Cir. 1991) (relying on United States v. Starr, 816 F.2d 94, 98 (2d Cir. 1987) to hold that government must present proof that defendant possessed fraudulent intent); United States v. Mandel, 591 F.2d 1347, 1363 & n.11 (4th Cir.) (stating that prosecutor had to prove the requisite intent to defraud), *conviction aff'd per curiam in relevant part on reh'g en banc*, 602 F.2d 653 (4th Cir. 1979), *cert. denied*, 445 U.S. 961 (1980); *Regent Office Supply Co.*, 421 F.2d at 1179–81 (holding that government must show defendant contemplated doing actual harm, that is, something more than merely deceiving victim). Specific intent, however, may be proven by inferences from the circumstances. *See* John C. Coffee, Jr., *From Tort to Crime: Some Reflections on the Criminalization of Fiduciary Breaches and the Problematic Line Between Law and Ethics*, 19 AM. CRIM. L. REV. 117, 126 n.47 (1981).

[32] *See* Morano, *supra* note 13, at 57 n.36 (commenting that those who plan to plan a scheme may be punished).

Carried to its natural conclusion, this principle may lead to liability for triple and even quadruple inchoate crimes—conspiracy to commit mail fraud is triply inchoate. *See* United States v. Isaacs, 493 F.2d 1124, 1152 (7th Cir.) (defendants charged with both mail fraud and conspiracy to commit mail fraud), *cert. denied*, 417 U.S. 976 (1974). Theoretically, attempted conspiracy to commit mail fraud is inchoate to the fourth degree; a person might be punished for trying to agree to plan a plan to defraud another person.

[33] *See* Coffee, *supra* note 31, at 167 (criminalization of deceit or failure to disclose when it is the means to an illegitimate end ignores the ends-means distinction).

[34] *See* Rakoff, *supra* note 8, at 776–77, 821 (urging courts to view mailing as overt act and jurisdictional element); Peter T. Barbur, *Mail Fraud and Free Speech*, 61 N.Y.U. L. REV. 942, 945 (1986) (noting that mailing requirement is similar to overt act requirement of conspiracy statute).

[35] *See* Hurson, *supra* note 13, at 459 (offering proposal that eliminates "abuse of multicount" indictments); Williams, *supra* note 6, at 144–45 & n.47 (noting that stacking counts of mail fraud to increase penalties is frowned upon by criminal law scholars).

B. *The Judicial Construction of a "Scheme to Defraud"*—Mandel, Margiotta, *and* McNally

The courts dramatically expanded the scope of the mail fraud statute by ignoring the principles of common-law fraud, which require a tangible loss, and, thereby, formulating the intangible rights doctrine. This doctrine defines "a scheme to defraud" to include conduct that does not cause a victim to lose money or property.[36] In classic fraudulent schemes, the actor defrauds a victim of money or property by misrepresenting a material fact. The victim suffers a loss of money or property and the perpetrator benefits by the amount of the loss. In contrast, fiduciary fraud,[37] which may be perpetrated by a failure to disclose a material fact, deprives the principal of a right to the fiduciary's honest and faithful services. Such cases typically occur when the fiduciary receives a benefit from a third party—such as the insurance company in the example given earlier—even though the fiduciary has not misled the third party. Theoretically, the arrangement between the fiduciary and the third party deprives the victim of a right to the intangible value of the fiduciary's faithful and loyal services.[38]

Mail fraud, through the intangible rights doctrine, has been expanded to the point that a fiduciary, agent, or employee commits an offense when, through a material deception or a failure to disclose, a beneficiary, principal, or employer suffers even an intangible, constructed detriment.[39] Since the mid-1970s, the

---

[36] Courts supported these prosecutions by constructing an intangible loss to the victim. A classic example is Learned Hand's formulation of harm:

> A man is none the less cheated out of his property, when he is induced to part with it by fraud, because he gets a quid pro quo of equal value. It may be impossible to measure his loss by the gross scales available to a court, but he has suffered a wrong; he has lost his chance to bargain with the facts before him. That is the evil against which the statute is directed.

United States v. Rowe, 56 F.2d 747, 749 (2d Cir. 1932) (finding no error in conviction even where victim received consideration of equal value to that which was relinquished).

[37] Fiduciary fraud involves intangible rights and interests when a fiduciary relationship, such as that of employer-employee, exists. *See* Morano, *supra* note 13, at 48.

[38] This theory was initially developed in prosecutions for fraudulent commercial schemes in the private sector. *See* United States v. George, 477 F.2d 508 (7th Cir.), *cert. denied*, 414 U.S. 827 (1973). In that case, the court upheld the mail fraud conviction of a purchasing agent who accepted kickbacks from the seller, although the employer suffered no economic harm. The court imputed economic harm, reasoning that the employer lost a possible discount or the opportunity to bargain with knowledge of a relevant fact.

[39] *See* Hurson, *supra* note 13, at 429 (reviewing expansion of mail fraud statute).

courts, pressured by federal prosecutors, have applied the intangible rights doctrine to prosecutions of public officials.[40] The prosecutions are based on the notion that a public official is a fiduciary of the citizenry and owes fiduciary duties of honesty and loyalty. By breaching those duties, an official deprives citizens of the intangible right to his honest government.

Although the prosecution of Marvin Mandel, Governor of Maryland, was not the first to utilize the intangible rights theory in a political corruption case,[41] the Fourth Circuit opinion upholding his conviction for mail fraud presents a fairly clear articulation of the doctrine.[42] Essentially, Mandel received financial benefits—clothing, jewelry, and an interest in two real estate ventures—from a racing association and failed to disclose his involvement with the association to state legislators who were considering a bill benefiting the association.[43] Mandel was

---

[40] In 1975, prompted by the Watergate scandal, President Gerald Ford directed federal prosecutors to target political corruption on the state and local levels. The Department of Justice established its Public Integrity Section in 1976. Interestingly, it was not until the mid-1980s that prosecutors began to focus on federal officials. *See* Andrew T. Baxter, *Federal Discretion in the Prosecution of Local Political Corruption*, 10 PEPP. L. REV. 321, 321–22 (1983) (recounting history of federal campaign against state and local political corruption); Arthur Maass, *Public Policy by Prosecution*, THE PUB. INTEREST, Fall 1987, at 107, 111 (noting that until the 1970s citizen committees, state legislative committees, and local and state prosecutors led campaigns against political corruption).

The campaign against local political corruption immediately led to more prosecutions. *See* Ralph E. Loomis, Comment, *Federal Prosecution of Elected State Officials for Mail Fraud: Creative Prosecution or an Affront to Federalism?*, 28 AM. U. L. REV. 63, 64 n.12 (1978) (noting that federal indictments of state and local officials increased 500%, from 63 in 1970 to 337 in 1976).

[41] The intangible rights theory was first suggested in Shushan v. United States, 117 F.2d 110, 115 (5th Cir. 1941), *cert. denied*, 313 U.S. 574 (1942). In *Shushan*, the court, in dicta, suggested that public corruption might come within the ambit of the mail fraud statute. The court wrote: "No trustee has more sacred duties than a public official and any scheme to obtain an advantage by corrupting such a one must in the federal law be considered a scheme to defraud." *Id.* Other early cases that applied the intangible rights doctrine to political corruption include United States v. States, 488 F.2d 761 (8th Cir. 1973), *cert. denied*, 417 U.S. 909 (1974); United States v. Keane, 522 F.2d 534 (7th Cir. 1975), *cert. denied*, 424 U.S. 976 (1976); United States v. Isaacs, 493 F.2d 1124 (7th Cir.), *cert. denied*, 417 U.S. 976 (1974); United States v. Bush, 522 F.2d 641 (7th Cir. 1975), *cert. denied*, 424 U.S. 977 (1976); United States v. McNeive, 536 F.2d 1245 (8th Cir. 1976).

[42] United States v. Mandel, 591 F.2d 1347, 1358 (4th Cir.), *conviction aff'd per curiam in relevant part on reh'g en banc*, 602 F.2d 653 (4th Cir. 1979), *cert. denied*, 445 U.S. 961 (1980).

[43] *Id.* at 1356. At the relevant time, the legislature was considering—and eventually passed—a "consolidation bill" providing the racing association with which Mandel was involved an increase of its racing days from 36 to 94 and the use of one-mile tracks. *Id.* at 1355. Mandel actively supported the legislation. Several legislators testified that the identity of the owners and Mandel's relationship with them would have been relevant to their voting decision. *Id.*

charged with mail fraud, that is, of devising a scheme to defraud the citizens of Maryland of their right to the governor's honest and faithful services.

The court invoked the purpose of the statute—to prohibit the misuse of the mails—to endorse a broad definition of "a scheme to defraud."[44] The opinion noted that bribery, failure to disclose, and concealment might not always support a charge of mail fraud because these acts, standing alone, were not fraudulent. After recognizing this potential limitation, however, the court abandoned it, holding that acceptance of a benefit, failure to disclose, or concealment satisfied the fraud element of the statute. In so holding, the court accepted the premise that citizens have a right to honest and loyal services that they could lose through fraud. According to the court, accepting a bribe satisfies the fraud element because the public is not receiving the public official's honest, faithful, and disinterested services to which it is entitled.[45] Failing to disclose the existence of a direct interest in a matter the official is considering satisfies the fraud element because it defrauds the public of its intangible right to honest, loyal, faithful, and disinterested government.[46] Similarly, concealing material information from a public body and receiving an illicit benefit through an action of that public body satisfies the fraud element because the public is deprived of the right to have its officials act on accurate information.[47]

Three years later, the Second Circuit expanded the intangible rights doctrine in the public sector by applying it to unelected officials. In *United States v. Margiotta*,[48] the Second Circuit

---

[44] The court stated that any scheme that is contrary to public policy or that conflicts with accepted standards of moral uprightness, fundamental honesty, fair play, and right dealing may constitute mail fraud. *Id.* at 1361 (relying principally on Badders v. United States, 240 U.S. 391, 393 (1916) and Gregory v. United States, 253 F.2d 104, 109 (5th Cir. 1958)). The court, noting that Congress had left the task of defining "a scheme to defraud" to the judiciary, acknowledged that the resulting judicial definition meant that a defendant could be guilty of mail fraud without having transgressed a federal or state statute or the common law. *Id.* at 1360.

[45] *Id.* at 1362 (relying on Shushan v. United States, 117 F.2d 110, 115 (5th Cir. 1941), *cert. denied*, 313 U.S. 574 (1942) and United States v. Brown, 540 F.2d 364, 374 (8th Cir. 1976)).

[46] The court noted that the duty to disclose is not based on statutory law but on the relationship of the parties. It found a fiduciary duty of honesty and loyalty between a government official and the official's constituents. *Id.* at 1363.

[47] *Id.* at 1363–64.

[48] 688 F.2d 108 (2d Cir. 1982), *cert. denied*, 461 U.S. 913 (1983). The indictment alleged that Margiotta, the chairman of both the county and town Republican parties, engaged in a scheme to distribute insurance commissions worth $500,000 for insuring municipal properties to political associates. *Id.* at 113–14.

ruled that an unelected individual, the chairman of the local Republican party, owed a fiduciary duty to the general electorate. The court held that a fiduciary duty, which encompassed a duty to disclose material information, arose from the implicit relationship between Margiotta and the electorate.[49] The court then ruled that Margiotta's failure to disclose a scheme to allocate insurance commissions to political associates breached that duty and deprived citizens of "intangible and abstract political and civil rights of the general citizenry."[50]

These developments did not go unnoticed by federal prosecutors, who acknowledged the power of the mail fraud statute by praising it.[51] Neither did these developments go unappreciated; federal prosecutors paid the statute the ultimate compliment of using it extensively.[52] The increased use of mail fraud stimulated criticism from commentators who lamented the debasement of principles of federalism and constitutional law.[53] In 1987, the Supreme Court responded.

---

[49] The court endorsed two "time-tested measures" of fiduciary duty: a reliance test, which finds a fiduciary duty when others rely on the official because of a special relationship in the government, and a de facto control test, which finds a fiduciary duty when a person in fact makes governmental decisions. *See id.* at 122.

[50] *Id.* at 121. *See also* Hurson, *supra* note 13, at 436–39 (providing comprehensive analysis of *Margiotta*).

[51] *See* Coffee, *supra* note 31, at 126 (quoting prosecutor's maxim, "when in doubt, charge mail fraud"); Coffee, *supra* note 13, at 3 (mail fraud statute "seems destined to provide the federal prosecutor with what Archimedes long sought—a simple fulcrum from which one can move the world"); Rakoff, *supra* note 8, at 771 (quoting prosecutors as referring to the statute as "our Stradivarius, our Colt 45, our Louisville Slugger, our Cuisinart").

[52] *See* Adam H. Kurland, *The Guarantee Clause as a Basis for Federal Prosecutions of State and Local Officials*, 62 S. CAL. L. REV. 367, 398 (1989) (noting expanded use of mail fraud by United States Attorneys); *see supra* text accompanying note 5 for statistics demonstrating increase in political corruption cases.

[53] *See generally* Baxter, *supra* note 40 (noting constitutional issues of federalism and separation of powers arising from prosecutorial discretion and recommending Department of Justice guidelines to limit enforcement); Coffee, *supra* note 31 (criticizing federal enforcement of breach of fiduciary duty on ground of over-criminalization); Coffee, *supra* note 13 (arguing that intangible rights doctrine should be constrained); Hurson, *supra* note 13 (arguing for specific definition of "scheme to defraud"); John C. Jeffries, Jr., *Legality, Vagueness, and the Construction of Penal Statutes*, 71 VA. L. REV. 189 (1985) (critiquing *Margiotta*); Maass, *supra* note 40 (arguing that federalism requires states to be free of federal government interference in establishment of political institutions); Miner, *supra* note 16 (criticizing mail fraud statute as example of federalization of criminal law impinging on area of state concern and discussing negative consequences); Morano, *supra* note 13 (arguing that due process is offended by use of intangible rights doctrine); W. Robert Gray, Comment, *The Intangible Rights Doctrine and Political-Corruption Prosecutions Under the Federal Mail Fraud Statute*, 47 U. CHI. L. REV. 562, 587 (1980) (surveying history of mail fraud statute and concluding that use of statute to fight political corruption is not supported by congressional intent); Loomis, *supra* note 40, at 85 (arguing that although mail fraud prosecu-

McNally v. United States[54] overturned two decades of circuit and trial court decisions implementing the intangible rights doctrine, and suggested an awakened interest in federalism. The Court held that the mail fraud statute was limited to the protection of property rights.[55] The facts of *McNally* raise the same issues seen in *Mandel* and *Margiotta*—fiduciary fraud, intangible rights, and public officials. The defendants in *McNally* were convicted of mail fraud for operating a self-dealing patronage scheme that channeled insurance commissions to companies controlled by the defendants, one of whom was the chairman of Kentucky's Democratic Party.[56] The scheme did not result in any monetary or property loss to Kentucky,[57] violate Kentucky law, or violate any other federal law.[58] Consequently, the prosecution, the district court, and the appellate court relied on the substantial

---

tion implicates federal interests, potential impact on state political process mandates enhanced judicial review).

[54] 483 U.S. 350 (1987). *McNally* has been the subject of considerable comment. *See generally* Carey et al., *supra* note 6, at 336 (asserting that *McNally* created a gap in prosecution of public officials); John C. Coffee, Jr., *Hush!: The Criminal Status of Confidential Information After* McNally *and* Carpenter *and the Enduring Problem of Overcriminalization*, 26 AM. CRIM. L. REV. 121, 127 (1988) (commenting that *McNally*'s failure to define property vitiated its holding); Oxman, *supra* note 9 (predicting that the impact of *McNally* on prosecutions for mail fraud in political context would be severe); Williams, *supra* note 6 (applying different methods of statutory interpretation to test *McNally* reasoning); Caroline Aiken et al., Note, *Mail and Wire Fraud*, 28 AM. CRIM. L. REV. 591 (1991) (analyzing mail fraud in light of *McNally*); Jeffrey J. Dean & Doye E. Green, Jr., Note, McNally v. United States *and Its Effect on the Federal Mail Fraud Statute: Will White Collar Criminals Get a Break?*, 39 MERCER L. REV. 697 (1988) (analyzing decision); Donna M. Ducey, Note, McNally v. United States: *The Demise of the Intangible Rights Doctrine*, 66 N.C. L. REV. 1035 (1988) (placing *McNally* in context of movement toward strict construction of statutes); M. Diane Duszak, Note, *Post-*McNally *Review of Invalid Convictions Through the Writ of Coram Nobis*, 58 FORDHAM L. REV. 979 (1990) (tracing effect of *McNally* on previous prosecutions); Donna M. Maus, Note, *License Procurement and the Federal Mail Fraud Statute*, 58 U. CHI. L. REV. 1125 (1991) (concluding that licenses fall outside of *McNally*'s property requirement); Keith S. Hampton, Casenote, 19 ST. MARY'S L.J. 1115 (1988) (concluding that *McNally* correctly limited unchecked discretion of federal officials because state voters have undisputed authority over their state governments).

[55] *McNally*, 483 U.S. at 360. The decision was unexpected. *See id.* at 376 (Stevens, J., dissenting) (noting that 20 years of circuit and trial court decisions spoke against limiting the statute); *see also* Williams, *supra* note 6, at 138 (noting that critics had not succeeded in generating circuit court reinterpretation, legislative amendment, or Department of Justice guidelines).

[56] *See McNally*, 483 U.S. at 352–53. One defendant controlled the selection of the insurance agents who received the state's insurance contracts. *Id.*

[57] *See id.* at 360. The budgeted funds for insurance were simply channeled to certain firms, one of which was owned by the defendants; although the defendants benefited financially, the state did not suffer a financial loss.

[58] *See* Kurland, *supra* note 52, at 400 (analyzing *McNally*).

precedent establishing that a mail fraud conviction could be based on a loss of the intangible right to honest government.[59]

The Supreme Court based its reversal on a close reading of the statute and the plain meaning of "defraud."[60] The Court concluded, despite the disjunctive "or", that the phrase "for obtaining money or property" clarified "scheme to defraud."[61] Thus, the Court constructed the statute to criminalize only those frauds that involve money or property.[62] Furthermore, the Court found no evidence in its review of the legislative history that Congress intended anything other than protection from schemes to defraud people of money or property.[63] An intangible right to honest and impartial government was not property, therefore the mail fraud statute did not address its loss.[64] The Court's technical holding was informed by a concern for federalist principles: "Rather than construe the statute in a manner that leaves its outer boundaries ambiguous and involves the Federal Government in setting standards of disclosure and good government for local and state officials, we read section 1341 as limited in scope to

---

[59] *See McNally*, 483 U.S. at 355.

[60] The relevant text of 18 U.S.C. § 1341 reads: "Whoever, having devised or intending to devise any scheme or artifice to defraud, or for obtaining money or property . . . ." For the full text of the mail fraud statute, *see supra* note 14.

[61] The Court rejected the argument that the clauses should be read independently, which is what had led to the open-ended interpretation of "scheme to defraud." *See McNally*, 483 U.S. at 358–59.

[62] This interpretation was required by the principle of lenity. *Id.* at 359–60 (noting that when two rational readings of criminal statute are possible, lenity allows courts to choose harsher reading only when congressional intent is clear).

[63] *See id.* at 356. *See also* Gray, *supra* note 53, at 569–72 (concluding that 1909 amendment was designed to remove only common-law defense regarding future promises and legislators did not intend that scheme to defraud be independent of "obtaining money or property").

[64] This argument is circular because the right to "honest service" would be property if the Court had assigned citizens a right to it by finding that it was protected by the statute. Felix S. Cohen identified the logical flaw indulged in by courts that limit themselves to legal definitions, such as "property," previously formulated to define other concepts. According to Cohen, "[t]o justify or criticize legal rules in purely legal terms is always to argue in a vicious circle." Felix S. Cohen, *Transcendental Nonsense and the Functional Approach*, 35 COLUM. L. REV. 809, 814 (1935).

Arguably, the Court erred in its restricted definition of property because property rights also encompass legal relations between people. "Sometimes [the word 'property'] is employed to indicate the physical object to which various legal rights, privileges, etc., relate; then again—with far greater discrimination and accuracy—the word is used to denote the legal interest." Wesley N. Hohfeld, *Some Fundamental Legal Conceptions as Applied in Judicial Reasoning*, 23 YALE L.J. 16, 21 (1913). *See generally* Kenneth J. Vandevelde, *The New Property of the Nineteenth Century: The Development of the Modern Concept of Property*, 29 BUFF. L. REV. 325 (1980). *See also* Maus, *supra* note 54, at 1137 & n.71 (criticizing the Court for using the word "property" as if it had some intrinsic meaning and thereby masking the circularity of its reasoning).

the protection of property rights."[65] In a prescient final statement, the Court added that if Congress wanted to do more, "it must speak more clearly than it has."[66]

C. *The Congressional Response*

Congress immediately responded to the Court's challenge and, within the year, amended the mail fraud statute and reinstated the intangible rights doctrine. The new law defines "scheme or artifice to defraud" to include a scheme "to deprive another of the intangible right of honest services."[67]

The amendment, the result of an eleventh-hour compromise in the passage of the "amendment-ridden," election-year Anti-Drug Abuse Act of 1988, has a clear but sparse legislative history.[68] Supporters of the amendment stated that the purpose of the legislation was to overrule *McNally* and to enable the government to prosecute under the intangible rights doctrine whenever public corruption schemes deprive citizens of honest government. Representative Conyers (D-Mich.) explained that the amendment "restores the mail fraud provision to where [it] was before the *McNally* decision."[69] Conyers also stated that the amendment, if passed, would render it "no longer necessary to determine whether or not the scheme . . . involved money or property."[70]

---

[65] *McNally*, 483 U.S. at 360.

[66] *Id.*

[67] Anti-Drug Abuse Act of 1988, Pub. L. No. 100-690, § 7603, 102 Stat. 4181, 4508. The full text of the amendment reads:

> Definition of "scheme or artifice to defraud"
>
> For the purposes of this chapter, the term "scheme or artifice to defraud" includes a scheme or artifice to deprive another of the intangible right of honest services.

18 U.S.C. § 1346 (1988).

[68] *See* Kurland, *supra* note 52, at 489 n.452 (criticizing amendment).

[69] 134 CONG. REC. H11,251 (daily ed. Oct. 21, 1988) (Part II) (statement of Rep. Conyers (D-Mich.)); *id.* at S17,308 (statement of Sen. McConnell (R-Ky.)); *see generally Mail Fraud Hearings*, *supra* note 6.

Representative Conyers ultimately voted against the bill, so it is unclear what weight should be accorded his interpretation of the amendment. *See* David B. Sweet, Annotation, *Validity, Construction, and Application of Federal Mail Fraud Statute*, 97 L. Ed. 2d 863, 878 n.24 (1987) (outlining legislative history of § 1346); United States v. Berg, 710 F. Supp. 438, 440–45 (E.D.N.Y. 1989) (reviewing legislative history of § 1346), *rev'd sub nom.* United States v. Schwartz, 924 F.2d 410 (2d Cir. 1991).

[70] 134 CONG. REC. H11,251 (daily ed. Oct. 21, 1988) (Part II) (statement of Rep. Conyers) (noting that amendment also applies to wire fraud).

Through the amendment, Congress rejected the Supreme Court's decision that the mail fraud statute, as written, did not protect intangible rights. The Supreme Court based its ruling in *McNally* on technical considerations of statutory construction and did not consider whether the intangible rights doctrine violated constitutional provisions.[71] Furthermore, there is no record that Congress considered the constitutionality of the intangible rights doctrine. The codification of the pre-*McNally* standard, therefore, merely transferred the focus of concerns from case law to statutory law, and neither considered nor resolved them.[72]

The resurrection of the intangible rights doctrine requires courts to evaluate the constitutional implications of federal prosecutions based on the doctrine. That evaluation should result in findings that the amended statute is void for vagueness because intangible rights are undefinable. Before turning to the vagueness challenge, however, it is necessary to analyze the effect of the doctrine on three constitutional provisions—an effect that renders the statute facially vague and therefore unconstitutional.

II. The Constitutional Implications of the Intangible Rights Doctrine

The statutory right to honest services, intended by Congress to mean a right to good government, is an undefined, amorphous

---

[71] *See* Coffee, *supra* note 54, at 127 (commenting that *McNally* only implicitly vindicated the view that the intangible rights doctrine had resulted in a serious over-extension of criminal law).

[72] *See* Kurland, *supra* note 52, at 489–90 (amendment "merely reinstates and perpetuates the irrational loopholes and inconsistencies" inherent in the combination of intangible rights and use of the mail); Williams, *supra* note 6, at 167–68 (amendment merely forces courts to devise definition of "honest services"); Carey et al., *supra* note 6, at 327 (unclear whether language accomplished purpose articulated in legislative history because "honest service" may not be equivalent to "good government").

The courts have construed tangible property loss very broadly so that they need not reach the ramifications of "honest services." *See, e.g.*, United States v. Johns, 742 F. Supp. 196 (E.D. Pa. 1990) (salary paid in reliance that employee's services would be free of conflict of interest constituted loss of property). *But see* United States v. Telink, Inc., 910 F.2d 598 (9th Cir. 1990) (upholding dismissal because indictment was permeated with pre-*McNally* "right to honest services").

Such circumvention is facilitated by the holding of United States v. Carpenter, 484 U.S. 19 (1987), in which the Court upheld a mail fraud conviction based on a loss of intangible property rights in the private sector. This decision, which recognized intangible property rights in information, has been viewed as emasculating *McNally*. *See* Oxman, *supra* note 9, at 768 (*Carpenter* has allowed courts to continue hearing intangible rights cases under mail fraud statute); Coffee, *supra* note 54, at 140–42 (discussing implications of *Carpenter*).

right with uncertain boundaries. Although the intangible rights doctrine may allow federal authorities to control state and local political corruption more effectively than a more narrowly defined statute, it diminishes state and local autonomy. Further, the intangible rights doctrine offends the separation of powers doctrine and discourages political activity protected by the First Amendment.[73]

The issue of prosecutorial discretion permeates all of these constitutional considerations. The discretion of prosecutors is enhanced in cases involving intangible rights because prosecutors are not constrained by clearly defined statutory elements.[74] Excessive prosecutorial discretion exacerbates the deleterious effect of the intangible rights doctrine on federalism, the separation of powers, and the First Amendment.

A. *The Intangible Rights Doctrine and Federalism*

In *McNally*, the Supreme Court rejected federal efforts to dictate the standard for good government at the state and local levels.[75] The following discussion explores the basis of the Court's concern for federalism.[76]

Adding to the confusion that *Carpenter* caused, Justice Stevens, dissenting in *McNally*, suggested that salary paid to an employee is a form of property belonging to the employer, and that illicit gains are subject to a constructive trust. *See McNally*, 483 U.S. at 377 n.10 (Stevens, J., dissenting). In light of *Carpenter* and Justice Stevens's dissent, the Department of Justice advises prosecutors to formulate mail fraud indictments to incorporate those theories. *See* U.S. DEP'T OF JUSTICE, PROSECUTION OF PUBLIC CORRUPTION CASES 445 (1988) (suggesting salary, control over property, and constructive trusts as possible substitutes for tangible property).

An ancillary mail fraud case law has developed that interprets the related effects of *McNally* and the passage of § 1346. For instance, defendants convicted before *McNally* are applying for writs of *coram nobis* or acquittal. Defendants sentenced after the passage of § 1346 for acts committed before its passage but after *McNally* are claiming the benefits of ex post facto bars. *See generally* Duszak, *supra* note 54, at 979.

[73] The application of the doctrine may also offend various other provisions of the Bill of Rights. *See infra* note 118 and sources cited therein. My focus on three structural constitutional provisions is not intended to imply that such violations are less compelling or less damaging to individuals, and the mail fraud statute proposed in Part IV also alleviates the effect of the intangible rights doctrine on individual civil rights.

[74] *See supra* notes 14–35 and accompanying text (discussing elements of mail fraud).

[75] *See McNally*, 483 U.S. at 360 (rejecting interpretation that would involve "Federal Government in setting standards of disclosure and good government for local and state officials").

[76] *McNally* renewed the discussion about the limitations federalism imposes on the authority of Congress to enact criminal statutes. *See* Evans v. United States, 112 S. Ct. 1881, 1894–96 (1992) (Thomas, J., dissenting) (invoking federalist principles in argument rejecting expansive interpretation of the Hobbs Act). *See generally* Symposium, *supra* note 12; Miner, *supra* note 16.

The subject of federalism polarizes people: some invoke federalism as a significant constraint on national legislation while others reject federalism as an irrelevant consideration. The conclusion required by the theoretical dogmatism of strict construction, a rejection of almost all federal criminal law, goes too far.[77] Even if one acknowledges that the textual basis for federal criminal legislation is infirm, the federalist argument is academic. For all practical purposes, the question of whether there should be a federal criminal law anchored only by general constitutional mandates, such as the Commerce Clause, has been mooted by the public demand for, and the judicial and public acceptance of, sweeping federal criminal legislation.[78] Nevertheless, the position that federalism places no constraints on national legislation is also unacceptable.[79] An examination of the benefits of a federalist system leads to the conclusion that the application of the intangible rights doctrine to state and local political corruption diminishes those benefits and damages the federalist system.[80]

1. Why Federalism?

The Constitution mandates a tenuous, unfixed, and complex division of power between the federal and state governments that

---

[77] *See* Symposium, *supra* note 12, at 1744–47 (summarizing rejection by strict constructionists of federal criminal law).

In their consideration of the constitutional text and the Framers' original intent, strict constructionists largely ignore the implications of the post-Civil War amendments. The amendments, of course, profoundly altered the balance of power between state and federal governments. *See* New York v. United States, 112 S. Ct. 2408, 2444–45 n.3 (1992) (White, J., concurring in part and dissenting in part) (critiquing majority's formalistic and rigid obeisance to federalism); *see also* Akhil R. Amar, *Of Sovereignty and Federalism*, 96 YALE L.J. 1425, 1429–66 (1987) (recounting federalist debate from pre-Revolution through post-Civil War).

[78] *See, e.g.*, Anti-Drug Abuse Act of 1988, Pub. L. No. 100-690, 102 Stat. 4181; The Racketeer Influenced and Corrupt Organizations Act, 18 U.S.C. §§ 1961–1968 (1988 & Supp. IV 1992). The mail fraud statute would probably escape a strict constructionist's invalidation; the original mail fraud statute was passed pursuant to the constitutional provision establishing postal services and is justified under the Necessary and Proper Clause. *See* Symposium, *supra* note 12, at 1738.

[79] This position is the logical extension of the Court's decisions regarding congressional power to legislate in areas traditionally allocated to the states. *See* Garcia v. San Antonio Metro. Transit Auth., 469 U.S. 528 (1985) (repudiating Court's power to hold national statutes unconstitutional on federalism grounds); South Dakota v. Dole, 483 U.S. 203 (1987) (upholding legislation affecting drinking age on basis of spending power). *But see* Gregory v. Ashcroft, 111 S. Ct. 2395, 2398–401 (1991) (discussing federalist values); New York v. United States, 112 S. Ct. 2408 (1992) (rejecting federal government power to require state regulation as counter to federalist principles).

[80] *See* Symposium, *supra* note 12, at 1760–62 (concluding that use of mail fraud statute to pursue corruption in state government violates prerogative of states).

guarantees state, and in turn local, autonomy.[81] Beyond that elemental starting point, it is not easy to ascertain what federalism encompasses; the concept has been stripped of substantive content through its tendentious and selective use by politicians and agenda-driven theorists.[82] Federalism in practice encompasses more than a structural division of authority between units of government. It has meaning beyond a talismanic homage to the constitutional text and the theoretical principles of original intent that are used to justify political positions. Whether federalism is a result of adherence to theoretical Enlightenment principles or pragmatic eighteenth-century deal-making,[83] the Framers' division of the power to govern has bestowed significant and practical benefits.[84] Given the challenge to the federal structure presented by the negative centralizing effects of an increasingly powerful federal government, these benefits merit our attention and protection.

Shared authority between sovereigns protects and promotes individual freedom and control and, in turn, forges a connection among individuals and between individuals and their communi-

---

[81] This structure is achieved through distinct constitutional provisions: Article I of the Constitution, delegating legislative powers to the national Congress, is limited by the Tenth Amendment, which reserves to the states all powers not enumerated. *See* New York v. United States, 112 S. Ct. at 2417–18 (discussing federalist structure).

[82] *See, e.g.*, WORKING GROUP ON FEDERALISM, U.S. DOMESTIC POLICY COUNCIL, THE STATUS OF FEDERALISM IN AMERICA (1986) (arguing for greater attention to traditional division of power between federal and state governments in areas of federal regulation, programs, and grants, but ignoring threat to federalism of federal policing of local politics); Symposium, *supra* note 12, at 1739 (noting irony that administration committed to federalism approved mail fraud amendment that granted federal government authority to prosecute political corruption at state and local level); Michael W. McConnell, *Federalism: Evaluating the Founders' Design*, 54 U. CHI. L. REV. 1484, 1488 (1987) (reviewing RAOUL BERGER, FEDERALISM: THE FOUNDERS' DESIGN (1987)) (noting inconsistency of individual justices who, having dissented in *Garcia* on the ground that the Constitution protected state autonomy, voted in *Dole* to diminish that autonomy in light of congressional spending power).

[83] *See* Deborah Jones Merritt, *The Guarantee Clause and State Autonomy: Federalism for a Third Century*, 88 COLUM. L. REV. 1, 3 (1988) (noting federalism "was born as a political compromise rather than as a theoretical ideal").

[84] *See* McConnell, *supra* note 82, at 1493–511.

McConnell examines three major objectives of federalism envisaged by the Framers and concludes that they have been corroborated by modern public choice theory. Those objectives are: (1) preservation of decentralized decision-making, because small governmental units are more responsible to the interest and welfare of the people; (2) better protection of private rights and liberty; and (3) encouragement of a republican form of government. *Id.*

One might also add the benefits that federalism (1) checks "the oppressive power of a strong central government"; (2) increases opportunities for citizen participation in the democratic process; (3) creates diverse cultural and political environments; and (4) allows the states to operate as laboratories for experimental programs. *See* Merritt, *supra* note 83, at 3–10.

ties.[85] Decentralization encourages participation in the political process;[86] this leads to civic interaction between individuals and connects each individual to the larger community.[87] Through involvement in local legislation, individuals exercise control over their lives and the lives of their communities. People are more likely to abide by laws that they have initiated, and a vested citizenry indirectly enforces laws by demanding, encouraging, and reinforcing obedience to them.[88] An empowered citizenry, although occasionally an inconvenience to government administrators and elected leaders, ultimately benefits all levels of government as well as the general population because an engaged electorate is the source of innovation.[89] The federalist system creates "an all-pervading and restless activity, a superabundant force, and an energy which is inseparable from it, and which may, however unfavorable circumstances may be, produce wonders."[90]

The federal government acts in its own ultimate self-interest when it defers to and respects state and local autonomy. Although centralized government can be administratively efficient, the combination of its size and the unavoidably intense concentration of its authority tends to disempower citizens by reducing their role to that of "passive spectator" and "dependent actor."[91] Diffused power, in contrast, facilitates participatory democracy; this, in turn, can cause national, as well as state and local,

---

[85] These ideas are not new. *See* ALEXIS DE TOCQUEVILLE, DEMOCRACY IN AMERICA (Henry Reeve trans. & Francis Bowen trans. & ed., Cambridge, Sever & Francis 2d ed. 1863); *see generally* JOHN STUART MILL, *Considerations on Representative Government* (1861), and *On Liberty* (1859), *in* ON LIBERTY; REPRESENTATIVE GOVERNMENT; THE SUBJECTION OF WOMEN: THREE ESSAYS (Oxford Univ. Press 1912); THE FEDERALIST No. 10 (James Madison); KARL MARX, THE CIVIL WAR IN FRANCE (1871).

The work of Gerald Frug presents a modern view of the relation between state, local, and federal governments. *See* Gerald E. Frug, *The Ideology of Bureaucracy in American Law*, 97 HARV. L. REV. 1276 (1984); Gerald E. Frug, *The City as a Legal Concept*, 93 HARV. L. REV. 1057 (1980); Joan C. Williams, *The City, The Hope of Democracy: The Casebook as Moral Act*, 103 HARV. L. REV. 1174 (1990) (reviewing GERALD E. FRUG, LOCAL GOVERNMENT LAW (1988)).

[86] *See* Merritt, *supra* note 83, at 7–8 (noting entrance of women and minorities into state and local government positions).

[87] *See generally* DE TOCQUEVILLE, *supra* note 85; *see also* McConnell, *supra* note 82, at 1510 (noting that citizens who are actively engaged in public debate have a stake in the community).

[88] *See* McConnell, *supra* note 82, at 1508 (noting Framers' recognition of importance of voluntary compliance with the law, especially given a government that has minimal coercive apparatus).

[89] *See* Merritt, *supra* note 83, at 9 (listing national programs that began as experiments at state level).

[90] 1 DE TOCQUEVILLE, *supra* note 85, at 321.

[91] *Id.* at 113–14.

governments to respond more vigorously to the electorate. Participatory democracy promotes a strong electorate that ultimately strengthens and sustains the centralized bureaucracy.[92] Finally, decentralization through state and local autonomy provides a salutary counterweight to the centralization of government services and administrative regulations within the federal bureaucracy.

2. Federal Intervention Harms State and Local Autonomy

The Framers rejected a strong central authority for reasons historically grounded in a distrust of central government and a concomitant desire to hold rulers accountable.[93] The accountability of elected officials to the electorate, which is crucial to the state's role as a political decision-making unit in its own right,[94] is impaired when federal prosecutors monitor state and local officials.[95] This impairment is a result of the phenomenon that federal prosecutions for political corruption make state and local officials more accountable to an extrinsic entity, the federal government, than to those who voted for them. An interventionist federal presence encourages citizens to abdicate their responsibility for self-government at the state and local levels.[96] The ultimate result is a diminished demand on state and local legislative and executive branches to control political corruption.[97] The power federal prosecutors exercise over the political affairs

---

[92] *Id.* (concluding that when the central authority monopolizes power, it ultimately places itself at risk).

[93] *See* McConnell, *supra* note 82, at 1507–11 (noting Framers' concern for accountability and preservation of "Popular Government").

[94] *See* New York v. United States, 112 S. Ct. 2408, 2424 (1992) (invoking political accountability in context of federal legislation directing state to regulate); D. Bruce La Pierre, *Political Accountability in the National Political Process—The Alternative to Judicial Review of Federalism Issues*, 80 NW. L. REV. 577, 635 (1985).

[95] *See* Baxter, *supra* note 40, at 337 n.80 (arguing that unrestrained federal criminal jurisdiction would destroy state responsibility and autonomy in preservation of public order and administration of criminal law, quoting *Hearings Before the Subcomm. on Criminal Law and Procedures of the Senate Comm. on the Judiciary*, 92d Cong., 1st Sess. 7 (1971)).

[96] *See* Miner, *supra* note 16, at 127 ("What we are witnessing is an abdication of responsibility for self-government . . . ."); Loomis, *supra* note 40, at 73–74 (concluding that monitoring performance of government officials is best left to citizens of the state because federal intervention dampens rapport between elected representatives and constituents).

[97] *See* Charles F.C. Ruff, *Federal Prosecution of Local Corruption: A Case Study in the Making of Law Enforcement*, 65 GEO. L.J. 1171, 1214 (1977) (noting also that even successful prosecutions may have minimal long-term effect on political climate if electorate did not demand the reform).

of states and cities is particularly troublesome in the area of criminal legislation, in which states traditionally have the "principal responsibility for defining and prosecuting crimes."[98]

Prosecutorial discretion makes the federal infringement on state and local power more acute. An effective enforcement system requires that prosecutors be generally free to decide whom to prosecute.[99] In the context of mail fraud, that discretion is enhanced because the statutory elements of the crime provide only attenuated constraints.[100] Internal guidelines and supervision, which would provide an effective limitation on the discretion of prosecutors to charge mail fraud, are either nonexistent or ineffectual.[101] Due in part to the decentralized organization of the Department of Justice, federal prosecutors often operate independently, with minimal direction or supervision from their superiors in Washington.[102] Further, the secrecy and informality

---

[98] United States v. Bass, 404 U.S. 336, 349 (1971) (requiring clear congressional statement of intent to alter sensitive relation between federal and state criminal jurisdiction); *see also* Abbatte v. United States, 359 U.S. 187, 195 (1959) (holding that states bear principal responsibility for defining and prosecuting crimes); Screws v. United States, 325 U.S. 91, 109 (1945) (stating that administration of criminal justice under federal system rests with the states); Miner, *supra* note 16, at 117 (noting that detection, apprehension, prosecution, and punishment of criminals has historically been considered state function of most basic kind).

[99] To the extent that prosecutorial discretion may result in increased numbers of prosecutions and convictions, and to the extent that by serving their self-interest prosecutors may also serve the community interest, prosecutorial discretion produces short-term benefits. Nevertheless, an inescapable tension arises in the course of balancing short-term effectiveness in controlling political corruption against long-term derogation of state and local autonomy.

To mitigate such derogation, one commentator has suggested that federal prosecutors pass the results of their political corruption investigations to state and local officials for prosecution. *See* Coffee, *supra* note 31, at 170–71. *But see* Hurson, *supra* note 13, at 453 n.249 (it is "difficult to imagine why anyone would wish to work in the Justice Department or a United States Attorney's office if he was merely to be a foot soldier for state prosecutors").

[100] *See* United States v. Margiotta, 688 F.2d 108, 143 (2d Cir. 1982) (Winter, J., dissenting) (finding troubling the potential for abuse because of the raw political power mail fraud statute affords federal prosecutors to prosecute selectively), *cert. denied*, 461 U.S. 913 (1983); *see also supra* notes 14–35 and accompanying text (discussing elements of crime of mail fraud).

[101] The Department of Justice has not formulated meaningful guidelines for mail fraud prosecutions, as it has for RICO prosecutions, relying instead on the judgment of the individual attorney in the field. *See* Maass, *supra* note 40, at 119–22 (stating that uncontrolled discretion of prosecutors was due to nature of Department of Justice Public Integrity Section, characteristics of the office of United States Attorney, and deference to prosecutorial independence); Williams, *supra* note 6, at 146 (noting that Department of Justice *United States Attorneys Manual* guideline for initiating mail fraud indictments—"if the scheme is in its nature directed to defrauding . . . the general public"—does not provide meaningful direction to prosecutors).

[102] *See* Maass, *supra* note 40, at 121 (Department of Justice criminal division provides advice to United States Attorneys when it is requested); Ruff, *supra* note 97, at 1201–05 (explaining decentralized operation of Department of Justice in historical and pragmatic

with which prosecutors decide to press charges shield them from both outside scrutiny and internal examination.[103] Finally, courts defer to prosecutors' decisions in the belief that to do otherwise imposes judicial control over the Executive Branch in violation of the doctrine of the separation of powers.[104] When combined with the intangible rights doctrine, broad prosecutorial discretion means that prosecutors decide whom to charge, and more significantly for our purposes, what behavior is defined as criminal.[105]

Prosecutors do not apply the standards of an articulated federal policy to the behavior of state and local officials. Because of the manner in which the intangible rights doctrine evolved, no such policy exists.[106] Instead, prosecutors impose their personal and subjective standards of good government on those subject to their jurisdiction.[107] Understandably, there is disagreement over the nature of those standards of public morality, and the prosecutor's standards may or may not be those of the electorate. The result is inadequate notice of criminality.[108]

The absence of standards has thus far not been a tenable defense to charges of mail fraud. Courts have not had a vehicle

---

terms and noting that Department of Justice procedures do not enable it to control manner in which individual prosecutors exercise their prosecutorial discretion).

[103] *See* HERBERT L. PACKER, THE LIMITS OF THE CRIMINAL SANCTION 89 (1968) (noting that prosecutors "are not subjected to even the minimal psychological constraints that flow from self-perceived deviation from an acknowledged role"). This phenomenon is illustrated by the shock with which a federal prosecutor received the news that it was not his job to "pursue corruption" until Congress had enacted a statute outlawing it. *See* Symposium, *supra* note 12, at 1761.

[104] *See* Maass, *supra* note 40, at 122 (explaining excessive judicial deference to prosecutorial discretion). *But see* United States v. Archer, 486 F.2d 670, 678 (2d Cir. 1973) (federal courts have power to correct abuses of discretion where initial investigation was based on a "grossly inflated conception of the role of federal criminal law"). The irony, of course, is that unbridled discretion effectively allows the prosecutor to invade the province of the lawmaker. *See infra* text accompanying notes 111–117.

[105] *See* Miner, *supra* note 16, at 126–27 (characterizing prosecutorial discretion under the intangible rights doctrine as "unbridled"); *see also* Rakoff, *supra* note 8, at 779 (unique design and interpretation of mail fraud statute renders it more liable to "irrational, unpredictable or extreme applications and hence, to abuse").

[106] *See* Maass, *supra* note 40, at 111–19 (stating that mail fraud prosecutions are uncontrolled, unsupervised, and unauthorized by any legislative body); *supra* notes 36–53 and accompanying text (tracing development of intangible rights doctrine).

[107] *See* Maass, *supra* note 40, at 119–24 (criticizing "selective, nonuniform, and uncertain use of criminal law to foster the morality of individual federal prosecutors"); Coffee, *supra* note 31, at 144 (questioning desirability of expansion of intangible rights in adversarial political setting where winners have incentive to indict losers).

[108] *See* United States v. Siegel, 717 F.2d 9, 24 (2d Cir. 1983) (Winter, J., dissenting) (noting the impossibility of predicting what "acts may after the fact attract a prosecutor's suspicions (or ire) and a judicial stamp of impropriety"); *see infra* text accompanying notes 174–194 (discussing notice prong of vagueness test).

by which to consider such a defense because there is no statutory standard of good government to measure, no administrative program to evaluate, and no specific political corruption law to consider.[109] Courts routinely reject defenses based on abuse of prosecutorial discretion, the *cri de coeur* of criminal defendants. As we have seen, the sources of that discretion are diffuse, its authority is questionable, and its impact is far from concrete. More significantly, prosecutorial discretion is based on constitutional considerations of its own. Finally, courts also routinely reject defenses grounded on the debasement of federalism.[110]

B. *The Intangible Rights Doctrine and the Separation of Powers*

The intangible rights doctrine, combined with prosecutorial discretion, implicates the Constitution's structural requirement of separation of powers.[111] Legislatively enacted criminal codes define criminal conduct by specifying formal elements of the crime. The judicial function is to enable a jury to compare the conduct of an accused to the elements of the crime. When the

---

[109] *See* Maass, *supra* note 40, at 113 (noting lack of authorizing statute for Department of Justice Public Integrity Section).

Individual justices have occasionally raised the issue of prosecutorial discretion. *See* Evans v. United States, 112 S. Ct. 1881, 1904 (1992) (Thomas, J., dissenting); Morrison v. Olson, 487 U.S. 654, 727–32 (1988) (Scalia, J., dissenting).

[110] For instance, the Fourth Circuit summarily rejected Mandel's argument that his conviction constituted an impermissible federal intrusion into the political affairs of the state. *See* United States v. Mandel, 591 F.2d 1347, 1359 (4th Cir.), *conviction aff'd per curiam in relevant part on reh'g en banc*, 602 F.2d 653 (4th Cir. 1979), *cert. denied*, 445 U.S. 961 (1980). The court invoked the purpose of the statute—to prevent the post office from being used to carry out fraudulent schemes—to find that even though the alleged scheme involved traditional state concerns, prosecution for mail fraud was not precluded. *Id.* The unarticulated conclusion is that the federal interest in the mail system trumps the state interest in regulating criminal conduct. *See also* United States v. States, 488 F.2d 761, 766–67 (8th Cir. 1973) (invoking mail fraud's purpose to reject argument that Congress did not explicitly authorize intervention into state affairs), *cert. denied*, 417 U.S. 909 (1974).

The Supreme Court has only rarely invalidated a federal statute because it offended federalist principles. *See* Oregon v. Mitchell, 400 U.S. 112, 117–18 (1970) (invalidating federal statute lowering voting age in state and local elections, which was superseded by passage of Twenty-Sixth Amendment); Carter v. Carter Coal Co., 298 U.S. 238, 290–97 (1936) (invalidating sections of New Deal legislation and discussing federalist principles).

[111] *See* Baxter, *supra* note 40, at 334 (federal prosecutors' unlimited discretion to define the meaning of a statute is "incompatible with constitutional notion of separation of powers"); *see also* Ruff, *supra* note 97, at 1228 (generally approving of prosecutor's role in formulating law enforcement policy but cautioning that prosecutor "must take care he does not assume the legislative role").

elements of a crime are unfixed and unclear, however, the judge and jury fix the elements, usurping the role properly left to the legislature and violating the separation of powers. An unclear criminal statute forces the judge and jury to derive the elements of the crime from the factual circumstances of the case.[112] The judge and jury, in formulating and in utilizing jury instructions, transform those factual circumstances into the "elements" of the crime.[113]

This fact-specific innovation of new crimes invites further innovation on new facts, and results in an open-ended, progressive creation of crimes. Indeed, the incremental progression of precedents that culminated in the application of the intangible rights doctrine to elected and nonelected officials is an excellent example of judicial crime creation.[114] Prosecutors and judges, by repeatedly using the facts of a specific case to fill the interstices left by the vague mail fraud statute, effectively exercise lawmaking powers.[115] The process not only blurs the distinction between legislative, executive, and judicial functions, but also allows the executive and judicial branches to usurp legislative functions. Judicial crime creation invites and encourages prosecutors to bring previously undefined conduct to trial in the hope that the court will criminalize it.[116] When prosecutors select cases according to their own agendas of social control, bureaucratic necessity, or personal advancement, lawmaking devolves to law enforcers.[117]

---

[112] A judicial definition of mail fraud that is based in effect on the factual circumstances of a particular defendant's specific behavior may also implicate the Ex Post Facto Clause. The Ex Post Facto Clause provides: "No Bill of Attainder or ex post facto Law shall be passed." U.S. CONST. art. I, § 9, cl. 3.

[113] *See* Morano, *supra* note 13, at 47 n.3 (analogizing mail fraud statute to the bed that Procrustes ensured would fit his guests by either stretching or cutting off the legs of his guests).

[114] *See* United States v. Margiotta, 688 F.2d 108, 141–42 (2d Cir. 1982) (Winter, J., dissenting) (tracing judicial development of intangible rights doctrine), *cert. denied*, 461 U.S. 913 (1983).

[115] *See, e.g.*, United States v. Maze, 414 U.S. 395, 405–08 (1974) (Burger, C.J., dissenting) (arguing that mail fraud should be used as a stopgap against new forms of fraud until Congress passes appropriate legislation). *But see* United States v. Holzer, 816 F.2d 304, 309 (7th Cir.) (stating that broad definition of fraud to include whatever is not morally upright "would put federal judges in the business of creating . . . common law crimes, i.e., crimes not defined by statute"), *vacated*, 484 U.S. 807 (1987).

[116] *See* Jeffries, *supra* note 53, at 222–23 (warning that the danger of fact-specific innovation is that police and prosecutors play too large a role in deciding what conduct is punished). Prosecutors, whose success is at least partially defined by convictions, naturally exploit the uncertainty created by fact-specific judicial innovation. *See supra* text accompanying notes 99–108 (discussing prosecutorial discretion).

[117] *See* Miner, *supra* note 16, at 127 ("Where there are many criminal violations but

C. *The Intangible Rights Doctrine and the First Amendment*

The intangible rights doctrine, coupled with prosecutorial discretion, diminishes more than the abstract concepts of federalism and separation of powers. The mail fraud statute chills protected First Amendment political activity because it allows the selection of certain individuals for prosecution.[118] Selective mail fraud

---

only a few can be chosen for prosecution, the prosecutor invades the province of the lawmaker.").

Criminal law theorists also decry the impact of this ad hoc criminalization of conduct on the development of criminal law and on the criminal justice system. *See* Williams, *supra* note 6, at 157 (stating that use of judicial interpretation of mail fraud statute to develop federal law results in superficial treatment of a very important problem). Allowing federal prosecutors to define criminal behavior, and thus criminal policy, on an ad hoc basis ignores the appropriate legislative role in defining criminal conduct rationally and carefully. *See id.* (noting that state legislatures are proper arena for significant political judgments involved in developing political corruption legislation). In addition, judicial expansion of existing law may ultimately produce law that differs from that which the legislature would have adopted. *See* Coffee, *supra* note 13, at 10 (judicial expansion of a statute may alter and distort the direction of legislative reform; for example, evolution of mail fraud law is at odds with the Model Penal Code).

[118] *See* Williams, *supra* note 6, at 140 n.18 (mail fraud is "one of the most selectively enforced statutes in the history of the federal criminal law"). Because the mail fraud statute is often enforced arbitrarily and selectively, it may well violate other individual civil rights not considered in this Article. These include due process, Grayned v. City of Rockford, 408 U.S. 104, 108–09 (1972) (holding that lack of explicit standards allows arbitrary enforcement and violates due process); Papachristou v. City of Jacksonville, 405 U.S. 156 (1972) (overturning city vagrancy statute in part because selective enforcement violates due process); substantive due process, because of a confusion between immorality and illegality, *see* Morano, *supra* note 13, at 76 (rejecting synonymy between morality and criminality); and equal protection, *see* Yick Wo v. Hopkins, 118 U.S. 356, 373–74 (1896) (a fair law applied "with an evil eye and an uneven hand" denies equal justice and violates Constitution); *see also* Mark L. Amsterdam, *The One-Sided Sword: Selective Prosecution in Federal Courts*, 6 RUT.-CAM. L.J. 1, 25 (1974) (arguing that equal protection is violated if prosecution displays purposeful discrimination and an arbitrary or unjustifiable standard of selection).

Excessive prosecutorial discretion may also result in diminished horizontal equality between offenders as prosecutors treat similar conduct differently, depending on the prosecutor's particular charging policy. *See* Baxter, *supra* note 40, at 343–44 (examining reasons for dissimilar treatment of similar offenders); Ruff, *supra* note 97, at 1203 (noting that basis of decision to prosecute varies from district to district).

The prosecution's policies may vary on several fronts. First, the prosecutor has power to indict defendants for more than one offense because the jurisdictional elements of mail fraud mirror those of other federal statutes. *See* United States v. Computer Sciences Corp., 689 F.2d 1181, 1187 (4th Cir. 1982) (upholding prosecutions under RICO, False Claims Act, mail fraud, and wire fraud because statutes do not require that "prosecution under one must be at the expense of prosecuting under the other"), *cert. denied*, 459 U.S. 1105 (1983), *overruled on other grounds by* Busby v. Crown Supply, Inc., 896 F.2d 833, 841 (4th Cir. 1990); United States v. Gillock, 587 F.2d 284 (6th Cir. 1978), *rev'd on other grounds*, 445 U.S. 360 (1980) (example of prosecution under Hobbs Act, Travel Act, and RICO); Morano, *supra* note 13, at 48 n.6 (listing cases in which multiple charges were based on single course of conduct). The prosecutor can also bring state charges in addition to federal charges. *See* Baxter, *supra* note 40, at 344. Finally, the prosecutor decides how many counts of mail fraud to charge. *See supra* text accompanying note 35.

prosecution may result from personal motives of political aspirations,[119] simple careerist concerns,[120] or a desire to eliminate political opponents or activists.[121] This potential misuse of the mail fraud statute creates a danger to the political process that may eclipse the danger posed by corrupt officials.[122] The threat of criminal sanctions may discourage people from participating in the political process at all.[123]

---

[119] *See* Loomis, *supra* note 40, at 78–79; ROSCOE POUND, CRIMINAL JUSTICE IN AMERICA 182–89 (DeCapo Press 1972) (1930) (concluding that the political ambition of a public prosecutor is an inherent and negative component of the American criminal justice system).

Successful prosecutions of public officials create sufficient capital to launch a political career. Two examples will suffice to make the point. James Thompson, a Republican and former United States Attorney for the Northern District of Illinois, successfully prosecuted many political associates of Chicago's Democratic mayor Richard Daley, including former governor Otto Kerner, former state revenue director Theodore Issacs, former city council leader Thomas Keane, and Daley adviser Earl Bush. Subsequently, Thompson was elected governor of Illinois, and was later considered for United States Attorney General. *See* Kurland, *supra* note 52, at 398; Ann Devroy, *Speculation At Justice on Thornburgh*, WASH. POST, May 2, 1991, at 17.

Former Attorney General Richard Thornburgh's career is also based on successful political corruption prosecutions. As United States Attorney for the Western District of Pennsylvania, he successfully prosecuted Pittsburgh area mayors and city council members from 1969–75 and was elected District Attorney of Allegheny County. In 1975, he headed the newly created Public Integrity Section of the Department of Justice. In 1978, he was elected governor of Pennsylvania. *See* Maass, *supra* note 40, at 125.

[120] *See* Ruff, *supra* note 97, at 1211 (stating that it is unrealistic to deny that prosecutor's motivation is occasionally colored by personal ambition); Coffee, *supra* note 13, at 21 (commenting on self-interest of prosecutor). A highly visible, successful prosecution can significantly advance a prosecutor's career and there are few internal constraints to counter the temptation to pursue such cases. *Id.*

[121] Some commentators have drawn a connection between the prosecutions of associates of Chicago Mayor Richard Daley and the pivotal role of Daley's machine in delivering the Illinois vote to John F. Kennedy in the presidential election of 1960. *See* HANK MESSICK, THE POLITICS OF PROSECUTION: JIM THOMPSON, MARJE EVERETT, RICHARD NIXON AND THE TRIAL OF OTTO KERNER (1978); Williams, *supra* note 6, at 148. *But see* Maass, *supra* note 40, at 124 (concluding that evidence to support charge that Republican Department of Justice focused on Democratic officeholders is unconvincing); Loomis, *supra* note 40, at 79 (noting that vigorous prosecution of state and local officials will always appear politically motivated, at least from viewpoint of defendant's supporters). Interestingly, Chicago continues to be a focus of federal prosecutions. *See* Kurland, *supra* note 52, at 398 n.111. The political nature of federal prosecutions is heightened by the fact that the President appoints United States Attorneys from candidates nominated by state and local officials from the President's political party. *See* Maass, *supra* note 40, at 120.

[122] *See* United States v. Margiotta, 688 F.2d 108, 144 (2d Cir. 1982) (Winter, J., dissenting) (stating that use of mail fraud creates "a danger of corruption to the democratic system greater than anything Margiotta is alleged to have done"), *cert. denied*, 461 U.S. 913 (1983); United States v. Archer, 486 F.2d 670 (2d Cir. 1973); Ingber v. Enzor, 664 F. Supp. 814 (S.D.N.Y. 1987).

[123] *See Margiotta*, 688 F.2d at 143 (Winter, J., dissenting) (arguing that decision upholding conviction of unelected official for mail fraud will discourage political participation); Coffee, *supra* note 31, at 143–44 (concluding that vague standard of *Margiotta* may chill willingness of individuals to participate in politics because it

These observations should not be dismissed as an alarmist assertion or an overdrawn parade of horribles. The judiciary has expanded the intangible right to honest government to include nonelected officials[124] and has made failure to disclose a political motive the equivalent of misrepresenting a material fact.[125] These developments may inhibit participation in the political process because they expose politicians to criminal sanctions for what they do not say.[126] Politicians and political activists could conceivably be indicted for conduct traditionally considered to be partisan campaigning, interest group building, or furthering the interests of their party.[127] Arguably, the freedoms of speech and political association are being chilled, and thereby violated.[128]

In summary, federal mail fraud prosecutions of state and local officials present three major constitutional concerns. First, state and local decision-making power, the basis of a viable federalist system, is diminished by federal prosecutions of state and local

---

criminalizes "cronyism"); Hurson, *supra* note 13, at 441 (stating that theory expressed in *Margiotta* sweeps protected First Amendment political activity into crime of mail fraud); Jeffries, *supra* note 53, at 239 (concluding that result of *Margiotta* is "ill-defined prospect of criminal liability for influential private citizens whose participation in the political process falls short of civics-book standards").

[124] *See Margiotta*, 688 F.2d 108 (upholding conviction of non-elected party official who had directed insurance commissions to political associates); *see also supra* notes 48–50 and accompanying text (discussing *Margiotta*).

There is some justification for the proposition that an elected official owes a fiduciary duty to the electorate as a whole; a fiduciary duty, however, should not be inferred from the status and power of a nonelected politician. *See Margiotta*, 688 F.2d at 141–42 (Winter, J., dissenting) (rejecting view that nonelected political participant in a pluralistic, partisan political system owes a fiduciary duty to electorate as a whole).

[125] *See* United States v. Mandel, 591 F.2d 1347 (4th Cir.), *conviction aff'd per curiam in relevant part on reh'g en banc*, 602 F.2d 653 (4th Cir. 1979), *cert. denied*, 445 U.S. 961 (1980); *Margiotta*, 688 F.2d 108; *supra* text accompanying notes 36–53 (discussing development of intangible rights doctrine).

[126] *See Margiotta*, 688 F.2d at 140 (Winter, J., dissenting) (noting that majority's logic encompasses campaign literature and public speeches); Coffee, *supra* note 13, at 14–15 (stating that *Margiotta* could be read to require officials to disclose ideological, as well as economic, conflicts of interest). Judge Winter's examples of conduct giving rise to mail fraud include a candidate who mails a brochure containing a promise that the candidate knows cannot be kept; an official who performs an act imposing unnecessary costs on taxpayers without disclosure; and a partisan political leader who throws decisive support behind a less qualified candidate, similarly without disclosure. *See Margiotta*, 688 F.2d at 140 (Winter, J., dissenting).

[127] *See* Coffee, *supra* note 31, at 144 (commenting that *Margiotta* ruling may interfere with coalition formation, which demands reciprocity); Hurson, *supra* note 13, at 441–42 (noting that the role of party activists is to advance the position of the party and its members).

[128] *See* Grayned v. City of Rockford, 408 U.S. 104, 114–15 (1972) (political participation may not be deterred); Citizens Against Rent Control/Coalition for Fair Housing v. Berkeley, 454 U.S. 290, 299 (1981) (freedom of political association may not be deterred); United Mine Workers of America v. Pennington, 381 U.S. 657, 670 (1965) (efforts to influence public officials may not be deterred).

officials. Prosecutions by the central authority weaken the ties between elected officials and the electorate, impose the individual moral standards of prosecutors on the electorate, and may skew state and local politics when prosecutors target political enemies or prosecute to advance personal ambition. Second, the use of the intangible rights doctrine has grave implications for the separation of powers because it encourages judicial crime creation and diminishes the authority of the legislature. Third, the doctrine intrudes upon traditional political activity and compromises individual First Amendment rights.

D. *An Evaluation of the Federal Interest*

The use of the intangible rights doctrine has, of course, positive results. Successful prosecutions take corrupt politicians "off the streets," eliminating some quantum of political corruption,[129] and the resultant publicity probably has some deterrence value. Prosecutions for political corruption are generally supported by the public, who perceive the underlying conduct as immoral.[130] Nevertheless, public support does not mitigate the specific negative consequences of the use of the intangible rights doctrine. Furthermore, reliance on the number of successful prosecutions and on public support begs the question of whether an independent federal interest exists to justify federal intervention into state and local politics.[131] Most justifications for these federal

---

[129] *See* Thomas H. Henderson, Jr., *The Expanding Role of Federal Prosecutors in Combatting State and Local Political Corruption*, 8 CUMB. L. REV. 385, 401 (1977) (praising expansion of intangible rights doctrine).

[130] *See generally* Philip B. Heymann, *The Risks of Corruption*, THE PUBLIC INTEREST, Fall 1987, at 128 (decision to prosecute is based on a general sense of moral outrage at the underlying conduct and the strength of the evidence supporting the charge).

[131] *See* L.B. Schwartz, *Federal Criminal Jurisdiction and Prosecutors' Discretion*, 13 LAW & CONTEMP. PROBS. 65, 73 (1948) (presenting specific federal interests that justify intervention into state and local affairs); Ruff, *supra* note 97, at 1213 (federal intervention in local corruption cases should be founded on balance of three factors: (1) the capacity of the federal prosecutor to manage both the investigation and the prosecution effectively; (2) the capacity of the local prosecutor, with the help of the federal prosecutor, to manage such cases; and (3) the adverse social consequences that would result from a failure to prosecute); Maass, *supra* note 40, at 114 (allowing federal prosecutions for (1) violations of basic constitutional rights of citizens; (2) crimes that infringe on primary federal interests; and (3) criminal behavior for which any citizen would be prosecuted).

Commentators have also argued the converse—that there are federal interests, judicial economy and efficiency, to support limiting federal criminal jurisdiction. HENRY J. FRIENDLY, FEDERAL JURISDICTION 55–73 (1973) (proposing minimal federal criminal jurisdiction on ground that most federal criminal legislation has "greatly outreached any true federal interest").

prosecutions are based on mere expediency and do not constitute a federal interest.

One significant federal interest justifying intervention is the obligation of the federal government to ensure that local and state governments conduct their affairs in accordance with constitutional principles.[132] This consideration, however, must be tempered by the distinction between federal intervention to enforce constitutional rights and federal intervention to punish political corruption. The federal government certainly intervened in state politics when it sent federal troops to Montgomery, Alabama on behalf of students who sought to enter a local, segregated high school. Intervention was justified because students have a constitutional right to an integrated classroom. The Constitution, however, does not grant citizens a right to good government. Rather, the Constitution, by providing a governmental structure, safeguarding civil liberties, and guaranteeing democratic elections, makes good government possible. The electorate itself must effectuate that promise. Only when supervening events that have been clearly and statutorily articulated as unlawful interfere with the ability of the electorate to effectuate good government should federal authorities intervene.[133] Examples of situations that justify federal interference include cases in which corruption involves the election process, infiltrates law enforcement agencies, or involves multi-state investigations. Ironically, because mail fraud reaches only those schemes that involve a mailing, the statute will not always reach the criminal conduct that most justifies its use.

A sophisticated justification for the federal role in state and local political corruption cases has been derived from the constitutional clause guaranteeing a republican form of government.[134] The same provision, however, is also invoked to support

---

[132] *See* Williams, *supra* note 6, at 155–56 (discussing reasons why federal prosecutors have power to prosecute local instances of corruption).

[133] *See* Schwartz, *supra* note 131, at 73–75 (discussing justifications for federal criminal law).

[134] *See* Kurland, *supra* note 52, at 435 (concluding that Guarantee Clause is the source of federal authority to prosecute cases of state and local official corruption under the intangible rights doctrine); Symposium, *supra* note 12, at 1769–73 (presenting clause as justification for federal prosecution of state and local political corruption and use of mail fraud statute following *McNally*).

The Guarantee Clause states: "The United States shall guarantee to every State in this Union a Republican Form of Government, and shall protect each of them against Invasion; and on Application of the Legislature, or of the Executive (when the Legislature cannot be convened) against domestic Violence." U.S. CONST. art. IV, § 4.

the converse concept, state autonomy.[135] Although this theory has generated some interest,[136] it does not necessarily support federal intervention based on the amorphous intangible rights doctrine: there remains the significant task of defining "republican." A strong argument can be made that the citizens who live under a state's republican form of government, not members of the federal executive and judicial branches, should determine whether certain conduct is destructive of their form of republicanism.

Other advocates of the federal role assert that the continued viability of the republic rests on the faith people have in government institutions at all levels, and that the public perception of the integrity of those institutions is of paramount importance.[137] Accordingly, anything that damages that perception constitutes a federal interest.[138] This contention is misguided in its emphasis on perception over reality, and is likely to cause problems more significant than the corruption it seeks to stop. Its logic justifies the manipulation of reality in order to engender a positive perception, no matter what is really going on. In the long run, such a policy is likely to create cynicism, disenchantment with politics, and eventually alienation from the political process.[139]

Federal prosecution of state and local political corruption is more successfully defended on pragmatic grounds. It is argued that local prosecutors, whether co-opted by the corrupt power structure or personally involved in it, will not lodge charges against popular or prominent politicians,[140] and that even when local prosecutors are conscientious and honest, human nature

---

[135] *See* Merritt, *supra* note 83, at 2 (arguing that the clause implies a restraint on federal power because states cannot enjoy republican governments unless they retain sufficient autonomy to establish and maintain their own form of republicanism).

[136] *See* Symposium, *supra* note 12, at 1766 (recounting debate within Department of Justice on ways to counter *McNally*); *id.* at 1769–73 (presenting argument in favor of using Guarantee Clause).

[137] *See* Carey et al., *supra* note 6, at 313–14 (positing that political corruption impairs belief that the people can govern themselves and thus strikes at fundamental tenet of community life).

[138] Kurland, *supra* note 52, at 377 (describing Department of Justice statement designating public corruption as a criminal enforcement priority, 3 *U.S. Department of Justice, Criminal Division, Law Enforcement Alert*, Jan. 1985, at 1, to the effect that any action that erodes faith of citizenry creates a substantial federal interest).

[139] *See generally* Richard D. Parker, *The Past of Constitutional Theory—And Its Future*, 42 OHIO ST. L.J. 223 (1981) (critiquing assumptions of process theory).

[140] *See* Carey et al., *supra* note 6, at 307–09 (detailing practical impediments to prosecuting one's superiors and colleagues); Kurland, *supra* note 52, at 377 (noting political difficulty of acting against entrenched power). *But see* Maass, *supra* note 40, at 116 (rejecting this justification as no substitute for statutory authority).

and political reality make such prosecutions unlikely. There is some evidence, however, that local prosecutors do take action.[141] An alternate argument suggests that local prosecutors would act enthusiastically if the federal presence was eliminated because they can fulfill individual political ambitions by prosecuting corrupt officials.[142] Successful federal prosecutions discourage the efforts of local enforcement officials,[143] resulting in fewer state prosecutions and a justification of more federal intervention.

Another expedient reason for federal action is the frequent absence of state statutes authorizing such prosecutions by state and local prosecutors.[144] The absence of state statutes may simply indicate that the electorate—theoretically harmed by the behavior at issue—is not disturbed enough by such activity to support legislation outlawing it. When a state or local entity has declined to enact legislation addressing certain kinds of political corruption, the wisdom and propriety of federal intervention is suspect.[145] In addition, this argument exposes the fiction that the use of mail fraud to pursue state and local political corruption is auxiliary to state law.[146]

It is argued that federal prosecution is more efficient and productive than state prosecution because the federal prosecutor

---

[141] *See* Maass, *supra* note 40, at 117 (presenting several instances during 1986 in which New York prosecutors who were actively investigating corrupt officials were preempted by federal investigators); Williams, *supra* note 6, at 155 n.126 (citing Mingo County, West Virginia prosecution of 15 public officials on political corruption charges, State *ex rel.* Owens v. Brown, 351 S.E.2d 412 (W. Va. 1986)); *see, e.g.*, Sharon Walsh, *Justice Content to Take Back Seat in the BCCI Case*, WASH. POST, Apr. 9, 1993, at F1.

[142] *See* Williams, *supra* note 6, at 155 (state prosecutors, as elected officials, know that failure to prosecute political corruption raises questions about their own integrity); *supra* note 119 (providing examples of correlation between prosecutions and successful political careers).

[143] *See* Schwartz, *supra* note 131, at 70.

[144] *See* United States v. Bush, 522 F.2d 641 (7th Cir. 1975), *cert. denied*, 424 U.S. 977 (1976) (example of prosecution in which there was "no concrete evidence of . . . a violation of any [state] criminal law"); Kurland, *supra* note 52, at 377 (noting that until recently state anti-corruption statutes were ineffective or nonexistent).

In addition, other barriers may hamper the efforts of state and local prosecutors. *See* Carey et al., *supra* note 6, at 304–09 (listing structural barriers to state and local enforcement in West Virginia, including absence of statewide prosecuting authority, part-time nature of county prosecutor's position, jurisdictional boundary of county line, constraints on grand jury, and procedural limitations at trial).

[145] *See* Loomis, *supra* note 40, at 75 (in the absence of state legislation, federal jurisdiction "actually is predicated upon a basic difference of opinion between the federal courts and the state legislature as to what local activity should be criminalized"); *see also supra* text at notes 93–98 for discussion of paramount interest of state and local governments in criminal law and the relation of that interest to federalism.

[146] *See* Coffee, *supra* note 31, at 169 (noting that federal effort to enforce intangible rights cannot be auxiliary to state criminal law enforcement when state law does not criminalize the conduct at issue).

has more effective procedural tools at the investigative stage to gather evidence, protect witnesses, and procure wiretaps.[147] If efficiency were the sole criterion for federal intervention, however, the federal government would properly usurp many traditional state functions.[148] The Framers rejected efficient, centralized authority when they devised the federal system, and the nation has endorsed that choice. Finally, the resources of the federal authority are not unlimited, and complex mail fraud investigations divert those resources from the increasing amount of more traditional fraudulent activity, especially in the federal regulatory context.[149]

Such justifications of federal intervention as the failure of state and local officials to act, the relative efficiency of federal law enforcement, and the putative federal interest supporting federal prosecutions wrongly focus on the short-term solution of ridding state and local politics of corrupt practices. This myopic concern for the short term ignores the possibility that tolerating political corruption until those governed reject it may best serve society's long-term interests. Moreover, the justifications fail to counter the specific constitutional problems raised by the intangible rights doctrine.

III. The Mail Fraud Statute Is Facially Void for Vagueness

The mail fraud statute now explicitly prohibits schemes to defraud citizens of the intangible right to the "honest services"

---

[147] *See* Kurland, *supra* note 52, at 380 (citing authority of federal prosecutor to offer limited use immunity and to utilize non-consensual wiretaps); *Mail Fraud Hearings*, *supra* note 6, at 16–17 (1988) (statement of John C. Keeney, Acting Assistant Attorney General) (explaining that state and local prosecutors face limited investigative authority, geographical confinement, ineffective grand juries, and general lack of resources and personnel); Williams, *supra* note 6, at 156 (noting vast police and prosecutorial resources of federal government).

[148] *See* Baxter, *supra* note 40, at 341 (concluding that despite advantage of highly specialized investigatory techniques, federal interest amounting to more than efficiency must exist).

[149] Williams, *supra* note 6, at 156 (commenting, in the face of increased demands, that resource advantage of federal prosecutors may be more theoretical than real); Steve Lohr, *A New Breed of Sam Spade on the Trail of Hidden Assets*, N.Y. Times, Feb. 20, 1992, at A1 (noting increasing reliance of federal regulatory agencies on private investigative firms to locate hidden assets of white collar criminals); *supra* note 99 (discussing proposal that federal prosecutors share files with state counterparts); Coffee, *supra* note 13, at 19–20 n.84 (arguing that current practice is an inefficient use of federal resources because it diffuses prosecutorial energies over a wide spectrum of behavior and often results in inconsequential cases).

of public officials. The use of the term "honest services," added in a 1988 amendment to the mail fraud statute,[150] does not, however, meet the Supreme Court's requirement that criminal statutes should be clear and unambiguous, a requirement that is consistent with the boundaries of federal power.

The statute is unconstitutionally vague because it fails to provide notice to state and local officials of what conduct is prohibited and, more crucially in the context of political corruption, fails to provide standards of enforcement to federal prosecutors. In addition, the statute is facially vague because it implicates significant constitutional provisions that cannot be protected on a case-by-case basis.

In the following discussion, I examine the theoretical foundations of the vagueness doctrine and the bases of both prongs of the vagueness test. Application of the test reveals that the mail fraud statute fails both prongs. I conclude, after considering the constitutional implications of the intangible rights doctrine in the political context, that the mail fraud statute is vague on its face.

A. *The Vagueness Doctrine*

The operative effect of the void for vagueness doctrine[151] is simply stated: when a court deems a criminal statute overly vague, it voids the law and will not enforce it. Perhaps because of its simplicity, the power and authority of this doctrine are often underestimated.

---

[150] Section 1346, enacted in 1988, states, in full: "For the purposes of this chapter, the term 'scheme or artifice to defraud' includes a scheme or artifice to deprive another of the intangible right of honest services." 18 U.S.C. § 1346 (1988).

See supra text accompanying notes 67–72 (recounting legislative history demonstrating congressional intent that a right to honest or good government is included in "honest services").

[151] The vagueness doctrine has engaged the attention of many scholars. *See, e.g.*, WAYNE R. LAFAVE & AUSTIN W. SCOTT, JR., CRIMINAL LAW 90–97 (2d ed. 1986) [hereinafter LAFAVE]; PACKER, *supra* note 103, at 71–72; Duke, *supra* note 6; Jeffries, *supra* note 53; Sanford H. Kadish, *Legal Norm and Discretion in the Police and Sentencing Processes*, 75 HARV. L. REV. 904 (1962); Mark Kelman, *Interpretive Construction in the Substantive Criminal Law*, 33 STAN. L. REV. 591, 660–62 (1981); Austin W. Scott, Jr., *Constitutional Limitations on Substantive Criminal Law*, 29 ROCKY MTN. L. REV. 275 (1957); Anthony G. Amsterdam, Note, *The Void for Vagueness Doctrine in the Supreme Court*, 109 U. PA. L. REV. 67 (1960); Michael S. Kelley, Note, *"Something Beyond": The Unconstitutional Vagueness of RICO's Pattern Requirement*, 40 CATH. U. L. REV. 331, 370 (1991); Jeffrey I. Tilden, Note, Big Mama Rag: *An Inquiry Into Vagueness*, 67 VA. L. REV. 1543 (1981).

Vagueness both derives from and effectuates "legality,"[152] a principle that is closely related to the rule of law[153] and has been called the keystone of constitutional government.[154] The relationship between vagueness and legality is symbiotic. The vagueness doctrine rests within the broader concept of legality, which insists on notice of prohibited conduct and prevents government authorities from treating conduct as criminal unless the legislature has defined it as unlawful.[155] Thus, legality and vagueness share the principles that the community may punish only clearly defined conduct and that arbitrary enforcement of the law is not consonant with the rule of law. The vagueness doctrine, through application of the void for vagueness test, is the vehicle that effectuates legality. Courts have used the vagueness analysis to safeguard four distinct components of legality: providing notice to citizens of prohibited behavior; providing standards to prevent

---

[152] *See* Jeffries, *supra* note 53, at 196 ("vagueness doctrine is the operational arm of legality").

Lenity, the rule of strict construction that requires courts faced with vague laws to rule against the government, also evolved from and supports legality. *See* Duke, *supra* note 6, at 912. Lenity is a rule of construction and, as such, states have no constitutional obligation to apply it. States must, however, honor the constraints of the vagueness doctrine. *Id.* at 913 n.12.

[153] *See* Jeffries, *supra* note 53, at 212 (rule of law constrains arbitrary exercise of government power).

For the purposes of this Article, the rule of law is synonymous with governance authorized by a written constitution and effectuated by statutes created by a representative assembly. *See* OLIVER WENDELL HOLMES, JR., THE COMMON LAW 111 (Boston, Little, Brown & Co. 1923) (1881) ("[A]ny legal standard must, in theory, be capable of being known."); ROSCOE POUND, THE SPIRIT OF THE COMMON LAW 183 (Marshall Jones ed., 1921) (supremacy of law is "a doctrine that the sovereign and all its agencies are bound to act upon principles, not according to arbitrary will; are obliged to follow reason instead of being free to follow caprice"). For an application of these principles, *see* Mattison v. Dallas Carrier Corp., 947 F.2d 95, 101 (4th Cir. 1991) (invalidating South Carolina's method of awarding punitive damages because "due process embraces a rule of law which contains standards that can be known in advance, conformed to, and applied rationally").

[154] *See* Kadish, *supra* note 151, at 904 ("root principle"); PACKER, *supra* note 103, at 79–80 ("*The* first principle"); Anthony G. Amsterdam, *Federal Constitutional Restrictions on the Punishment of Crimes of Status, Crimes of General Obnoxiousness, Crimes of Displeasing Police Officers, and the Like*, 3 CRIM. L. BULL. 205, 216 (1967) (rule of law is embodied in the Due Process Clause).

Legal scholars disagree about the source of the legality principle. *See* Jeffries, *supra* note 53, at 195 n.15 (considering the possibility that although legality derives from Enlightenment philosophers, it was resurrected as a reaction to the use of analogical reasoning by the Soviet Union and Nazi Germany to create crimes); Amsterdam, *supra* note 151, at 74 n.38 (noting that vagueness was born in the reign of substantive due process to resist regulatory or economic legislation and later evolved as a means to enforce First Amendment rights).

[155] *See* PACKER, *supra* note 103, at 80 (explaining that conventional reason for legislation is that people must know what conduct is prohibited so that they may act accordingly).

arbitrary and selective prosecution; preventing judicial lawmaking by maintaining the separation of powers; and protecting constitutional rights, especially those of the First Amendment.[156] The vagueness doctrine, and its test, should also be used by courts as the mechanism to preserve the federalist system.

B. *The Vagueness Test*

The Due Process Clauses of the Fifth and Fourteenth Amendments[157] require that criminal offenses be defined precisely.[158] The two-pronged vagueness test gauges compliance with that requirement: the "notice" prong tests whether a statute provides notice to citizens of prohibited behavior, and the "standards" prong tests whether a statute provides standards of enforcement for police and prosecutors. If the language of a statute fails to

---

[156] It has been suggested that courts apply the vagueness test inconsistently, designating legislation as vague only when other rights are at stake. Thus, even vague statutes survive scrutiny unless another constitutional principle is implicated. *See* LAFAVE, *supra* note 151, at 92 (stating that provision of fair notice, vigilance against arbitrary enforcement, and importance of First Amendment "breathing space" form the bases of void for vagueness rulings); PACKER, *supra* note 103, at 85–91 (positing that controlling abuse of discretion and judicial lawmaking is the central concern of the judiciary); Amsterdam, *supra* note 151, at 73–75 (suggesting that "vagueness alone, although helpful and important, does not provide a full and rational explanation of the case development" and arguing that the doctrine is used to protect individual freedoms); Jeffries, *supra* note 53, at 196 (noting that judicial concern includes protecting civil liberties and limiting excessive prosecutorial discretion). Amsterdam's 1960 study offers the best analysis of the vagueness doctrine, and its subtitle, "A Means to an End," still resonates today.

More recently, the attempts to reconcile the vagueness doctrine with the decisions purporting to apply it have been described as "Herculean." Duke, *supra* note 6, at 915.

[157] The Fifth Amendment provides that "[n]o person shall be . . . deprived of life, liberty, or property, without due process of law . . . ." U.S. CONST. amend. V.

The Fourteenth Amendment applies the same standard to the states:

> No State shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.

U.S. CONST. amend. XIV, § 1.

[158] *See* Connally v. General Constr. Co., 269 U.S. 385, 391 (1926) (holding that criminal statute violated due process when the average person "must necessarily guess at its meaning and differ as to its application"); United States v. Harriss, 347 U.S. 612, 617 (1954) (stating that no one "shall be held criminally responsible for conduct which [that person] could not reasonably understand to be proscribed"); Scott, *supra* note 151, at 287 (outlining constitutional limits to criminal prosecutions).

The Sixth Amendment, providing that the accused has a right "to be informed of the nature and cause of the accusation," U.S. CONST. amend. VI, has been used to accomplish the same objective—to provide notice. *See* LAFAVE, *supra* note 151, at 90 (vague statutes have been construed to violate the Sixth Amendment right to be informed of the "nature and cause of the accusation") (citing Yu Cong Eng v. Trin., 271 U.S. 500 (1926) and United States v. L. Cohen Grocery Co., 255 U.S. 81 (1921)).

meet either prong, the statute is void.[159] Although the vagueness doctrine and the test require both notice and standards, the "more important aspect of the doctrine is the requirement that the legislature establish minimal guidelines to govern law enforcement."[160] The familiar incantation, stating that a statute is vague if "men of common intelligence must necessarily guess at its meaning and differ as to its application,"[161] generally describes the standard for both prongs.

1. The Standards Prong

The standards prong, the more sophisticated component of the void for vagueness test, mandates that a statute provide clear standards for enforcement of the law.[162] This prong prevents the abuse of official discretion because the most effective way to control prosecutorial discretion is to require prosecutors to confine their attention to what has already been clearly defined as criminal.[163]

The common flaw of vague statutes is that they inject "into the governmental wheel so much free play that . . . [government] is likely to function erratically [and become] responsive to whim or discrimination."[164] The mandate of clear language reflects the concern that punishment for universally recognized wrongful acts that are not criminal when engaged in may lead to punishment for behavior that is not recognized as wrongful or criminal.[165] The failure to provide such standards is unconstitutional because legislators may not delegate basic policy matters to police officers, judges, and jurors for resolution on an ad

---

[159] Giaccio v. Pa., 382 U.S. 399, 402–03 (1966) (invalidating statute on ground of lack of standards); LAFAVE, *supra* note 151, at 94.

[160] Kolender v. Lawson, 461 U.S. 352, 358 (1983) (citing Smith v. Goguen, 415 U.S. 566, 574 (1974) (stating that most important aspect of vagueness doctrine is requirement that legislature establish minimal guidelines to govern law enforcement)). *See* LAFAVE, *supra* note 151, at 94 (stating that arbitrary and discriminatory enforcement and First Amendment considerations are more likely to be controlling than the notion of fair warning in void for vagueness test); Jeffries, *supra* note 53, at 218 (arbitrary and discretionary enforcement is "most persuasive justification for vagueness review").

[161] *Connally*, 269 U.S. at 391.

[162] *See* Amsterdam, *supra* note 154, at 221 (concluding that Supreme Court's vagueness decisions embody concern about arbitrary and erratic enforcement).

[163] *See* PACKER, *supra* note 103, at 89–90 (commenting on need for devices to ensure fair decisions initiating criminal prosecutions and noting that most such devices are "post-audits" of the decisions of police and prosecutors).

[164] Amsterdam, *supra* note 151, at 90 (concluding that Supreme Court uses vagueness doctrine to control scope and regularity of use of governmental force).

[165] *See* PACKER, *supra* note 103, at 85.

hoc and subjective basis.[166] Vague statutes violate due process by placing "unfettered discretion" in the hands of the police, thereby permitting or even encouraging arbitrary, discriminatory enforcement.[167] Vague statutes also furnish a convenient tool that police and prosecutors may use "against particular groups deemed to merit their displeasure."[168] As the Supreme Court has stated, "the absence of specificity in a criminal statute invites abuse on the part of prosecuting officials who are left free to harass any individuals or groups who may be the object of official displeasure."[169]

The standard prong of the vagueness test constrains the arbitrary exercise of government power by retarding the evils of "caprice and whim, the misuse of government power for private ends, and the unacknowledged reliance on illegitimate criteria of selection."[170] The standards embodied in clear language also guide the determination of guilt,[171] reduce judicial crime creation, and facilitate judicial review of trial court decisions.[172] Consequently, a statute is vague unless it includes standards that guide police and prosecutors in its enforcement.

The incorporation of "honest services" renders the mail fraud statute vague because it does not provide standards to prosecutors which clearly delineate prohibited conduct. The prohibition against dishonest services, in the political context, leads to differential, nonuniform, and inconsistent prosecutions. The undefined term allows prosecutors to decide whether to prosecute

---

[166] *See* Grayned v. City of Rockford, 408 U.S. 104, 108–09 (1972); Giaccio v. Pa., 382 U.S. 399, 403 (1966) ("Certainly one of the basic purposes of the Due Process Clause has always been to protect a person against having the Government impose burdens upon him except in accordance with the valid laws of the land.").

[167] *See* Papachristou v. City of Jacksonville, 405 U.S. 156, 168 (1972) (invalidating a vagrancy law).

[168] *See id.* at 170 (quoting Thornhill v. Ala., 310 U.S. 88, 97–98 (1940)).

[169] *See* Smith v. Goguen, 415 U.S. 566, 575 (1974) (stating that when legislature fails to provide minimal guidelines, a criminal statute may permit "policemen, prosecutors, and juries to pursue their personal predilections"); NAACP v. Button, 371 U.S. 415, 435 (1963) (overly vague law "lends itself to selective enforcement against unpopular causes").

[170] Jeffries, *supra* note 53, at 212 (maintaining that rule of law mandates that "agencies of official coercion should . . . be guided by rules—that is, by openly acknowledged, relatively stable, and generally applicable statements of proscribed conduct").

[171] *See* Scott, *supra* note 151, at 288 (providing standards against which conduct can be measured is one rationale behind the void for vagueness doctrine).

[172] *See* Amsterdam, *supra* note 151, at 80 (noting that review is obstructed by statutory vagueness because overreaching and prejudiced, discriminatory prosecutorial decisions are concealed beneath factual findings that reviewing courts may not generally reexamine).

on the basis of their individual definitions of honesty and their understanding of the political process.[173] Given the possible political motivation for enforcement on either the personal, party, or ideological level, the absence of enforcement standards is particularly troublesome.

2. The Notice Prong

Statutes that do not provide notice to those potentially affected by them offend the most basic principles of due process of law.[174] Due process requires that a penal statute be "sufficiently explicit to inform those who are subject to it what conduct on their part will render them liable to its penalties."[175] Notice that certain behavior is prohibited comprises the most obvious requirement guaranteed by the vagueness doctrine.[176]

Simple fairness requires criminal statutes to be definite because their violation may lead to the deprivation of freedom.[177] If criminal statutes are to guide future conduct and deter wrongful behavior, they must tell people what conduct merits punishment. By providing the opportunity to exercise free will, the notice requirement reinforces a basic tenet of just punishment—that individual culpability and responsibility are prerequisites to a criminal sanction.[178] When behavior is clearly prohibited, individuals must choose affirmatively to disobey the law.

The argument that the term "honest services" fails to provide notice and is therefore void may be received suspiciously by the courts because of the jurisprudence developed to avoid findings of vagueness.[179] Despite the centrality of notice to the fairness

---

[173] *See supra* text accompanying notes 99–110 (discussing prosecutorial discretion).

[174] *See* Connally v. General Constr. Co., 269 U.S. 385, 388, 391 (1926) (invalidating Oklahoma statute that prohibited paying laborers "less than the current rate of per diem wages"). The Constitution also specifically mandates prior notice through the prohibition against ex post facto lawmaking. U.S. CONST. art. I, § 9, cl. 3.

[175] *Connally*, 269 U.S. at 391.

[176] *See* United States v. Protex Indus., Inc., 874 F.2d 740, 743 (10th Cir. 1989) (essence of vagueness doctrine is that potential defendant must have "fair warning") (quoting Bouie v. City of Columbia, 378 U.S. 347, 352 (1964)).

[177] *See Connally*, 269 U.S. at 391 (requirement that criminal statutes be definite is "consonant alike with ordinary notions of fair play and the settled rules of law").

[178] *See* Duke, *supra* note 6, at 912. Sufficiently explicit criminal laws avoid the argument over the indeterminate limits of free will by focusing on choice. To the extent that anyone acts out of free will, the requirements of criminal law are satisfied if individuals are given an opportunity to choose affirmatively a particular course of action.

[179] Courts must also consider the standards prong of the vagueness test. Even if the defendant knew that particular conduct was prohibited, a statute may still be unconsti-

of the criminal justice system, courts have frequently rejected vagueness challenges to legislation even when the language is palpably vague and ambiguous.[180] This is, in part, the result of a strong presumption of validity that attaches to acts of Congress.[181] Congress, however, should not be accorded deference when the vagueness of the legislation allows the government to deny individual constitutional rights or compromise important structural constitutional principles, such as federalism and the separation of powers.

Vagueness challenges based on statutory definitions that contain "an element of degree as to which estimates may differ" also often fail.[182] Courts that rejected vagueness challenges to prosecutions based on "degree," however, developed their rationale in the context of economic regulatory legislation. This jurisprudence is arguably inappropriate in the context of criminal fraud.

Another typical justification for rejecting vagueness challenges is a variation of the tenet that ignorance of the law is not an excuse.[183] This theory holds defendants responsible for "lawyers' notice," that is, for judicial interpretation of a statute.[184]

---

tutionally vague if its vagueness gives police and prosecutors arbitrary power to determine whether to enforce the statute. *See* LAFAVE, *supra* note 151, at 94 n.39. *See also supra* text accompanying notes 162–173 (discussing standards prong).

[180] *See supra* note 156 (discussing inconsistency of vagueness jurisprudence).

[181] *See* United States v. Nat'l Dairy Prods. Corp., 372 U.S. 29, 32 (1963); United States v. Harriss, 347 U.S. 612, 618 n.6 (1954).

[182] *See* Nash v. United States, 229 U.S. 373 (1913) (considering and rejecting vagueness challenge to Sherman Act).

Terms such as "unduly restricting trade" in antitrust actions and "unreasonable allowance" in tax cases are examples of terms that define criminal conduct in terms of degree. The most famous defense of such statutes is that of Justice Holmes, who wrote that charges of unfair notice based on ambiguous terms must fail because they could potentially be levelled at all criminal conduct. *Id.* at 377. According to Holmes, "the law is full of instances where a man's fate depends on his estimating rightly, that is, as the jury subsequently estimates it, some matter of degree. If his judgment is wrong, not only may he incur a fine or a short imprisonment . . . ; he may incur the penalty of death." *Id.* Therefore, courts place the risk that the conduct may be criminal on the potential defendant: "[T]hose who choose . . . to sail as close as possible to the wind inevitably run some risk." *See* Knuller v. Director of Pub. Prosecutions, 1973 App. Cas. 435, 488 (opinion of Lord Simon of Glaisdale); *see also* Boyce Motor Lines, Inc. v. United States, 342 U.S. 337, 340 (1952) (stating that it is not unfair that one who "deliberately goes perilously close to an area of proscribed conduct shall take the risk that he may cross the line").

[183] *See* HOLMES, *supra* note 153, at 47–48 (justifying precept that everyone must be presumed to know the law because otherwise ignorance of the law is encouraged). *But see* JEROME HALL, GENERAL PRINCIPLES OF CRIMINAL LAW 38–83 (2d ed. 1960) (questioning Holmes's thesis).

[184] *See* United States v. Bohonus, 628 F.2d 1167, 1174 (9th Cir.), *cert. denied*, 447 U.S. 928 (1980) (rejecting vagueness challenge to mail fraud statute in part because

The fiction that judicial decisions provide notice should not be accorded undue deference. Formal notice, in the form of judicial precedents, does not rise to the level of notice that due process requires.[185]

Finally, courts employ the requirement of specific intent to uphold vague statutes.[186] Courts reason that since specific intent requires the government to prove an intention to defraud, the defendant cannot claim that he or she was unaware that the act was prohibited once such intent is established.[187] Although substantial precedent supports this reasoning,[188] it is difficult to understand how specific intent nullifies the notice prong of the vagueness test in the context of mail fraud. Defendants in political corruption cases may or may not intend to reap a personal financial benefit, but may not intend to defraud the state or citizenry of money or property, and in fact do not. It is also unlikely that political corruption defendants specifically intend to deprive the citizenry of their right to good government. Moreover, such defendants cannot have intended to defraud insurance companies, or other third parties, of either money or intangible rights, because typically such parties initiate and benefit from the arrangements that are the subject of the prosecution. Finally, even if the defendants intended to divert funds, they arguably did not intend to violate the statute. It is, after all, possible to cause a result intentionally and yet be without fair warning that such conduct is criminal.[189]

The mail fraud statute fails the notice test because the term

---

the numerous judicial decisions applying it to deprivation of intangible rights afforded defendant reasonable notice).

[185] *See* Jeffries, *supra* note 53, at 208 (arguing that when notice must be garnered from various inaccessible sources, the concept becomes distinctly unrealistic).

[186] *See* Bohonus, 628 F.2d at 1174 (mail fraud statute's specific intent requirement negates the notice problem); *see also* United States v. Margiotta, 688 F.2d 108, 129 (2d Cir. 1982) (rejecting vagueness challenge to mail fraud statute because of specific intent requirement), *cert. denied*, 461 U.S. 913 (1983); United States v. Bailin, No. 89-CR-668, 1990 U.S. Dist. LEXIS 9168, at *15 n.6 (N.D. Ill. July 17, 1990) (rejecting vagueness challenge to RICO statute because mail fraud is specific intent crime).

[187] *See Bohonus*, 628 F.2d at 1174.

[188] *See* Village of Hoffman Estates v. Flipside, Hoffman Estates, Inc., 455 U.S. 489, 499 (1982) (stating that "scienter requirement may mitigate a law's vagueness, especially with respect to the adequacy of notice"); *see also* United States v. Gaudreau, 860 F.2d 357 (10th Cir. 1988); Nova Records, Inc. v. Sendak, 706 F.2d 782 (7th Cir. 1983).

[189] *See* LAFAVE, *supra* note 151, at 93–94 (commenting that unclear language is not clarified by scienter because scienter requires only knowledge of consequences of action and is distinct from knowledge of prohibition by statute); Boyce Motor Lines, Inc. v. United States, 342 U.S. 337, 345 (1952) (Jackson, J., dissenting) (arguing that scienter cannot cure vagueness).

"honest services" does not clearly state the content of criminal conduct.[190] "Honest services" is an evolving, aspirational term that describes a level of conduct that may never be attained. In addition, the term "honest," like the concept of good government, embodies a relative value.[191] In real world politics, only a blurred and shifting line separates political corruption from political patronage, and honest from dishonest service.[192] Since state law also often fails to prohibit political corruption in a well-defined manner, local political practices become a significant factor in state and local cases.[193] The political tradition of a local area may approve or tolerate officials of the winning

---

[190] The inherent notice problems of the term "honest services" were discussed at congressional hearings on proposed amendments to the mail fraud statute following *McNally*. The following exchange, between Ronald Stroman, assistant counsel for the subcommittee, and John C. Keeney, Acting Assistant Attorney General, is instructive:

> Mr. STROMAN. Well, honest services of public official, do you think that is specific? I mean what does "honest services" mean? Certainly if I am a public official
>
> Mr. KEENEY. Well, it means that—it means what the circuit courts of appeals have been saying for years that when a Mandel or a Kerner corruptly uses his office he is depriving the citizens of that State of his honest services.
>
> . . . .
>
> Mr. STROMAN. I would wholeheartedly agree with that, but certainly the concept of intangible rights has been interpreted by a whole host of cases as well. To use the term "honest government" and say that is more specific than intangible rights when you have got the same history of case law, quite frankly I do not see the distinction.
>
> If I am an official in the Government and I see the term "honest government," that certainly does not alert me anymore than the existing statute as to what you are trying to cover. I do not know what that means. I would have to read the cases that you referred to.
>
> If I read the mail fraud statute, the same situation applies. I would have to read the cases to specifically understand what the statute is attempting to get at. And my point is that if you say that what you are trying to do is create a new statute because it is more specific, quite frankly it is not anymore [sic] specific.

Mail Fraud Hearings, *supra* note 6, at 48–49.

[191] *See* Kelman, *supra* note 151, at 661–62 (noting that statutes referring to values are inevitably vague because there can be no precise understanding of the content of a value that, by definition, is not communally shared).

[192] *See* Rutan v. Republican Party of Ill., 497 U.S. 62, 102–10 (Scalia, J., dissenting) (arguing that patronage system is inherent in American politics and is also necessary to the two-party system), *reh'g denied*, 497 U.S. 1050 (1990). Notwithstanding this view, the Supreme Court has narrowed the scope of permissible patronage. *See id.*; Branti v. Finkel, 445 U.S. 507 (1980); Elrod v. Burns, 427 U.S. 347 (1976).

[193] *See* Edward J. Imwinkelried & Ephraim Margolin, *The Case for the Admissibility of Defense Testimony About Customary Political Practices in Official Corruption Cases*, 29 AM. CRIM. L. REV. 1, 6–11 (1991) (presenting various theories under which other legislators' testimony about customary political practices are logically relevant to political corruption defense).

party awarding those who helped secure the election with jobs or insurance contracts. In such cases, the defendant is without notice that such patronage may be, nonetheless, considered criminal under a federal statute—much less a federal statute written in large part by the judiciary. Democratic principles protect those traditions if a majority of the citizens of a state choose to endorse or to tolerate political practices that others may find "dishonest."[194]

3. Facial Vagueness

A statute may be vague on its face or vague as applied to a particular defendant.[195] Section 1346, the legislation that incorporates the intangible rights doctrine into the mail fraud statute, is facially vague because it fails to provide standards to guide police and prosecutors in the exercise of their discretion.[196] In addition, the decisions of the Supreme Court finding statutes unconstitutionally vague vindicate two of the three constitutional interests that are implicated by the intangible rights doctrine: the threat to political activity protected by the First Amendment, and the prohibition against judicial lawmaking, which is constitutionally forbidden by separation of powers doctrine. The third interest, the protection of the federalist structure and the concomitant prevention of the diminution of state and local autonomy, is analogous and also justifies a finding of facial vagueness.

The vagueness doctrine protects individual civil liberty by authorizing courts to give heightened scrutiny to vague statutes that may chill First Amendment or other constitutionally protected activity.[197] The vagueness principle, in practice, has been

---

[194] *See* 86 CONG. REC. 2557 (1940) (relating remarks of Senator Pepper (D-Fla.) that it is "the privilege of a State in a democratic Government even to have bad government, if its people want it to be bad"). The Senator's remark may reflect the perception of De Tocqueville that

> It is incontestable that the people frequently conduct public business very [poorly]; but is it impossible that the [people] should take part in public business without extending the circle of their ideas, and quitting the ordinary routine of their thoughts.

1 DE TOCQUEVILLE, *supra* note 85, at 320.

[195] Courts prefer to avoid holding a statute facially vague. *See* United States v. Mazurie, 419 U.S. 544, 550–53 (1975) (citing United States v. Nat'l Dairy Prods. Corp., 372 U.S. 29 (1963)).

[196] *See* Kolender v. Lawson, 461 U.S. 352, 361 (1983) (invalidating loitering statute as facially vague on sole ground that it encourages arbitrary enforcement).

[197] *See* LAFAVE, *supra* note 151, at 96 (reviewing Supreme Court's concern that

used to maintain a buffer zone for civil rights.[198] The doctrine ensures that individuals who may otherwise be vulnerable to arbitrary or selective prosecution and judicial criminalization will not hesitate to engage in lawful conduct.[199] Providing police and prosecutors with clear enforcement standards by invalidating laws that allow arbitrary enforcement guarantees that unpopular political expression is not suppressed and that political participation is not unconstitutionally limited.[200]

Invalidating the mail fraud statute will remove the threat of criminal sanctions for conduct that is essential to the political process. We assume that those active in politics pursue power for reasons somehow related to self-interest—whether to enact a program to benefit the majority, a minority, or just a few.[201] To expect political activists to act impartially is unrealistic and inimical to the political process.[202]

The vagueness doctrine, by demanding specific criminal elements, reduces judicial lawmaking.[203] Vague statutes force the judiciary to define criminal conduct. Courts, in order to give

---

exercise of First Amendment rights not be inhibited); Amsterdam, *supra* note 151, at 80 (maintaining that vague statutes are dangerous because they allow the state to "get away with more inhibitory regulation than it has a constitutional right to impose"); Duke, *supra* note 6, at 912–13 (concluding that ambiguous criminal prohibitions defeat the values served by dissuaded conduct when that conduct is constitutionally protected).

[198] *See* Amsterdam, *supra* note 151, at 75 (concluding that Supreme Court uses unconstitutional indefiniteness to protect the peripheries of civil rights by mediating between the "organs of public coercion" and the constitutional protections accorded to liberty). The Court has repeatedly applied the vagueness doctrine strictly to First Amendment rights, even when the statute does not specifically infringe the claimant's constitutional rights. *See* Plummer v. City of Columbus, 414 U.S. 2 (1973); Dombrowski v. Pfister, 380 U.S. 479 (1965); Thornhill v. Ala., 310 U.S. 88 (1940). *But see* Buckley v. Valeo, 424 U.S. 1, 41–45 (1976) (holding that Federal Election Campaign Act's disclosure and reporting requirements were not unconstitutionally vague in violation of First Amendment speech and association rights because of narrow construction of Act).

[199] *See Buckley*, 424 U.S. at 41 n.48 (rationale for facial invalidation is that vague laws inhibit protected expression by inducing "citizens to steer far wider of the unlawful zone than if boundaries of the forbidden areas were clearly marked") (internal quotation marks omitted; citations omitted).

[200] *See* NAACP v. Button, 371 U.S. 415, 433 (1963) (invalidating legislation on ground that it discouraged people from exercising their First Amendment rights).

[201] *See Margiotta*, 688 F.2d 108, 143 (2d Cir. 1982) (Winter, J., dissenting) (stating that pluralistic political system assumes that the politically active will pursue power and self-interest), *cert. denied*, 461 U.S. 913 (1983).

[202] *Id.* (concluding that unless Congress passes specific legislation specifying what political acts are criminal, enforcement of obligation that public affairs be conducted honestly should be by political processes of public debate, free press, and alert electorate).

[203] *See* PACKER, *supra* note 103, at 91 (noting public's negative reaction to "judge-made offenses" following *Shaw v. Director of Public Prosecutions*, 2 Weekly L.R. 897 (1961)).

content to vague statutes, convert the facts of the case to the elements of the offense, effectively incorporating those "elements" into the statute.[204] Federal prosecutors are encouraged to pressure the courts to expand their arrogation of the legislative function of defining criminal conduct. Such usurpation, especially when accomplished through indirect means, is fundamentally undemocratic[205] and violates the principles of the separation of powers.[206] By voiding vague statutes, courts eliminate the problem that requires trial courts to employ common-law methodology—a methodology that is particularly inappropriate in criminal law.

The Constitution mandates a federalist structure. As we have seen, federal intervention based on the intangible rights doctrine is inimical to state and local autonomy.[207] When state and local public officials are prosecuted, there is "intense public interest in ensuring the integrity of the prosecution."[208] Although the vagueness doctrine has not yet been used to protect the federalist structure, there is no principled reason why it should not be. The constitutional value of federalism is as deserving of protection as those of the separation of powers and the First Amendment. The vague mail fraud statute allows untoward and often unwanted federal intervention into state politics, results in the diminution of state and local autonomy, and upsets the balance of power between the national and state governments.

IV. A Legislative Solution

Given that "honest services" is unconstitutionally vague, Congress is presented with a problem. The determination that the

---

[204] *See supra* text accompanying notes 111–128 (discussing judicial lawmaking).

[205] *See* Scott, *supra* note 151, at 288 (concluding that a concern for fairness caused legislatures to codify criminal law and to abolish common-law crimes); *see also* Jeffries, *supra* note 53, at 221 (common-law process should be avoided in criminal law because fact-specific innovation is inimical to the rule of law and violates democratic values); LaFave, *supra* note 151, at 92.

[206] *See* Duke, *supra* note 6, at 912 (examining values served by vagueness and lenity doctrines); Williams, *supra* note 6, at 149 (tracing principle of legality to separation of powers theory that separates the functions of defining crime and determining punishment). According to Enlightenment philosophy, which largely provides the philosophical framework of the Constitution, liberty requires the separation of the judiciary from the legislative and executive branches. *See* Charles Montesquieu, The Spirit of the Laws 152 (T. Nugent trans., 1949).

[207] *See supra* text accompanying notes 93–110 (discussing effect of mail fraud prosecutions in political corruption cases on state and local autonomy).

[208] *See* Imwinkelried & Margolin, *supra* note 193, at 2 (noting federalist ramifications of increased number of federal prosecutions for political corruption).

amended mail fraud statute is unconstitutionally vague restores *McNally*'s ruling that the language of the mail fraud statute, a "scheme to defraud," does not protect intangible rights to honest services or good government.[209] Congress should redefine a "scheme to defraud" by codifying the objective standard of *McNally*,[210] and should not enact new criminal legislation that is based on the intangible right to honest services.

A. *A New Mail Fraud Statute*

The following draft of a new mail fraud statute attempts to accomplish three objectives.[211] The most compelling objective is to define the crime clearly. Second, the new statute seeks to mitigate the inchoate nature of mail fraud, which places undue emphasis on intent. Third, the proposal changes a mailing from a substantive element of mail fraud to a jurisdictional basis.

(a) Whoever knowingly—
(1) schemes to defraud with intent to—
(A) obtain money or property belonging to another person or a government entity; or
(B) cause economic loss to another person or to a government entity; or
(2) engages in a substantial step in a course of conduct planned to culminate in the realization of such scheme; shall be guilty of a felony, punishable by a fine of not more than $_____ or imprisonment for not more than ___ years, or both;
(b) There is federal jurisdiction over an offense in this section if, in furtherance of such scheme, the actor uses or causes the use of the United States mail or an interstate or foreign communication facility, including but not limited to a facility of wire, radio, television, or telecommunication;

---

[209] *See supra* text accompanying notes 54–66 (discussing *McNally*).

[210] Although it is not strictly necessary to codify *McNally*, Congress should do so because the legislative process of converting judicial law into statutory law provides a public forum for debate, educates the public, provides notice, and encourages uniform application.

[211] This proposal is based on a model offered in 1983. *See* Hurson, *supra* note 13 (offering model statute codifying the right to honest services and limiting its application to elected or appointed public officials); Coffee, *supra* note 54, at 147–52; Coffee, *supra* note 13, at 12–13; Oxman, *supra* note 9, at 778–89.

(c) As used in this section—

(1) "scheme to defraud" means any plan, devise, contrivance, or program of action involving—

(A) deception; and

(B) any one or more of the following additional characteristics:

(a) the receipt of money or property by the actor, or the loss of money or property by the victim, of an aggregate potential or actual value of $_____ or more;

(b) a sale and purchase of property, goods, services, or financial instruments, provided the deception involved relates to the quality, quantity or price of the property, goods, services, or instruments, and not solely to the manner in which they were marketed, promoted, or sold, and the aggregate value (or amount paid) for the property, goods, services or instruments is $_____ or more.

(2) "property" includes tangible or intangible interests but does not include a deprivation of exclusive use.

(3) "substantial step" includes conduct that is strongly corroborative of the actor's purpose.

(4) "deception" means the taking by the actor or his or her causing to be taken by others, with the intent to mislead, any of the following actions:

(A) knowingly making a material false statement;

(B) knowingly failing to make a material statement, or omitting material information, or concealing a material fact necessary to make a statement not misleading;

(C) knowingly using a trick, scheme, or device;

(D) knowingly submitting or inviting reliance upon a writing or physical object that is misleading in a material respect.

(5) "material statement" is one that a reasonable decisionmaker would likely find important in making the decision.

(6) "government entity" includes the federal and state governments and any political subdivision or agency of either.

1. Money or Property

The *McNally* decision struck an effective balance between protecting structural constitutional provisions and maintaining a

federal role in the control of political corruption on the state and local level. The mechanism that accomplishes that balance, the money or property element, is therefore incorporated into the proposed statute. The imposition of an objective standard of pecuniary loss properly defines a "scheme to defraud" and prevents the development of future vagueness problems. Because people commonly understand that theft of money or property through fraud is a criminal act, defendants can be charged fairly with prior notice. The requirement of cognizable harm serves as a threshold standard to guide prosecutors in the decision to prosecute and reduces the possibility of selective or arbitrary prosecution. The objective standard also reduces prosecutorial discretion, ensuring that criminal conduct may not be defined by an individual prosecutor's subjective concept of good government. The restoration of the money or property element is thus a constructive solution to the problems inherent in the undefined "scheme to defraud."

The pecuniary element largely eliminates the three constitutional objections to the intangible rights doctrine considered in this Article. First, the property qualification properly allocates primary authority to monitor state and local politics to state governments.[212] The restoration of the *McNally* decision acknowledges the paramount interest of the states in controlling political corruption on the state and local level. Without a federal presence, the decision to criminalize the political practices of state and local politicians is correctly left to state legislators who are well-positioned to enact appropriate legislation. State legislators understand both the values of the electorate and the realities of state politics, patronage, and party loyalty.[213]

Second, the money or property standard discourages judicial crime creation and restores legislative authority over criminal legislation.[214] A pecuniary loss becomes an element of mail fraud

---

[212] *See supra* text accompanying notes 93–110 (discussing effect of intangible rights doctrine on federalism).

[213] *See* Williams, *supra* note 6, at 157 (noting that states have recognized the significant political judgments necessary to legislate in this area and criticizing the use of mail fraud as an ad hoc and superficial method of addressing the problems of political corruption).

Reliance on state legislators may reduce the passage of strict laws, but this cost of representative democracy is mitigated by the public debate accompanying such legislation. Public debate results in public accountability, so that the electorate can fairly hold state and local politicians to whatever standard is enacted.

[214] *See supra* text accompanying notes 111–117 (discussing effect of intangible rights doctrine on separation of powers).

that the government must prove, foreclosing the judicial practice of formulating broad jury instructions that essentially restate the indictment. When jury instructions no longer define "honest services" in terms of the defendant's conduct, juries will determine guilt by applying a standard that exists outside of the facts and circumstances of the case.

Third, the elimination of the intangible right to honest services limits the possibility that engaging in party politics or aggressively representing or working for an interest group will result in criminal charges.[215] Political activists may pursue the goals of the interest groups they represent without fear that failing to disclose that interest will expose them to criminal liability. Conversely, anyone who defrauds the state of money or property may be prosecuted, making it unnecessary to determine whether a nonelected volunteer has assumed a fiduciary duty to the electorate as a whole.

The element of economic loss in section (a)(1) of the proposal properly applies to both the public sector and the private sector.[216] Notwithstanding substantial precedent that allows mail fraud prosecution of private sector employees and fiduciaries for failing to provide "honest services," these private sector prosecutions also infringe on separation of powers values.[217] It is also clear that the original purpose of the statute, protecting the federal interest in the mail, is also served by prosecutions in the private sector.[218] In addition, there is no principled basis on which to distinguish between public and private fraud because the consequences of private sector fraud often affect the entire community.[219] Furthermore, if a distinction is made between pri-

---

[215] *See supra* text accompanying notes 118–128 (discussing effect of intangible rights doctrine on First Amendment political activity).

[216] *McNally* does not distinguish between private and public sector cases. Thus, private sector employers and fiduciaries have no claim to an intangible right to honest services. *But see Mail Fraud Hearings*, *supra* note 6, at 95 (statement of Terrance G. Reed, National Association of Criminal Defense Lawyers, Washington, D.C.) (expressing view that *McNally* decision affects only public sector cases).

[217] The potential for judicial crime creation is equally likely in private as in public cases, and thus the separation of powers doctrine weighs in favor of a facial vagueness finding. However, neither federalism nor the First Amendment appear to be offended by prosecutions of private defendants under the mail fraud statute. The general complaint about the encroachment of the federal government into state responsibility for criminal law is unpersuasive given the development of a significant body of federal criminal law.

[218] Economic gain is usually the object of private sector fraud, so the property limitation may not affect many private sector prosecutions anyway.

[219] The recent explosion of private sector fraud cases has profoundly affected the savings and loan industry, Wall Street, and the banking industry.

vate and public sector mail fraud, result-oriented prosecutorial decisions are inevitable. Someone must determine whether the defendant schemed to defraud the public of a property right or schemed to defraud an employer or principal of a right to honest service that resulted in a property loss to the public. That "someone" is likely to be the prosecutor. Thus, the objections based on prosecutorial discretion argue for rejecting any distinction between the two sectors.

Limiting the mail fraud statute to the deprivation of a pecuniary interest does not entirely eliminate federal intervention because federal authorities will continue to prosecute schemes to defraud a state or local government entity of money or property. In addition, the objective standard of money or property does not entirely eliminate prosecutorial flexibility and therefore discretion because the standard is not confined to traditional notions of property.[220] As the Supreme Court has held, the concept of property includes and therefore protects intangible property that has been accorded the status of property.[221] The flexibility of the concept of property introduces a problematic potential for expansion, which reviewing courts should resist. When courts determine whether protected property has been lost, the nuances of the law of property rights may result in an unwarranted expansion of the kinds of property protected by criminal sanctions.[222]

2. Intent

It is impossible to eliminate entirely the inchoate nature of the crime of mail fraud. First, the statute must encompass the incipient offense because there is no general federal attempt statute.[223] Second, the purpose of the mail fraud statute has evolved to protect the public from a wide range of fraudulent schemes. It would be presumptuous, as well as unavailing, to reframe that

---

[220] *See* Section (c)(2) of the proposed statute, *supra* part IV.A. This definition ensures that victims suffer economic loss, whether the property is tangible or not.

[221] *See* United States v. Carpenter, 484 U.S. 19 (1987); *supra* note 72 and accompanying text (discussing property right in information, *Carpenter*, and cases interpreting *McNally* property qualification).

[222] An example of such a problematic concept is the constructive trust. This judge-made civil law concept creates a property right by allocating the defendant's gain to the victim on the theory that the gain was effectively held in trust by the trustee/defendant for the owner/victim. *See supra* note 72 (noting attempts to avoid implications of *McNally*).

[223] *See supra* text accompanying notes 29–35 (discussing inchoate nature of mail fraud statute).

purpose at this stage. Consequently, this proposal continues to define the crime in terms of scheming to defraud. Nevertheless, it is possible and desirable to reduce the undue emphasis on the actor's intent.[224] Section (a)(2) of the proposed statute endeavors to accomplish this goal by requiring the prosecution to prove that the actor did more than devise a plan to defraud. Accordingly, this provision requires the actor to engage in a substantial step that strongly corroborates a criminal purpose.[225]

This language, taken from the Model Penal Code, requires some act beyond preparation to verify the actor's intent to carry out the criminal plan.[226] This standard emphasizes what has been done, rather than what remains to be done. Consequently, the imposition of criminal liability is not allowed unless the actor demonstrates a firm criminal purpose.[227] A mailing or similar jurisdictional communication would be strong evidence of purpose, although other conduct would also satisfy the requirement.

The proposal specifically defines "deception." Deception includes omissions of a material nature and those omissions that fail to correct an impression made by a former statement. A material statement is determined by the victim's need to know the information, which is in turn defined by the objective standard of whether a reasonable person would use the information to reach a decision.

3. Jurisdiction

The proposal acknowledges that the federal interest in the mail has been converted to a jurisdictional base.[228] Thus, section (b) bases jurisdiction on the use of the mail. The provision also confers jurisdiction when the actor uses any interstate communication facility over which the federal government presides.

---

[224] *See supra* text accompanying notes 22–28 (discussing emphasis on actor's intent in mail fraud).

[225] This language follows the practice adopted by federal courts. *See* United States v. Cruz-Jimenez, 977 F.2d 95 (3d Cir. 1992); United States v. Valencia, 907 F.2d 671 (7th Cir. 1990); United States v. Savaiano, 843 F.2d 1280 (10th Cir.), *cert. denied*, 488 U.S. 836 (1988); United States v. Cartlidge, 808 F.2d 1064, 1067 (5th Cir. 1987); United States v. Forbrich, 758 F.2d 555 (11th Cir. 1985); United States v. McFadden, 739 F.2d 149 (4th Cir.), *cert. denied*, 469 U.S. 920 (1984); United States v. Mowad, 641 F.2d 1067 (2d Cir.), *cert. denied*, 454 U.S. 817 (1981); *see also* Coffee, *supra* note 13, at 13 (suggesting same language).

[226] *See* MODEL PENAL CODE § 5.01(1)(c)–(2) (1962); GEORGE P. FLETCHER, RETHINKING CRIMINAL LAW § 3.3.1 (1978).

[227] *See* MODEL PENAL CODE § 5.01 cmt. (Tentative Draft No. 10, 1960).

[228] *See supra* text accompanying notes 14–21 (discussing use of the mails).

This jurisdictional definition allows for technological advances that will include new forms of federally regulated communication. The national government's interest in preventing the misuse of these facilities to defraud those who rely on them is identical to the government's interest in preventing misuse of the mail. There is no reason to limit that interest to the post office.

This formulation of jurisdiction effectively eliminates wire fraud as a separate offense. The mail fraud and wire fraud statutes, whose language and purposes are so similar that courts apply the same case law to either statute, are virtually indistinguishable.[229] In recognition of the constraints of a federal interest, the provision retains the requirement of an interstate communication. If an intrastate communication conferred federal jurisdiction, jurisdiction would effectively be based on the Commerce Clause. Basing jurisdiction on an effect on interstate commerce inappropriately converts mail fraud to simple fraud, a crime that is the subject of state laws. For the same reason, federal jurisdiction does not include the use of private mail carriers.

Making the federal interest in the mail jurisdictional and removing a mailing as a criminal element also eliminates the prosecutorial practice of stacking mail fraud counts.[230] Prosecutors can presently add a mail fraud count for every mailing, resulting in multiple counts even when the actor intended only one scheme to defraud. There is no principled justification for this practice.

B. *Public Corruption Legislation*

Congress is presently considering public corruption legislation based on the intangible rights doctrine in the public sector.[231] Based on pre-*McNally* case law, this proposal codifies the vague

---

[229] *See* 18 U.S.C. § 1343 (1988 & Supp. IV 1992). *See, e.g.*, United States v. Von Barta, 635 F.2d 999, 1005 n.11 (2d Cir. 1980), *cert. denied*, 450 U.S. 998 (1981).

[230] *See* Hurson, *supra* note 13, at 435 (discussing practice); Morano, *supra* note 13, at 48 (discussing practice); *supra* text accompanying note 35.

[231] H.R. 3355, 103d Cong., 1st Sess. § 4402, (1993) [hereinafter McConnell Amendment]. H.R. 3355, the Violent Crime Control and Law Enforcement Act of 1993, was in conference as this article was printed, and it was unclear whether the McConnell Amendment would remain a part of the crime bill that was ultimately enacted.

A substantially similar proposal, the Anti-Public Corruption Bill, was presented in 1988 as the Department of Justice's attempt to counter the *McNally* decision, but did

intangible rights doctrine whose constitutional implications have been the subject of this Article. The bill reverses the Supreme Court's *McNally* holding, as did the amendment to the mail fraud statute,[232] by specifically criminalizing the deprivation of the honest services of a state or local government official or employee.[233] It does not, however, circumvent the fact that the intangible rights doctrine is inherently and unconstitutionally vague. All of the objections to the term "honest services" in the mail fraud statute apply also to this proposed legislation.[234]

Among its most obvious infirmities, the public corruption bill expands even further the emphasis on intent by prohibiting an "endeavor" to deprive the inhabitants of a state or state subdivision of honest services.[235] The authors purposely declined to define the specific level of intent required, relying instead on case law holding that a mailing, "for the purpose of executing or concealing such scheme or artifice," satisfies the element of intent.[236]

Not only is the public corruption legislative proposal fundamentally flawed on grounds of constitutional vagueness and principles of common law, it is also premature. Significant research and analysis should be completed before Congress enacts legislation directed to political corruption. First, the putative

---

not survive House Judiciary Committee hearings. *See Mail Fraud Hearings*, *supra* note 6, at 23 (presenting text of Anti-Public Corruption Bill).

Subsequently, the Senate included public corruption legislation in its proposed crime bills of 1989, 1990, 1991, and 1992. *See* 135 CONG. REC. S12,430–32 (daily ed. Oct. 3, 1989) (statement of Sen. Biden (D-Del.)); 136 CONG. REC. S6638–39 (daily ed. May 21, 1990) (statement of Sen. Biden); 137 CONG. REC. S9382–83 (daily ed. July 9, 1991); 138 CONG. REC. S6911–12 (daily ed. May 19, 1992).

[232] *See supra* text accompanying notes 67–72 (discussing amendment).

[233] *See* McConnell Amendment, *supra* note 231, § 226(a)(1). The provision states:

> Whoever, in a circumstance described in paragraph (3), deprives or defrauds, or endeavors to deprive or to defraud, by any scheme or artifice, the inhabitants of a State or political subdivision of a State *of the honest services of an official or employee of the State, or political subdivision*, shall be fined under this title, imprisoned not more than 10 years, or both. (emphasis added)

A similar provision covers federal public officials. *Id.* § 226(b). The bill also prohibits election fraud and protects whistle-blowers. *Id.* § 226(a)(2), (c).

[234] *See supra* text accompanying notes 73–128 (discussing constitutional implications of the intangible rights doctrine).

[235] *See* McConnell Amendment, *supra* note 231, § 226(a).

[236] *See* McConnell Amendment, *supra* note 231, § 226(a)(3)(A).

In addition, the bill's definition of "official" is over inclusive and includes unelected officials. An official is any person who is employed or exercises authority derived from a government position or holds a position in a government subdivision; is nominated, appointed, or selected for office; or has been informed that he or she will be nominated, appointed, or selected. *See id.* § 226(d)(1).

federal interest should be evaluated and defined, focusing on whether that interest justifies an infringement of state and local autonomy. Next, Congress should explore alternative methods of accommodating the federal interest with the interests of state and local governments. For instance, Congress could achieve both goals by placing the federal investigative advantage at the disposal of state prosecutors to supplement state prosecutorial efforts.[237] Congress should examine the efforts of those states that have adopted anti-corruption laws, evaluating such laws by reference to their statutory language and their success in effecting convictions. Congress should also determine how such an anti-corruption statute might affect existing federal criminal law.[238] Finally, and most important, in order to pass constitutional muster, such legislation must clearly ban specific behavior rather than assert an undefined right to honest services.[239] There is no evidence that Congress or any appropriate subcommittee has formulated, addressed, or resolved any of these issues.

CONCLUSION

We face a choice. We can allow state and local legislators and prosecutors, in response to the voice of their electorate, to define corrupt behavior and to remove from office those state and local officials who fall within that definition. Or we can accept federal intervention in state and local politics in an effort to impose a properly working democratic process. The choice is profoundly political, depending as it does on personal visions of democracy

---

The bill also expands jurisdiction to include actions that affect interstate commerce. *See id.* § 226(a)(3)(A)(iii)–(iv). Although most state and local political corruption is, by definition, confined to state boundaries, jurisdiction is conferred over wholly intrastate activity on the accepted theory that intrastate commerce affects interstate commerce. *See id.* § 226(d)(5).

[237] Although this suggestion has been made before, it has not been given serious consideration. *See supra* note 99 (discussing suggestion that federal prosecutors turn over results of investigations to state prosecutors).

[238] Public corruption legislation should take into account other federal statutes that are used to prosecute crimes of a political nature, such as extortion. *See, e.g.*, McCormick v. United States, 111 S. Ct. 1807 (1991) (interpreting Hobbs Act); Evans v. United States, 112 S. Ct. 1881 (1992) (interpreting Hobbs Act).

[239] A useful starting point is the Model Penal Code provision that covers offenses against public administration. *See* MODEL PENAL CODE §§ 240–243 (1962). Section 240.7, which prohibits selling political endorsements and using special influence, is especially pertinent.

and faith in self-government. Both alternatives pose a final problem: Who will guard the guardians? Although this problem will never be perfectly resolved, the electorate, who have ultimate responsibility, can more effectively safeguard the preservation of constitutional and political integrity when the persons entrusted with monitoring political corruption are state and local officials.

NOTE

INTERPRETIVE DIRECTIONS IN STATUTES

ALAN R. ROMERO*

Legislatures have employed a variety of strategies in their attempts to influence the interpretation of statutes. One strategy involves the use of "interpretive directions," statutory provisions that direct courts to interpret statutes in a particular way.

In this Note, Mr. Romero analyzes the effectiveness and desirability of interpretive directions. He raises both constitutional and practical objections to the use of interpretive directions, and discusses possible judicial responses to such legislative action. He concludes that legislatures can best control the application of the laws by taking more care in drafting the substantive provisions of statutes rather than by enacting interpretive directions.

Legislatures direct the application of laws mainly by detailing their substance. If a legislature wants to require certain conduct, it can simply specify that conduct in the law it passes. However, no amount of description or definition will remove all ambiguities or cover all possible circumstances. Even carefully drafted statutes are often open to more than one fair interpretation. Courts therefore must decide how to apply statutes in particular cases. Inevitably, sometimes they will apply statutes in ways the legislature did not expect and might not have desired.

Scholars have discussed a number of ways to increase the likelihood that courts will apply the laws as the legislature intended. Many of these are directed toward courts rather than the legislature. For example, some have argued that legislative history can help to identify the intention of the legislature.[1] On the other hand, some contend that using legislative history is actually an obstacle to determining what a statute really means.[2]

---

* Law Clerk to Hon. Stephen H. Anderson, United States Court of Appeals for the Tenth Circuit. B.A., Brigham Young University, 1990; J.D., Harvard Law School, 1993. The author would like to thank Professor David Shapiro for his helpful suggestions.

[1] *See, e.g.*, Stephen Breyer, *On the Uses of Legislative History in Interpreting Statutes*, 65 S. CAL. L. REV. 845 (1992); James M. Landis, *A Note on "Statutory Interpretation"*, 43 HARV. L. REV. 886, 889–90 (1930); Abner J. Mikva, *A Reply to Judge Starr's Observations*, 1987 DUKE L.J. 380.

[2] *See, e.g.*, Wallace v. Christensen, 802 F.2d 1539, 1559 (9th Cir. 1986) (Kozinski, J., concurring in the judgment) (arguing that the use of legislative history "creates strong incentives for manipulating legislative history to achieve through the courts results not achievable during the enactment process"); Kenneth W. Starr, *Observations About the Use of Legislative History*, 1987 DUKE L.J. 371 (recommending restraint in the use of legislative history for both pragmatic and theoretical reasons).

Fewer scholars have suggested ways that legislatures could control statutory interpretation by the courts. Some have emphasized the importance of improved drafting, reasoning that clearer laws would need less interpretation.[3] Some have considered whether legislatures might better implement their will by increasing dialogue with the judiciary, responding to court interpretations by amending statutes and passing further legislation to make the legislative will known, or even by adopting more formal facilitating mechanisms.[4] Several also have pointed out that the legislature can better control the application of its laws if it drafts them in light of judicial interpretive theories.[5]

Of course, judges probably will always disagree about interpretive rules and processes, which makes it difficult for the legislature to anticipate how statutes will be interpreted. But maybe the legislature could impose consistent interpretive rules on all judges by statute. If this were possible, the legislature would not have to speculate as much about how the courts would interpret the ambiguities and omissions left in a statute.[6] Instead, the legislature could direct the interpretation of its own statutes.

Legislatures have already experimented with this strategy. Hundreds of statutes contain clauses requiring that they be lib-

---

[3] *See, e.g.*, Grayfred B. Gray, *Reducing Unintended Ambiguity in Statutes: An Introduction to Normalization of Statutory Drafting*, 54 TENN. L. REV. 433, 434 (1987).

[4] *See* Shirley S. Abrahamson & Robert L. Hughes, *Shall We Dance? Steps for Legislators and Judges in Statutory Interpretation*, 75 MINN. L. REV. 1045, 1093 (1991) (noting the "generally accepted view" is that there should be more interaction, but remaining unsure as to whether this is a sound notion); Ruth Bader Ginsburg & Peter W. Huber, *The Intercircuit Committee*, 100 HARV. L. REV. 1417 (1987) (advocating "second-look" committees in Congress to revise statutes for clarity and ease of applicability); Joseph F. Weis, Jr., *The Federal Courts Study Committee Begins Its Work*, 21 ST. MARY'S L.J. 15, 21 (1989) (reporting early returns from surveys of members of the federal judiciary, citizen groups, and court administrators suggesting that courts and Congress should exchange information about legislative defects).

[5] *See* Anthony D'Amato, *Can Legislatures Constrain Judicial Interpretation of Statutes?*, 75 VA. L. REV. 561 (1989); Abner J. Mikva, *Reading and Writing Statutes*, 28 S. TEX. L.J. 181, 183 (1986) (suggesting that the problem with drafting statutes and statutory interpretation is often "the unawareness that the legislative branch and the judicial branch have of each other's game rules," and recommending that court opinions interpreting statutes should be distributed to the drafters).

[6] *See* Robert H. Jackson, *The Meaning of Statutes*, 34 A.B.A. J. 535, 538 (1948); Craig W. Palm, Note, *RICO and the Liberal Construction Clause*, 66 CORNELL L. REV. 167, 182 & nn.51–52 (1980).

> [A]s matters stand today, I do not see how Congress can know, even roughly, the effect that will ultimately be given to any language it may use . . . . Though it would not dispel all the doubts which are inherent in the situation, it would help give objectivity to the process of interpretation and assurance to drafting of statutes, if we could have general acceptance by the bench as well as the Bar of a few basic principles of statutory construction.

Jackson, *supra* at 538.

erally or strictly construed.[7] Most legislatures have passed general laws specifying at least a few interpretive principles for the judiciary to follow in all cases or in certain types of cases.[8] Such laws, which I will call interpretive directions, might improve the chances that a judicial interpretation will be consistent with the legislature's intentions. However, there are many problems with interpretive directions. Although the idea is an appealing way to reduce the uncertainty of judicial interpretation, in practice it is not very effective for avoiding unexpected or unwanted interpretations.

In Part I, I define interpretive directions and discuss how legislatures have used them in the past and how legislatures might use them more effectively in the future. In Part II, I consider possible limitations on legislative and congressional power to direct judicial interpretation of statutes. Finally, in Part III, I examine practical obstacles that may prevent interpretive directions from having a significant impact even if the legislative power to pass them is unrestrained.

I. USE OF INTERPRETIVE DIRECTIONS

A. *Definition of Interpretive Directions*

In one sense, every statute can be seen as nothing but interpretive directions: the words the legislature chooses direct how the law will be interpreted and applied, but do not necessarily dictate a single interpretation. For example, the legislature's choice of the word "may" almost always will produce a different interpretation than if it chooses the word "shall," since the words to be interpreted are different.[9] Although words may not always have fixed meanings, at least they make certain interpretations more likely than others.

By "interpretive directions," however, I mean statutory provisions that direct interpretation not by specifying the substance to be interpreted, but by telling courts how to interpret the law. The most common type of interpretive direction tells courts to

---

[7] *See infra* notes 18–26 and accompanying text.

[8] *See infra* notes 27–32 and accompanying text.

[9] In this example, the court may also interpret the words differently because the legislature has defined them differently. *See, e.g.*, NEB. REV. STAT. § 49-802(1) (1988). However, that is not true of most words used in statutes.

resolve ambiguities in favor of certain purposes or parties, such as strict construction of penal statutes to favor the accused or liberal construction to effectuate the statute's purposes. Interpretive directions also may tell courts what sources to consult in resolving ambiguities, such as legislative history or the context of the statute. Similarly, they may list factors that judges may consider in choosing among possible interpretations, such as the consequences of a particular interpretation. An interpretive direction also might tell the courts to abandon a judicial interpretive rule, such as the common law rule of strict construction of penal statutes, in favor of an interpretation according to the "fair import" of the words used.

I do not include in this definition several other types of provisions that often appear in particular statutes or in general code sections on construing statutes. First, most codes have general definitions of terms that apply to all statutes unless otherwise stated.[10] These definition sections are not interpretive directions, because they only describe the substance to be interpreted. Second, many state codes have general provisions that tell courts how to deal with conflicts in statutes.[11] Again, this type of provision directly defines the scope and consequences of various statutes and does not establish interpretive principles.[12] Finally, one type of provision on the borderline of "interpretive directions" is the general presumption. Many states have enacted various presumptions for courts to apply in filling gaps or omissions in a statute.[13] Since these provisions create presumptions rather than default rules, they do guide the court's interpretation in a way. They also have some of the same problems that typical interpretive directions have.[14] Still, they are not true interpretive directions, according to my definition. If such a provision were written into a specific statute, it would be part of the substance

---

[10] *E.g.*, GA. CODE ANN. §§ 102–103 (Michie Supp. 1992); N.M. STAT. ANN. § 12-2-2 (Michie 1988).

[11] *E.g.*, COLO. REV. STAT. ANN. § 2-4-205 (West 1989) (explaining how courts should deal with conflicts between special and general provisions).

[12] However, directing courts to try to harmonize apparently conflicting statutes may be considered an interpretive direction, even though a rule that one type of statute will prevail over another if they cannot be reconciled may not.

[13] *E.g.*, TEX. GOV'T CODE ANN. § 311.022 (West 1988) ("A statute is presumed to be prospective in its operation unless expressly made retrospective.").

[14] For example, they may face similar problems of interpretation. *See* State v. Arellano, 801 S.W.2d 128, 130 (Tex. Ct. App. 1990) (holding that legislative history sufficiently revealed intent to make a statute retroactive despite statute requiring prospectivity unless "expressly made retrospective"); *infra* part III.C.

defining the scope of the statute, because it would no longer have the uncertainty that makes it seem like an interpretive direction. For example, a general provision that presumes prospectivity unless otherwise stated would become definite if included in a specific statute. A specific statute might say that each of its various provisions will be presumed to be prospective unless clearly stated otherwise, but such a clause establishes a substantive default rule, not an interpretive principle.

B. *History of Interpretive Directions*

The first interpretive directions were liberal construction clauses. In the nineteenth century, legislatures began to codify what previously had been left to the common law.[15] Courts resisted this change by invoking canons of restrictive interpretation, such as strict construction of penal statutes and statutes in derogation of the common law.[16] Legislatures tried to stop such restrictive interpretation by specifically directing courts to construe statutes liberally. "[I]t became standard practice in drafting statutes to insert a preamble stating broadly the purpose of the act and to close with a provision declaring that the statute should be liberally construed."[17]

Congress and state legislatures continue to use liberal construction clauses. Probably the best-known liberal construction clause is part of the Racketeer Influenced and Corrupt Organizations Act ("RICO").[18] The RICO liberal construction clause has drawn more attention from courts than other such clauses, probably because RICO is an especially broad and ambiguous statute and because RICO was the first substantive federal criminal statute to contain such a clause.[19] Some states also have included liberal construction clauses in their own RICO stat-

---

[15] *See generally* LAWRENCE M. FRIEDMAN, A HISTORY OF AMERICAN LAW 403–07 (2d ed. 1985) (describing codification movement of 19th century).

[16] *See* G. Robert Blakey, *Foreword*, 65 NOTRE DAME L. REV. 873, 876 n.23 (1990) (citing J. WILLARD HURST, THE GROWTH OF AMERICAN LAW: THE LAW MAKERS 186 (1950)); Roscoe Pound, *Common Law and Legislation*, 21 HARV. L. REV. 383, 386 (1908).

[17] DAVID WIGDOR, ROSCOE POUND: PHILOSOPHER OF LAW 174 (1974).

[18] Organized Crime Control Act of 1970, Pub. L. No. 91-452, 84 Stat. 922 (codified at 18 U.S.C. §§ 1961–1968 (1988)). The RICO liberal construction clause says: "The provisions of this Title shall be liberally construed to effectuate its remedial purposes." *Id.* § 904(a), 84 Stat. 941, 947 (reproduced following 18 U.S.C. § 1961 (1988)).

[19] *See* Russello v. United States, 464 U.S. 16, 27 (1983).

utes.[20] Worker's compensation statutes also typically direct liberal construction in favor of compensating workers.[21] Many other types of state laws include their own liberal construction clauses as well.[22]

Besides including liberal construction clauses in particular statutes, many state legislatures have abolished the common law rules of strict construction of penal statutes[23] and strict construction of statutes in derogation of the common law,[24] often replacing them with general requirements of "fair import" or "liberal" construction.[25] A few states instead have codified the common law canon requiring strict construction of penal statutes.[26]

---

[20] *E.g.*, OR. REV. STAT. § 166.735(2) (1991); R.I. GEN. LAWS § 7-15-10 (1992); *see also* Joseph E. Bauerschmidt, Note, *"Mother of Mercy—Is This the End of RICO?"—Justice Scalia Invites Constitutional Void-for-Vagueness Challenge to RICO "Pattern"*, 65 NOTRE DAME L. REV. 1106, 1147 & n.244 (1990).

[21] *E.g.*, CAL. LAB. CODE § 3202 (West 1988); GA. CODE ANN. § 34-9-150 (Michie 1987); MO. ANN. STAT. § 287.800 (Vernon 1969); WASH. REV. CODE § 51.12.010 (1990) ("This title shall be liberally construed for the purpose of reducing to a minimum the suffering and economic loss arising from injuries and/or death occurring in the course of employment."); *see also* Mark R. Whitmore, Note, *Denying Scholarship Athletes Worker's Compensation: Do Courts Punt Away a Statutory Right?*, 76 IOWA L. REV. 763, 768 (1991) ("[C]ourts and legislatures generally require liberal interpretation of worker's compensation statutes to benefit employees injured while fulfilling their employment agreements."). *But see, e.g.*, FLA. STAT. ANN. § 440.015 (West 1992) ("[I]t is the intent of the Legislature that the facts in a worker's compensation case are not to be interpreted liberally in favor of either the rights of the injured worker or the rights of the employer.").

[22] *See, e.g.*, ALASKA STAT. § 25.23.005 (1990) (liberal construction of adoption code in favor of best interests of adopted children); CAL. CIV. CODE § 82 (West 1993) (liberal construction of fair dealership law to promote underlying purposes); ILL. ANN. STAT. ch. 225, para. 65(2) (Smith-Hurd 1993) (liberal construction of nursing act to best carry out purposes); NEV. REV. STAT. § 30.140 (1992) (liberal construction of declaratory judgments chapter).

[23] *See* ARIZ. REV. STAT. ANN. § 13-104 (1989); CAL. PENAL CODE § 4 (West 1988); DEL. CODE ANN. tit. 11, § 203 (1987); HAW. REV. STAT. § 701-104 (1988); KY. REV. STAT. ANN. § 500.030 (Michie/Bobbs-Merrill 1990); LA. REV. STAT. ANN. § 14:3 (West 1986); MICH. STAT. ANN. § 28.192(2) (Callaghan 1990); MONT. CODE ANN. § 45-1-102(2) (1991); N.H. REV. STAT. ANN. § 625(3) (1986); N.J. STAT. ANN. § 2C:1-2(c) (West 1992); N.Y. PENAL LAW § 5 (McKinney 1988); N.D. CENT. CODE § 29-01-29 (1991); OR. REV. STAT. § 161.025(2) (1991); 18 PA. CONS. STAT. ANN. § 105 (1983); S.D. CODIFIED LAWS ANN. § 22-1-1 (1979); TEX. PENAL CODE ANN. § 1.05(a) (West 1974); UTAH CODE ANN. § 76-1-106 (1990).

[24] *See* ARK. CODE ANN. § 16-55-104(a) (Michie 1987); CAL. CIV. CODE § 4 (West 1988); IDAHO CODE § 73-102(1) (1989); IOWA CODE ANN. § 4.2 (West 1989); KAN. STAT. ANN. § 77-109 (1992); MO. ANN. STAT. § 1.010 (Vernon 1969); MONT. CODE ANN. § 1-2-103 (1991); NEB. REV. STAT. § 25-2218 (1988); OKLA. STAT. ANN. tit. 12, § 2, tit. 25, § 29 (West 1988); 1 PA. CONS. STAT. ANN. § 1928(a) (1983); UTAH CODE ANN. § 68-3-2 (1990).

[25] Some state legislatures have directed liberal construction of all statutes without expressly abrogating either type of strict construction canon. *See, e.g.*, COLO. REV. STAT. ANN. § 2-4-212 (West 1989); ILL. ANN. STAT. ch. 1, para. 1011 (Smith-Hurd 1993).

[26] *See* FLA. STAT. ANN. § 775.021(1) (West 1992); GA. CODE ANN. § 16-1-20(a)

Some states also have adopted interpretive directions telling courts what sources they may consult and what factors they may consider in construing statutes. Most of these directions are permissive, not mandatory, and selective rather than exhaustive.[27] For example, Colorado law permits courts to consider, among other matters, the object of the ambiguous statute, the circumstances of its enactment, the legislative history, the common law and similar laws, the consequences of a given interpretation, administrative constructions, and any legislative declaration of purpose.[28]

However, some states do have general mandatory interpretive directions. Many of these are unremarkable codifications of common law canons of interpretation. Minnesota law tells courts that "[w]hen the words of a law in their application to an existing situation are clear and free from all ambiguity, the letter of the law shall not be disregarded under the pretext of pursuing the spirit."[29] Similarly, the Texas Code states that "[i]n interpreting a statute, a court shall diligently attempt to ascertain legislative intent."[30] A few types of more specific mandatory interpretive directions have been used, such as New Jersey's direction that courts draw no implication or presumption from the arrangement of sections in a statute.[31] These more specific mandatory directions usually prohibit courts from consulting or considering certain things when construing statutes.[32]

C. *Potential Uses for Interpretive Directions*

Legislatures generally have adopted only a few types of interpretive directions. With few exceptions, they have not tried to

---

(Michie 1987); OHIO REV. CODE ANN. § 2901.04(A) (Baldwin 1993); 1 PA. CONS. STAT. ANN. § 1928(b).

[27] *See, e.g.*, COLO. REV. STAT. ANN. § 2-4-203 (West 1989) ("If a statute is ambiguous, the court, in determining the intention of the general assembly, may consider among other matters . . . ."); IOWA CODE ANN. § 4.6 (West 1989); MINN. STAT. ANN. § 645.16 (West 1983) ("When the words of a law are not explicit, the intention of the legislature may be ascertained by considering, among other matters . . . ."); 1 PA. CONS. STAT. ANN. § 1921(c) (1983).

[28] *See* COLO. REV. STAT. ANN. § 2-4-203; *see also* MODEL STATUTORY CONSTRUCTION ACT § 15 & cmt. (1990) (noting few states that have adopted similar lists of permissible extrinsic aids).

[29] MINN. STAT. ANN. § 645.16.

[30] TEX. GOV'T CODE ANN. § 312.005 (West 1988).

[31] N.J. STAT. ANN. § 1:1-5 (West 1992).

[32] *Compare* ALA. CODE § 1-1-14(b) (1982) ("Unless otherwise provided in this Code, the descriptive headings or catchlines immediately preceding or within the text of the

exploit more fully the potential of interpretive directions to influence judicial interpretation of statutes. One reason for this may be that legislatures have seen no need for innovative or aggressive interpretive directions. Legislatures are unlikely to take a great interest in judicial statutory interpretation unless they believe courts persistently frustrate legislative intent, as when judicial obstruction of legislation led to the widespread use of liberal construction clauses.[33] Legislatures also may not have used interpretive directions much because they suspect that they will not make much of a difference anyway. Finally, legislatures often may not realize the significance of the ambiguities in their statutes; when they do, they may be more likely to respond with more specific drafting than with interpretive directions.

Regardless of the reasons, legislatures could make better use of whatever power they have to direct interpretation. Justice Scalia suggested that "[w]hat is of paramount importance is that Congress be able to legislate against a background of clear interpretive rules, so that it may know the effect of the language it adopts."[34] Scalia meant that courts should follow consistent interpretive rules with predictable outcomes. One scholar has even suggested that courts should infer and follow the interpretive rules preferred by the legislature.[35] However, neither the Supreme Court nor the judiciary generally agrees on what interpretive rules to follow. If Congress and state legislatures could dictate interpretive rules by statute, they might achieve this consistency and predictability by themselves faster and on their own terms. They could even set different interpretive rules for different statutes to reflect the variety of statutes and statutory objectives.

---

individual sections of this Code, except the section numbers included in the headings or catchlines immediately preceding the text of such sections, do not constitute part of the law, and shall in no manner limit or expand the construction of any such section.") *with* 1 PA. CONS. STAT. ANN. § 1924 (1983) ("The title and preamble of a statute may be considered in the construction thereof . . . . The headings prefixed to titles, parts, articles, chapters, sections and other divisions of a statute shall not be considered to control but may be used to aid in the construction thereof."). *But see* ILL. ANN. STAT. ch. 815, para. 505/2 (Smith-Hurd 1993) ("In construing this section consideration shall be given to the interpretations of the Federal Trade Commission and the federal courts relating to Section 5(a) of the Federal Trade Commission Act."). Interpretive directions that create presumptions are an exception to this observation, if they are considered interpretive directions at all. *See supra* notes 13–14 and accompanying text.

[33] *See supra* notes 15–17 and accompanying text.

[34] Finley v. United States, 490 U.S. 545, 556 (1989); *see also* Jackson, *supra* note 6, at 537.

[35] Edward O. Correia, *A Legislative Conception of Legislative Supremacy*, 42 CASE W. RES. L. REV. 1129, 1145 (1992).

Legislatures could devise a variety of novel interpretive directions that could have a significant impact on judicial statutory interpretation. For example, legislatures could forbid courts to consult legislative history in determining the meaning of ambiguous statutes. This type of interpretive direction might be useful when a statute is the product of difficult compromises, because it might decrease the chances that a court would upset a compromise reflected in the statute's language by relying upon legislative history that favors one particular side. Legislatures also might prefer complex, technical statutes to be interpreted without reference to legislative history. With such statutes, courts may remain more faithful to the legislature's intentions by enforcing the words in all cases rather than interpreting them in light of legislative history and perceived intent.[36] A legislature also might want to limit the permissible sources of legislative history, rather than relying on courts to weigh items of legislative history, especially when the legislative history contains many conflicting perspectives.[37]

On the other hand, sometimes a legislature might want to require reference to legislative history.[38] A legislature might decide that a broad statute with obviously general and ambiguous provisions would benefit from the use of legislative history in

---

[36] *See* D'Amato, *supra* note 5, at 582 (suggesting that this is the reason for favoring judicial adherence to the text of the tax code). *But see* Bob Jones Univ. v. United States, 461 U.S. 574, 586 (1983) (interpreting charitable tax exemptions by relying on legislative intent rather than on the plain meaning of the statute).

[37] *See* Civil Rights Act of 1991, § 105(b), 105 Stat. 1071, 1075 ("No statements other than the interpretive memorandum appearing at Vol. 137 Congressional Record S 15276 (daily ed. Oct. 25, 1991) shall be considered legislative history of, or relied upon in any way as legislative history in construing or applying, any provision of this Act that relates to Wards Cove—Business necessity/cumulation/alternative business practice."); Maxwell O. Chibundu, *Delinking Disproportionality from Discrimination: Procedural Burdens as Proxy for Substantive Visions*, 23 N.M. L. REV. 87, 158 (1993) (noting that § 105(b) prevents courts from considering many competing definitions of "business necessity" that were inserted in legislative history).

[38] Some legislatures have authorized reference to legislative history, but I have found no statute that requires use of legislative history in some or all situations. *See* COLO. REV. STAT. ANN. § 2-4-203(1) ("If a statute is ambiguous, the court, in determining the intention of the general assembly, may consider . . . (c) The legislative history, if any . . . ."); IND. CODE § 29-1-1-4 (1989) ("The report of the probate code study commission . . . may be consulted by the courts to determine the underlying reasons, purposes and policies of this article, and may be used as a guide in its construction and application."); IOWA CODE ANN. § 4.6.3 ("If a statute is ambiguous, the court . . . may consider . . . [t]he legislative history."); MINN. STAT. ANN. § 645.16 ("When the words of a law are not explicit, the intention of the legislature may be ascertained by considering, among other matters . . . (7) The contemporaneous legislative history."); 1 PA. CONS. STAT. ANN. § 1921(c) ("[T]he intention of the General Assembly may be ascertained by considering . . . (7) The contemporaneous legislative history."); TEX. GOV'T CODE ANN. § 311.023(3) (West 1988).

all cases. Or a legislature might insert such a clause to counter what it perceives as extreme textualism by the courts.

Similarly, state legislatures might require consideration of federal judicial or administrative interpretations of similar federal statutes.[39] Such directions might help keep state statutes in harmony with related federal efforts. They also might take advantage of federal expertise and experience.

Legislatures also might forbid courts to use clear statement rules. These rules are judicial creations requiring explicit statutory text to overcome various presumptions that the legislature did not intend to take particular constitutionally sensitive actions.[40] If legislatures disliked these added burdens or their unexpected retroactive application, they could pass interpretive laws forbidding clear statement rules in some or all areas. On the other hand, legislatures might want to create their own clear statement rules so that courts will not mistakenly decide that the legislature intended to take certain actions.[41]

Another novel use of interpretive directions would be setting standards of interpretation for filling statutory gaps and omissions. Ordinarily, legislatures concerned about this problem have used general presumptions or default rules.[42] However, legislatures might experiment with interpretive standards for courts to follow in implying certain remedies or in filling gaps such as missing limitations periods. Since presumptions and default rules have more predictable effects than interpretive standards, such directions might be useful only where there are many

---

While clauses authorizing the use of legislative history might dampen criticism from opponents of its use, such clauses do not guarantee that judges will use it. Judges can still conclude that there is no need to turn to the history, or that it would do more harm than good. *See* Boykin v. State, 818 S.W.2d 782 (Tex. Crim. App. 1991) (acknowledging statutory permission to consult legislative history but discouraging its use).

[39] *See* ILL. ANN. STAT. ch. 815, para. 505/2 (requiring state courts to consider federal judicial and administrative interpretations of federal statute in interpreting state statute).

[40] *See, e.g.*, Gregory v. Ashcroft, 111 S. Ct. 2395, 2406 (1991) (requiring unambiguous language indicating intent to interfere with state authority to determine the qualifications of important government officials); Pennhurst State School & Hosp. v. Halderman, 451 U.S. 1, 17–18 (1981) (requiring unambiguous statement of intent to attach federal conditions to federal money given to states). *See generally* William N. Eskridge, Jr. & Philip P. Frickey, *Quasi-Constitutional Law: Clear Statement Rules as Constitutional Lawmaking*, 45 VAND. L. REV. 593 (1992) (discussing the Supreme Court's use of clear statement rules).

[41] *See* Religious Freedom Restoration Act of 1993, Pub. L. No. 103-141, § 6(b), 107 Stat. 1488, 1489 ("Federal statutory law adopted after the date of the enactment of this Act is subject to this Act unless such law explicitly excludes such application by reference to this Act.").

[42] *See supra* notes 13–14 and accompanying text.

different preferred interpretations, depending on the circumstances.[43]

II. Legislative Power to Direct Interpretation

Although interpretive directions may have the potential to increase the legislature's control over how its laws are interpreted and applied, the legislature's power to use them may be limited. In this section, I discuss possible constitutional limitations on legislative power to direct how the courts interpret statutes. Then, in the next section, I discuss practical limitations on the impact of interpretive directions.

A. *Separation of Powers*

At least some interpretive directions might violate the principle of separation of powers between the legislative and the judicial branches. The legislature has the power to make the laws, while the judiciary declares what those laws mean.[44] If declaring what laws mean is part of the essence of the judicial function, it would seem that telling the courts how to perform that function would be an impermissible invasion by the legislative branch.[45] If so, judges would not be bound by the efforts of the legislative branch to direct or restrict that process.[46]

---

[43] In such cases, once a legislature became aware of the problem, it might be more likely to apply its standards itself and then deal with the problem by amending statutes specifically rather than adopting interpretive standards. There might be some cases, however, where the legislature would prefer to let the courts apply such standards.

[44] The first Justice Harlan's statement of separation of powers is often quoted: "[The courts] cannot and will not judicially legislate, since [their] function is to declare the law, while it belongs to the legislative department to make the law." Standard Oil Co. v. United States, 221 U.S. 1, 102 (1911) (Harlan, J., concurring in part and dissenting in part).

[45] Justice Scalia has made a similar argument. He maintains that reference to legislative history violates the separation of powers because the courts thereby allow Congress to control the judicial function of interpreting statutes. *See* William N. Eskridge, Jr., *The New Textualism*, 37 UCLA L. Rev. 621, 673 (1990) (footnotes omitted) ("There are several possible twists to Justice Scalia's separation of powers argument. Most simply, it seems to be that after Congress has performed its article I duty of enacting legislation, its views become irrelevant to the very separate inquiry performed by article III judges . . . when they interpret and apply the legislation."); Allison C. Giles, Note, *The Value of Nonlegislators' Contributions to Legislative History*, 79 Geo. L.J. 359, 359 n.3 (1990) (citing Justice Antonin Scalia, Speech delivered at several law schools (1985–1986)).

[46] *Cf.* Henry P. Monaghan, *Stare Decisis and Constitutional Adjudication*, 88 Colum. L. Rev. 723, 754 (1988) ("If stare decisis is part of the judicial power of article III,

Yet very few have argued that a legislature violates the separation of powers when it passes an interpretive direction.[47] Interpretive directions have been used for decades, and many courts have implicitly or explicitly acknowledged their validity.[48] Even so, interpretive directions may impermissibly interfere with the judicial function in some instances.

One such instance may be where interpretive directions attempt to control the application of existing statutes to past acts. Over forty years ago, Jefferson Fordham and Russell Leach declared:

> Any serious suggestion at this day that since interpretation is a judicial function a general interpretive act, applicable only to future statutes, would be unconstitutional, could hardly be taken seriously . . . . [W]e have long proceeded on the basis that, although ultimate interpretation is for the courts, it is

---

it is an inalienable command binding the Court; its demands remain authoritative in the face of competing demands by the other branches. Constitutional common law, however, is subject to congressional control.").

[47] *See* Livingston Hall, *Strict or Liberal Construction of Penal Statutes*, 48 HARV. L. REV. 748, 757 (1935) (asserting that "only one court has doubted the power of the legislature to 'direct the judiciary in the interpretation of existing statutes'") (quoting State v. Parsons, 220 N.W. 328, 331 (Iowa 1928)); *see also* State v. Schlenker, 84 N.W. 698, 698 (Iowa 1900) ("[The legislature] has no power to direct the judiciary in the interpretation of existing statutes."); Commonwealth *ex rel.* Roney v. Warwick, 33 A. 373 (Pa. 1895); Titusville Iron-Works v. Keystone Oil Co., 15 A. 917, 919 (Pa. 1888) (declaring that under the Pennsylvania constitution, "[t]he legislature can no more exercise judicial powers than the courts can arrogate to themselves legislative powers . . . . The act . . . is in no respect a legislative declaration of the rights and privileges of the class of persons to whom it relates, but it is a judicial order or decree directed to the courts").

On the other hand, only a few have argued that interpretive directions are constitutional. *See* Jefferson B. Fordham & J. Russell Leach, *Interpretation of Statutes in Derogation of the Common Law*, 3 VAND. L. REV. 438, 448 (1950); W. David Slawson, *Legislative History and the Need to Bring Statutory Interpretation Under the Rule of Law*, 44 STAN. L. REV. 383, 384 (1992); Palm, *supra* note 6, at 180–81.

[48] *See, e.g.*, State v. Tramble, 695 P.2d 737, 740 (Ariz. 1985) (en banc); Commonwealth v. Trent, 77 S.W. 390, 393 (Ky. 1903) (declaring that a liberal construction clause abrogating strict construction of penal statutes "was within the competency of the Legislature to enact . . . and it is incumbent on the judiciary to enforce it"); State v. Harper, 498 A.2d 310, 313 (N.H. 1985) (declining to follow the rule of strict construction of penal statutes because the legislature had abolished the rule); Publix Asbury Corp. v. City of Asbury Park, 86 A.2d 798, 803 (N.J. Ch. 1951) (following legislative prohibition on drawing implications from location of statutory provisions), *aff'd*, 86 A.2d 806 (N.J. Super. Ct. App. Div. 1952); Turner v. State, 126 P. 452, 455–56 (Okla. Crim. App. 1912) (following "fair import" direction, declaring that "[w]hen the Legislature has made a change in legal procedure, it is the duty of the courts to lay aside their preconceived ideas, and construe such legislation according to its spirit and reason . . . . [M]any judges try to so twist and evade statutes as to enable them to substitute their own private views for regularly enacted statutes"); State v. Hemrich, 161 P. 79, 81 (Wash. 1916).

> within the legislative province to lay down rules of interpretation for the future.[49]

They were not so confident about legislative attempts to control how the judiciary interpreted and applied existing statutes to past acts, however:

> If it purports to control interpretation in a pending case or to give effect to an interpretation different from that employed in a case already decided and does that with respect to the very matter then in litigation it would surely meet with judicial disfavor as an effort to interfere with the independence or finality of judicial decision. Even in the absence of that factor, an interpretive act would not, generally speaking, be given effect by the courts as to past transactions.[50]

Not only might retroactive application interfere with the judicial function, but if an interpretive direction caused a change in the scope of a crime or an increase in punishment, it should be considered an ex post facto law.[51] Even if an interpretive direction did not violate the ex post facto prohibition, it should be treated like new substantive laws and presumed to apply only prospectively.[52]

It also might be constitutionally impermissible for legislatures to go beyond specifying the appropriate guiding principles or interpretive attitude (strict or liberal construction) to insisting that courts consult or not consult certain sources. Most mandatory interpretive clauses to date have been strict or liberal construction clauses. These intrude less on the judicial function because they are essentially just evidence of legislative intent. Since the meaning of words can vary with the intention of the speaker, the legislature is simply making clear what the law itself means. In fact, "shall be liberally construed" may be interpreted to mean simply that the legislature intended the ambiguous provisions of the statute to have the meaning that best

---

[49] Fordham & Leach, *supra* note 47, at 448 (footnotes omitted).

[50] *Id.* at 449 (footnotes omitted).

[51] *See* Collins v. Youngblood, 497 U.S. 37, 46 (1990) (explaining that the ex post facto clause prohibits laws, whatever their form and however subtle, that retroactively alter criminal offenses or increase their punishment).

[52] *See* Bowen v. Georgetown Univ. Hosp., 488 U.S. 204, 208 (1988) ("Retroactivity is not favored in the law. Thus, congressional enactments and administrative rules will not be construed to have retroactive effect unless their language requires this result."); United States v. Security Indus. Bank, 459 U.S. 70, 79 (1982) ("The principle that statutes operate only prospectively, while judicial decisions operate retrospectively, is familiar to every law student.").

furthers the declared purposes.[53] In other words, liberal construction clauses may be seen as just another way to indicate within the statute itself the meaning of the statute's ambiguous provisions, rather than as a command to the judiciary to go about its job in a particular way.[54] Many courts have treated liberal construction clauses as nothing more than evidence of legislative intent, although they have not explicitly offered this reasoning as a justification.[55]

Although this is only one possible interpretation of liberal construction clauses, it does highlight the difference between this type of interpretive clause and clauses specifying what sources courts may consider in construing a statute. In the latter case, the legislature is not telling courts in any way what it intended the language of the statute to mean. Rather, it is plainly telling the courts how to go about determining the meaning of the statute. Clauses forbidding or requiring reference to certain sources in interpreting statutes thus intrude more on the judicial function.

One weakness in this distinction between these two general types of directives is that it also suggests that general interpretive statutes might impermissibly invade the judicial function, even though such statutes have not been challenged any more than statute-specific construction clauses. It makes more sense to think of an interpretive clause within a specific statute as an

---

[53] *Cf.* G. Robert Blakey, *The RICO Civil Fraud Action in Context: Reflections on* Bennett v. Berg, 58 NOTRE DAME L. REV. 237, 290 n.150 (1982) (suggesting that the liberal construction clause should prevail over the common law rule of strict construction of penal statutes because the liberal construction clause is "a basis for ascertaining legislative intent," not just a rule of construction).

[54] Interpreting liberal construction clauses in this way leaves courts free to follow whatever course of interpretation they prefer, although they certainly must weigh the clause as an indication of the intended breadth of ambiguous provisions.

[55] *See, e.g.*, Russello v. United States, 464 U.S. 16, 26–27 (1983) (citing RICO liberal construction clause along with legislative history as evidence of broad legislative intent); *infra* part III.D.3.

Another way to justify traditional construction clauses while maintaining that legislatures cannot interfere with the judicial interpretive function is to argue that traditional construction clauses are really just directions to courts not to frustrate the legislative will.

> [S]uch provisions can be rationalized as requests to the courts to stop subverting normal meanings under the guise of "interpretation." Rather than a legislative intrusion into the province of the courts, it may merely be an attempt, in the form of a rule of law, to keep the courts from paying insufficient deference to its pronouncements in the field of policy making, and thus from usurping the functions of the legislature.

Palm, *supra* note 6, at 180 n.46 (quoting CHARLES B. NUTTING ET AL., LEGISLATION 397 (1969)).

internal indication of what an ambiguous provision was intended to mean. A general liberal construction clause applied to all legislative acts seems less like a convenient way to specify the intended meaning of ambiguities, since the interpretive statute applies to ambiguities in statutes that do not even exist at the time the interpretive statute is created.

However, there is still a significant difference between generally applicable strict or liberal construction clauses and interpretive directions that control what the courts may consult when interpreting statutes. Even when prospective, construction clauses tell the courts which of any two interpretations the legislature has chosen to prefer, such as the interpretation that furthers the statute's purposes. The legislative branch is free to decide and declare what interpretations it intends. But directions controlling source material do not tell the courts that the legislature has chosen to prefer a certain interpretation when the text is ambiguous. Rather, they simply tell the courts how to go about deciding what the statute means, without any indication of a preferred interpretation. Without any connection to choosing or explaining the meaning of the law, which is within the legislative power, interpretive directions requiring or forbidding consideration of particular extra-legislative sources probably unconstitutionally infringe on the judicial function.[56] However, legislatures could accomplish much the same thing in two other ways that would not raise the constitutional issue. First, they could explicitly incorporate the extra-legislative sources into the statute, making them part of the law rather than merely mandatory guides to interpretation. Second, they could declare those sources to be the only indication of their intent outside of the statute itself, in which case courts would not be bound to consider them but almost certainly would.

---

[56] *See* Fitzgerald v. Chicago Title & Trust Co., 361 N.E.2d 94, 96 (Ill. App. Ct. 1977) (noting that legislative direction to consider federal interpretations of federal act in construing state act was beyond legislature's authority), *aff'd*, 380 N.E.2d 790 (Ill. 1978); Ira C. Lupu, *Statutes Revolving in Constitutional Law Orbits*, 79 VA. L. REV. 1, 11 n.47 (1993) ("Such a limitation [on sources of legislative history that courts can consult] arguably violates separation of powers principles.").

> Nothing said or done in the legislative process *before* the statute is enacted can be fairly said to invade judicial functions, so long as judges are free to consider all the evidence of statutory meaning. If judges for institutional competence reasons choose to look at committee reports to inform their interpretation of statutes, this choice seems to be one they can make, and their liberty to make such choices is guaranteed, rather than precluded, by judicial independence and separation of powers precepts.

Eskridge, *supra* note 45, at 674.

B. *Bicameralism and Presentment*

The requirements of bicameralism and presentment[57] also may limit the types of interpretive directions that are within the legislative power. A clause requiring reference to legislative history in determining the meaning of a statute would implicate this limitation most clearly. Such a clause might be considered an attempt to elevate legislative history into law without satisfying the required constitutional procedures of passing both houses and being presented to the President. Although the clause requiring reference to legislative history would be properly approved, the legislative history that eventually would be incorporated into law would not have been properly approved.

Of course, the judiciary itself can use legislative history without violating these constitutional principles. When courts consult legislative history on their own initiative, they are simply trying to determine the meaning of the text by discovering the legislative intent reflected in the history. But if Congress insists that courts refer to legislative history, it is not merely confirming the judiciary's recognized power of interpretation. In effect, it is trying to modify the meaning of the statute by extra-legislative means.

Most interpretive directions would not be subject to this objection, because it only applies to directions that require courts to consider non-statutory materials as authoritative sources of meaning. Even then, courts probably would not invalidate an interpretive direction so long as it left courts to ultimately decide what the statute means. Furthermore, if all the specified legislative history existed when the law was passed and signed, the legislative history might have been properly approved because Congress and the President at least knew exactly what courts would be required to consider, even though they may not have known how that consideration would affect the interpretation and application of the law. The case where the argument would be strongest—where Congress declares a non-statutory source that has yet to be written, such as future committee pronouncements, to be binding law—would probably never come up.[58]

---

[57] U.S. CONST. art. I, §§ 1, 7; *see also* Immigration & Naturalization Serv. v. Chadha, 462 U.S. 919 (1983).

[58] Of course, delegations of authority to administrative agencies do just this. But

C. *Delegation of Law-Making Authority*

A related objection to interpretive directions is that the legislature cannot delegate its lawmaking authority to the courts. This non-delegation principle might impose similar constraints on how much the legislature can rely on interpretive directions instead of substantive statutory text to implement its will.[59]

An extreme example that probably would violate the non-delegation principle would be a general statement of the purpose of a statute accompanied by a direction to the courts to infer whatever remedies, procedures, and details are necessary to implement that general purpose. In that case, the legislature would seem to be delegating its lawmaking function to the courts, because courts would make all of the decisions about how to implement an expressed statutory purpose.[60]

A legislature probably would never pass such a statute, of course. But legislatures might want to delegate decisions on certain parts of a statutory scheme, such as remedies or politically sensitive issues.[61] Remaining silent on these issues might not be as objectionable as directing the courts to make certain decisions, however. For example, it would be more questionable if Congress explicitly directed the federal courts to apply the *Cort v. Ash*[62] standard or any other criteria in deciding whether to imply remedies under federal statutes.

Past experience suggests that the non-delegation principle is unlikely to limit the use of interpretive directions, however. Con-

---

where Congress delegates authority to a nonlegislative body, the question is treated as one of delegation, not as one of satisfying the requirements of bicameralism and presentment. *See infra* part II.C.

[59] *See* Donald H. Zeigler, *Rights Require Remedies: A New Approach to the Enforcement of Rights in the Federal Courts*, 38 HASTINGS L.J. 665, 724 (1987). Zeigler raises a similar question, but concludes that such delegation is not improper:

> While courts may properly decide the scope of their common-law powers themselves, Congress might be overstepping its authority in directing the courts to use those powers and in telling them what the scope of those powers should be. Congress plainly can affect the interpretive process by defining its terms, but directing the federal courts to make law goes much further.

Id. (footnote omitted).

[60] *See* James C. Quarles, *Some Statutory Construction Problems and Approaches in Criminal Law*, 3 VAND. L. REV. 531, 534 (1950) ("If a court should by interpretation or construction give vitality to a meaningless combination of words, it would undoubtedly be legislating and its action would be obnoxious to general principles of government in this country. But when a statute is ambiguous, interpretation is necessary.").

[61] *See* Correia, *supra* note 35, at 1146.

[62] 422 U.S. 66 (1975).

gress has already delegated to the courts broad power to make substantive law without being successfully challenged.[63] Such broad delegations to the courts are seen as simply authorizing the courts to create common law.[64] If anything, providing interpretive standards to guide the courts in creating common law would seem to be even less offensive to the constitutional separation of powers.[65]

III. Practical Limitations on the Impact of Interpretive Directions

Although there are few if any constitutional limitations on the legislative power to pass interpretive directions, there are many practical limitations on the potential of such statutes to influence judicial interpretation. These include other interpretive principles, the necessity of interpreting the interpretive statute itself, and the text of the statute. Perhaps the greatest practical limitation, however, is courts' ignorance of or unwillingness to follow interpretive directions from the legislature.

A. *Other Interpretive Principles*

Other interpretive principles applied by the courts often may limit the effect of interpretive directions. Virtually all interpretive directions apply only when a statute is ambiguous and needs to be interpreted. A court may conclude that it must apply other interpretive principles before deciding whether a statute is ambiguous and therefore subject to the legislature's interpretive direction, thus minimizing the impact of the interpretive direc-

---

[63] *See* Merrill Lynch, Pierce, Fenner & Smith, Inc. v. Curran, 456 U.S. 353, 375–76 (1982) (observing that "judicial recognition of an implied private remedy" does not violate separation of powers); Textile Workers Union v. Lincoln Mills, 353 U.S. 448 (1957) (holding that the Labor Management Relations Act of 1947 authorized federal courts to create federal common law governing labor contracts in industries affecting interstate commerce); Zeigler, *supra* note 59, at 725 ("If Congress can grant the federal courts broad power to create substantive law, it follows that Congress also can grant the courts the narrower power to create remedies to effectuate congressionally created substantive rights and duties.").

[64] *See Textile Workers*, 353 U.S. 448; *see also* Correia, *supra* note 35, at 1148 n.64 (observing that congressional power to ask courts to assist in making policy is well-established).

[65] *See* Hall, *supra* note 47, at 757 (arguing that liberal construction clauses are not "unconstitutional delegation[s] of legislative power to the judiciary" because "[l]iberal construction does not involve going beyond the intention of the legislature").

tion. A court also may admit that the text is ambiguous, but contend that other traditional tools of statutory interpretation should be applied before turning to legislative directions.

1. Substantive Gaps

Courts are especially likely to first apply judicially created interpretive standards for filling specific types of substantive gaps in statutes. For example, the court in *Kaushal v. State Bank of India*[66] ignored or overlooked the RICO liberal construction clause in deciding whether the statute provides equitable remedies to private plaintiffs. It reasoned that because RICO does not explicitly give equitable remedies to private plaintiffs, it was bound to apply the Supreme Court's stricter principles of construction for implying such remedies.[67]

The same issue would arise in other cases where the courts have established standards of interpretation for filling gaps in statutes. For example, RICO does not detail the accrual rules to apply to the relevant statutes of limitation. Courts might apply more liberal accrual rules in accordance with the liberal construction clause, since more liberal accrual rules would make the statute's remedies more available, or they might apply general interpretive principles of statutes of limitations that would require stricter accrual rules and thereby impose greater limits on plaintiffs' opportunities to seek redress. Because the RICO liberal construction clause does not clearly supersede such established judicial rules for dealing with particular gaps, courts probably should apply the more specific judicial rules.[68]

2. Constitutionally Based Canons

Interpretive canons based on constitutional concerns also might take priority over legislative interpretive directions. A

---

[66] 556 F. Supp. 576 (N.D. Ill. 1983).

[67] *Id.* at 584 (reasoning that "the Supreme Court's guidelines for directing [the inference of remedies] provide a mandate, not only a 'satisfactory rationale,'" for not following the liberal construction clause and inferring equitable remedies); *see also* Donald R. Lee, Note, *The Availability of Equitable Relief in Civil Causes of Action in RICO*, 59 NOTRE DAME L. REV. 945, 969 (1984) (criticizing this aspect of the reasoning in *Kaushal*, arguing that the court should have followed the liberal construction clause and implied the equitable remedies to further the act's purposes).

[68] *See* Paul B. O'Neill, Note, *"Mother of Mercy, Is This the Beginning of RICO?": The Proper Point of Accrual of a Private Civil RICO Action*, 65 N.Y.U. L. REV. 172, 238–39 (1990) (arguing that courts should not interpret the liberal construction clause

good example is the rule of lenity, which some judges have applied in spite of direct conflict with a liberal construction clause.[69] These judges and the commentators who agree with them reason that the due process principle of notice requires courts to construe ambiguities in favor of accused criminals. If that is so, interpretive directions that conflict with the rule of lenity, such as liberal construction clauses, would be unconstitutional.[70] If Congress did pass a conflicting interpretive direction, it would be ignored in favor of the rule of lenity.[71] Even though the federal constitution does not mandate it,[72] many courts will consider the rule of lenity and similar interpretive

---

to justify "any conceivable expansion of RICO liability, even one that contravenes a basic principle of statutes of limitation").

[69] *See* United States v. Anderson, 626 F.2d 1358, 1369–70 (8th Cir. 1980), *cert. denied*, 450 U.S. 912 (1981); United States v. Altese, 542 F.2d 104, 107 (2d Cir. 1976) (Van Graafeiland, J., dissenting) (arguing that, notwithstanding the liberal construction clause, courts should apply the harsher criminal penalty only when language is "clear Congressional direction"), *cert. denied*, 429 U.S. 1039 (1977); United States v. Mandel, 415 F. Supp. 997, 1022 (D. Md. 1976) ("While Congress may instruct courts to give broad interpretations to civil provisions, it cannot require courts to abandon the traditional canon of interpretation that ambiguities in criminal statutes are to be construed in favor of leniency."), *aff'd in part and vacated and remanded in part*, 591 F.2d 1347 (4th Cir.), *aff'd by equally divided court en banc*, 602 F.2d 653 (4th Cir. 1979), *cert. denied*, 445 U.S. 961 (1980); State v. Rackle, 523 P.2d 299, 302 (Haw. 1974) (ignoring statutory abrogation of strict construction and declaring that penal statutes must be strictly construed in favor of defendant); Watkins v. Wattle, 558 S.W.2d 705, 711 (Mo. Ct. App. 1977) (applying common law rule of strict construction despite legislative abrogation of the rule).

[70] *See* United States v. Grzywacz, 603 F.2d 682, 692 (7th Cir. 1979) (Swygert, J., dissenting), *cert. denied*, 446 U.S. 935 (1980); *Mandel*, 415 F. Supp. at 1022; *Report of the Ad Hoc Civil RICO Task Force*, 1985 A.B.A. SEC. CORP., BANKING & BUS. L. 13 ("If the liberal construction clause is applicable to determine the scope of criminal liability . . . the provision is unconstitutional."); Barry Tarlow, *RICO Revisited*, 17 GA. L. REV. 291, 309 (1983).

> Any application of the liberal construction clause to the RICO criminal provisions is . . . a violation of due process. The fundamental principle of statutory construction is the due process requirement that criminal statutes be construed in favor of lenity. If the liberal construction clause is applicable to determine the scope of criminal liability under Title IX, the provision is therefore unconstitutional.

Id.

[71] *But see* Blakey, *supra* note 53, at 290 n.150.

> The rule of lenity is a rule of construction that says that the interpretation more favorable to the defendant ought to be adopted when the text of the statute or its legislative history cannot be used to resolve an ambiguity according to congressional intent. The liberal construction clause, however, is a rule of statutory construction based on the text of the statute itself that provides a basis for ascertaining legislative intent. As such, it ought to govern.

Id. (citing Brown v. Barry, 3 U.S. (3 Dall.) 365, 367 (1797) ("[T]he intention of the Legislature, when discovered, must prevail [over] any rule of construction.")).

[72] *Id.* ("[T]he rule of strict construction has never been accorded independent constitutional status . . . . Due process values are at stake in the construction of a statute, but they are met when liberal construction is mandated by the legislature."); David L.

principles to take precedence over any legislative interpretive direction. Again, the propriety of giving precedence to constitutionally based canons depends on whether the legislature clearly meant the interpretive direction to take precedence. If it did, a court should acquiesce unless it holds that the result would be unconstitutional or that the legislature does not have authority to impose the interpretive direction on the courts.

3. Clear Statement Rules and Avoiding Constitutional Questions

More often, courts would properly apply clear statement rules and the canon of avoiding constitutionally sensitive areas before applying an interpretive direction.

A clear statement rule would prevent courts from applying interpretive directions in ways that would raise the constitutional concerns the rule was designed to avoid. If such concerns are great enough to reject the interpretation of the substantive text that otherwise would prevail, they are certainly great enough to deny the application of an interpretive direction when resolving an ambiguity in the substantive text. By their very nature, interpretive directions telling courts how to resolve ambiguities will never indicate the legislature's intention to enter specific constitutionally sensitive areas explicitly enough to satisfy a clear statement rule. Only making the substantive text unambiguous will do.[73]

Similarly, an interpretive direction would not satisfy the principle of construing statutes so as to avoid constitutional questions. That principle requires that if a given interpretation would raise constitutional questions, the court should apply it only if the legislature "clearly expressed" its "affirmative intention" to do so,[74] as long as another interpretation is "fairly possible."[75] If a statute is ambiguous, the mere fact that the legislature indicated a preference for resolving the ambiguity in one way would

---

DaCosta, Note, *Liberal Construction of Title IX of the Organized Crime Control Act of 1970: Section 904(a) v. the Doctrine of Strict Construction*, 12 RUTGERS L.J. 69, 71 n.20 (1980).

[73] A legislature might get around this obstacle by declaring that no clear statement rule may be applied to the act or its interpretation. If courts accept the legislative power to direct interpretation, they would have to comply. The Constitution does not require clear statement rules, since they are based not on legislative inability to intrude on the sensitive areas, but on a presumption that the legislature will not do so without saying so clearly. @FNOT_IND = [74] NLRB v. Catholic Bishop, 440 U.S. 490, 506 (1979).

[75] *Id.* at 510 (Brennan, J., dissenting); Machinists v. Street, 367 U.S. 740, 749–50 (1961); Crowell v. Benson, 285 U.S. 22, 62 (1932).

not be a clear expression of an affirmative intention to take the constitutionally questionable course.

Interpretive directions controlling the materials that judges consult when construing statutes raise a more difficult question. For example, if a particular statute forbade reference to legislative history, and included a liberal construction clause, a court would have to decide whether to consult the legislative history to make an otherwise unreasonable interpretation "fairly possible," rather than follow the unclear and constitutionally questionable interpretation. The court would have to stick with the text if it acknowledged legislative power to restrict what the courts can consider. On the other hand, if a legislature required reference to legislative history before deciding whether the statute was ambiguous, and if the text were insufficiently clear, a court would have to decide whether to accept a combination of text and legislative history in making a clear statement of intent. Again, if legislatures have the power to tell courts that a statute is not ambiguous if its legislative history makes it clear, the court would have to accept a clear statement in the legislative history.

4. More Specific Interpretive Principles

The first three types of conflicting interpretive principles suggest a fourth category that courts sometimes should apply before turning to statutory interpretive directions: more specific canons or standards. Courts often would apply judicial canons of interpretation before turning to the legislative direction if the judicial canons were more specifically directed at the particular type of ambiguity or interpretive problem at hand.

This practice would be justified by the principle that "general and specific provisions, in apparent contradiction, whether in the same or different statutes, and without regard to priority of enactment, may subsist together, the specific qualifying and supplying exceptions to the general."[76] This canon is based on the presumption that since the legislature does not intend to contra-

[76] Townsend v. Little, 109 U.S. 504, 512 (1883); *see also* Smith v. Robinson, 468 U.S. 992, 1024 (1984) (following "the familiar principle of statutory construction that conflicting statutes should be interpreted so as to give effect to each but to allow a later enacted, more specific statute to amend an earlier, more general statute only to the extent of the repugnancy between the two statutes"); United States v. Shewmaker, 936 F.2d 1124, 1128 (10th Cir. 1991) ("When general and specific statutory provisions apparently contradict, it is well-established that the two may exist together, the specific provision qualifying or limiting the general.").

dict itself or to include meaningless provisions, every part of a statute should be given effect if possible.[77] This can be done by modifying general passages to accommodate apparently conflicting but more specific passages.

Courts might treat interpretive directions similarly. When a legislature passes a general liberal construction clause, for example, courts may presume that the legislature does not intend to supersede judicial canons that would address specific types of interpretive questions. Instead, they intend for them to coexist, which is possible if the more general direction is interpreted to apply only if the statute is still ambiguous after applying traditional canons of statutory interpretation.

Some of the courts that have applied the rule of lenity rather than RICO's liberal construction clause seem to have implicitly followed such reasoning. Rather than declaring that the rule of lenity is constitutionally required, they have reasoned that Congress must not have meant for the liberal construction clause to defeat the rule of lenity in the particular situations where the rule of lenity applies, especially because the rule is at least related to constitutional concerns.[78]

This principle would severely restrict the legislature's power to influence interpretation by means of interpretive directions. After all, there is a judicially created interpretive standard or canon available for almost any type of ambiguity. If courts followed this principle, the legislature would have two choices if it really wanted to control interpretation: it could explicitly declare that its general interpretive direction takes precedence over some or all interpretive canons and standards; or it could pass specific interpretive clauses to deal with specific types of ambiguities. This second option would defeat the purpose of general interpretive clauses, because it would force the legislature to

---

[77] *See* Muller v. Lujan, 928 F.2d 207, 211 (6th Cir. 1991).

> [W]hen two statutes conflict to some degree they should be read together to give effect to each if that can be done without damage to their sense and purpose . . . . Where there is no clear intention otherwise, a specific statute will not be controlled or nullified by a general one, regardless of the priority of enactment.

Id. (citation omitted); *see also* Heckler v. Chaney, 470 U.S. 821 (1985) (interpreting apparently conflicting statutory provisions in a compatible way in order to give effect to both); United States v. Menasche, 348 U.S. 528, 538–39 (1955).

[78] *See* United States v. Anderson, 626 F.2d 1358, 1369–70 (8th Cir. 1980), *cert. denied*, 450 U.S. 912 (1981); United States v. Grzywacz, 603 F.2d 682, 692 (7th Cir. 1979) (Swygert, J., dissenting), *cert. denied*, 446 U.S. 935 (1980); United States v. Davis, 576 F.2d 1065, 1069–70 (3d Cir.) (Aldisert, J., concurring), *cert. denied*, 439 U.S. 836 (1978).

specifically consider each possible ambiguity almost as much as if it had removed the ambiguity in the first place. The first option is more realistic, but it would risk subordinating interpretive standards that the legislature would have expected or preferred to have been applied first. This would be especially likely with interpretive standards for filling statutory gaps or constitutionally related canons.

Unless legislatures declare otherwise, courts generally will have the power to decide where interpretive directions fit in the hierarchy of interpretive tools and principles.[79] Courts may even apply general common law interpretive canons before interpretive directions, on the theory that interpretive directions only apply after traditional tools of statutory interpretation have failed to resolve ambiguities.[80] This gives courts considerable freedom to avoid compliance with an interpretive direction, and reduces whatever influence interpretive directions might have. The way to achieve greater control through interpretive directions is to write directions that tell courts where the directions belong in the hierarchy and that are as specific as possible about what courts should do.

B. *Interpreting Interpretive Directions*

Another limitation on the impact of interpretive directions is the necessity of interpreting the direction itself. Like substantive law, interpretive directions require interpretation before judges can apply them. The necessity of interpretation further limits the constraining influence of interpretive directions and makes their effects less certain.

1. Ambiguities in Interpretive Directions

Interpretive directions inevitably will contain some ambiguity. This is especially true of strict construction, liberal construction,

---

[79] *See generally* CASS R. SUNSTEIN, AFTER THE RIGHTS REVOLUTION: RECONCEIVING THE REGULATORY STATE 157–59 (1990); Cass R. Sunstein, Interpreting Statutes in the Regulatory State, 103 HARV. L. REV. 405, 451–62 (1989).

[80] *See* Religious Technology Ctr. v. Wollersheim, 796 F.2d 1076, 1083–84 (9th Cir. 1986) (admitting that construction of act permitting injunctive relief to private plaintiffs in civil RICO actions was "certainly a plausible reading," but concluding on the basis of legislative history that such injunctive relief is not available), *cert. denied*, 479 U.S. 1103 (1987); *Grzywacz*, 603 F.2d at 690–92 (Swygert, J., dissenting) (consulting

"fair import," and other such clauses.[81] The RICO liberal construction clause illustrates well how many questions an interpretive direction can raise, and how such ambiguities can make an interpretive direction's effect hard to predict.

The RICO clause says that the statute "shall be liberally construed to effectuate its remedial purposes."[82] The RICO statute itself describes its purposes,[83] but some have suggested that Congress intended the clause to refer only to its civil remedial purposes, leaving criminal remedies to be strictly construed in accordance with the rule of lenity.[84] Judges also have refused to construe RICO liberally where they did not believe the apparently "liberal" interpretation would actually further the statute's stated purposes, while it might adversely affect other social goals with which RICO is not primarily concerned.[85]

Courts also might infer from the legislative history and the structure of the act certain purposes that are not expressed in the Statement of Findings and Purpose, since the statute itself does

---

legislative history and applying ejusdem generis and rule of lenity rather than applying liberal construction clause).

[81] The Texas presumption of prospectivity "unless expressly made retrospective," TEX. GOV'T CODE ANN. § 311.022, provides an example of the ambiguity that can arise under even relatively concrete directions to courts. Although the code provision seems to require a statutory statement that a law will be retrospective, one Texas court held that legislative history could be considered as evidence of legislative intent to give retroactive effect to a law. *See* State v. Arellano, 801 S.W.2d 128, 130 (Tex. Ct. App. 1990).

[82] Organized Crime Control Act of 1970, Pub. L. No. 91-452, § 904(a), 84 Stat. 941, 947 (reproduced following 18 U.S.C. § 1961 (1988)).

[83] The Statement of Findings and Purpose covering RICO declares: "It is the purpose of this Act to seek the eradication of organized crime in the United States by strengthening the legal tools in the evidence-gathering process, by establishing new penal prohibitions, and by providing enhanced sanctions and new remedies to deal with the unlawful activities of those engaged in organized crime." Pub. L. No. 91-452, § 1, 84 Stat. 922, 923 (1970) (reproduced but not codified following 18 U.S.C. § 1961 (1988)). The Senate Committee on the Judiciary said that RICO seeks to end "the infiltration of organized crime and racketeering into legitimate organizations." S. REP. No. 617, 91st Cong., 1st Sess. 76 (1969).

[84] *See Grzywacz*, 603 F.2d at 692 (Swygert, J., dissenting) ("It is unclear whether Congress intended its directive to apply to those sections which establish criminal liability or merely to the 'remedial' provisions of Title IX."); Barry Tarlow, *RICO: The New Darling of the Prosecutor's Nursery*, 49 FORDHAM L. REV. 165, 178 (1980) ("Significantly, the language of this clause permits liberal construction only to effectuate 'remedial purposes'; it does not mandate that RICO be liberally construed to determine what conduct constitutes a violation of Title IX.").

[85] *See* United States v. Anderson, 626 F.2d 1358, 1370 (8th Cir. 1980) ("[A]n expansive definition of the term 'enterprise' will not necessarily effectuate the remedial purpose of RICO to eliminate the infiltration of racketeers into our business economy. Instead, it will inject federal prosecutors into the realm of offenses traditionally considered to be of a local character." (citation omitted)), *cert. denied*, 450 U.S. 912 (1981).

not explicitly define "remedial purposes."[86] For example, one commentator has suggested that RICO provisions should be construed liberally not only to combat organized crime, but also to protect its victims.[87] If courts do not restrict themselves to furthering statutorily declared purposes, they might conceive of many different purposes that could have perverse effects. Most statutes can be thought to have negative purposes in addition to their overt positive purposes. Criminal statutes, for instance, implicitly intend to preserve individual rights and not to overdeter or over-punish. At the extreme, a court might base the directed interpretation on such a negative purpose.[88] Even if courts do not go that far, unexpected applications based on unintended purposes are likely, as are unexpected refusals to apply the interpretive direction because a court concludes that the legislature intended liberal construction only in favor of certain purposes but not others.

"Liberally construed" is also ambiguous. It may mean that if the text of the statute can have two or more reasonable meanings, the meaning which most furthers the remedial purposes must be followed without further inquiry.[89] But there are other

---

[86] At least one observer has even argued that the liberal construction clause itself requires reference to legislative history in order to determine the statute's remedial purposes. "[C]onstruction is to be liberal in order not to suit all potential plaintiffs, but 'to effectuate [the Act's] remedial purposes.' To ascertain the remedial purposes of the Act, one must go behind its language to discover what Congress' purposes were, as expressed during the creation of the Act." Andrew P. Bridges, *Private RICO Litigation Based Upon "Fraud in the Sale of Securities"*, 18 GA. L. REV. 43, 57 & n.92 (1983) (footnotes omitted) (arguing that "to effectuate its remedial purposes" itself calls for reference to legislative intent and history); *see also* United States v. Davis, 576 F.2d 1065, 1071 (3d Cir.) (Aldisert, J., concurring) ("[The RICO construction clause] must be read in light of the language of the statute and the legislative history. And in viewing the legislative purpose, I detect nothing that precludes the application of the rule of narrow construction of penal statutes."), *cert. denied*, 439 U.S. 836 (1978); DaCosta, *supra* note 72, at 97 (arguing that liberal construction of RICO should be based on "the remedial purposes of the statute as expressed on its face and in the legislative history").

[87] Michael C. Liebman, Note, *Governmental Civil RICO Actions and Labor Unions: Reorganization and Innocent Persons*, 58 GEO. WASH. L. REV. 125, 140 (1989) (suggesting that applying the innocent persons clause in RICO would advance the act's remedial purpose of protecting the victims of organized crime).

[88] *See* Moragne v. States Marine Lines, Inc., 398 U.S. 375, 392 (1970) (acknowledging the possibility that the legislature will recognize "other, conflicting interests" that may form the boundaries of an act's intended effect); *Davis*, 576 F.2d at 1071 (Aldisert, J., concurring) (arguing that "congressional acquiescence in the Justice Department's request to narrow the definition of state offenses" under RICO indicated that applying the rule of strict construction of penal statutes "effectuates, rather than defeats, the obvious legislative purpose").

[89] *See* Blakey, *supra* note 53, at 290 n.150 ("[I]f RICO's language is plain, it ought to control; if the language is ambiguous, that construction which would 'effectuate its

reasonable interpretations. The clause may mean that after a court applies regular interpretive methods, including reference to context, legislative history, and interpretive canons, it must resolve any remaining uncertainties in favor of the statute's remedial purposes.[90] Some also have questioned whether "liberally construed" means that every ambiguity must be resolved in the most liberal way that is reasonable, or whether there are limits to the command.[91] For example, if a purpose is fully "effectuated" without liberal resolution of an ambiguity, then the liberal construction clause might not require interpreting the ambiguous provision to favor the remedial purposes.[92] Although overenforcement may indeed frustrate the statute's purposes, the liberal construction clause does not seem to limit its own application. It does not describe some optimum level of fulfillment of its purposes beyond which the purposes need not be favored. However, the common law rule of lenity has been interpreted in

---

remedial purposes' 'by providing enhanced sanctions and new remedies' ought to be adopted."); G. Robert Blakey & Scott D. Cessar, *Equitable Relief Under Civil RICO: Reflections on* Religious Technology Center v. Wollersheim*: Will Civil RICO Be Effective Only Against White-Collar Crime?*, 62 NOTRE DAME L. REV. 526, 548–49, 552 (1987) (arguing that courts should not consult legislative history in resolving ambiguities in RICO because the liberal construction clause requires a court, once it finds ambiguity, to adopt interpretation furthering remedial purposes); Lee, *supra* note 67, at 953 ("Given the Liberal Construction Clause, the question of how this interpretation [that section 1964(a) allows a private plaintiff to seek equitable remedies] follows from the text is largely a matter of indifference. If the text is plain, the remedy is there; if the text is ambiguous, the ambiguity should be resolved in favor of enhancing the remedial purpose of RICO.").

[90] *See* Kaushal v. State Bank of India, 556 F. Supp. 576, 583 (N.D. Ill. 1983) (despite liberal construction clause, declaring that "[t]o the extent the statute may be viewed as ambiguous, this Court must turn to the legislative history to divine legislative intent"). This again raises the question of where interpretive directions fit in the interpretive hierarchy. *See supra* part III.A. But here the question is not where in the hierarchy courts would put them if it were left to their own judgment, but whether the legislature itself has implied some position in the hierarchy by its choice of words.

[91] *See* DaCosta, *supra* note 72, at 77 ("The clause, however, is not an invitation to the courts to expand the scope of RICO to the outer limits of imagination; rather, it is an instruction to the courts not to narrowly construe a statute meant to have broad application.").

[92] *See* Grider v. Texas Oil & Gas Corp., 868 F.2d 1147 (10th Cir.) (reasoning that liberal construction in favor of a remedy under one section of RICO was not required because an adequate remedy was available under a different section, even though the proposed remedy would burden the plaintiff less than would the available remedy, and would therefore have furthered the remedial purposes of the act), *cert. denied*, 493 U.S. 820 (1989); Schofield v. First Commodity Corp., 793 F.2d 28 (1st Cir. 1986) (same); P.M.F. Servs., Inc. v. Grady, 681 F. Supp. 549 (N.D. Ill. 1988) (same); Bruce J. Winick, *Forfeiture of Attorneys' Fees Under RICO and CCE and the Right to Counsel of Choice: The Constitutional Dilemma and How to Avoid It*, 43 U. MIAMI L. REV. 765, 848 n.405 (1989) ("In any case, the 'remedial' purpose of the forfeiture provisions can be effectuated fully without construing them to apply to legitimate attorneys' fees."); Liebman, *supra* note 87, at 141–43 (discussing *Grider*, *Schofield*, and *P.M.F. Servs.*).

a similar way, not to require the narrowest of the reasonable meanings, but instead to give the words "their fair meaning in accord with the manifest intent of the lawmakers."[93] Alternative interpretations like these obviously can limit the effect of an interpretive direction.

2. Self-Reference

Because an interpretive direction cannot be interpreted sensibly according to its own directions, courts are free to apply all of the traditional tools of statutory interpretation to the interpretive direction. The resulting uncertainty further reduces the legislature's ability to predict the effects of an interpretive direction.

For example, a court might think that it must liberally construe a liberal construction clause itself in order to effectuate the statute's "remedial purposes." If so, it would read "remedial purposes" broadly. The question is circular, however, because the liberal construction clause itself says that it and other provisions should be interpreted liberally only in those areas where doing so would effectuate a "remedial purpose." If criminal remedies are not included in that term, for example, then the liberal construction clause would not dictate liberal construction of itself to further such purposes. The scope of the clause therefore must be determined without reference to the clause's interpretive direction.

This will almost always be true of interpretive directions to some extent. Suppose an interpretive clause forbade reference to legislative history in resolving an act's ambiguities. This seems to be less ambiguous than a liberal construction clause. But if the legislative history indicated that the legislature did not intend to include certain types of materials in the term "legislative history," the court could not rely on the interpretive clause itself to prevent inquiry into those types of materials that indicated the clause's scope. If "legislative history" did not include certain things in the first place, the clause itself would not forbid consideration of them to determine its own scope. Similarly, a clause directing interpretation according to the "fair import" of

---

[93] United States v. Turkette, 452 U.S. 576, 587 n.10 (1981); United States v. Moore, 423 U.S. 122, 145 (1975); United States v. Bramblett, 348 U.S. 503, 510 (1955); United States v. Brown, 333 U.S. 18, 25–26 (1948).

the words used might be taken to be a sort of plain meaning rule, emphasizing reliance on the text. But if the legislative history indicated that "fair import" just means not construing statutes strictly, a court could not rely on the clause itself to adopt the former interpretation and block reference to legislative history. To do so would assume the issue to be decided: whether "fair import" means strict adherence to the text alone.

C. *Statutory Text and Identifying Ambiguity*

Even though interpretive directions may not say so explicitly, they generally operate only to resolve ambiguities in a statute. If the words of the statute are clear, the words will be enforced without construction or interpretation.[94] This would be true even for interpretive clauses requiring reference to certain materials in construing statutes, such as clauses declaring that the courts shall consult the legislative history or the preamble in interpreting an act. Since "interpreting" and "construing" are generally taken to mean the act of resolving ambiguities and uncertainties, courts would not be bound to follow the interpretive direction if the statute's words were clear.

A legislature might require certain procedures in all cases by explicitly declaring that the courts shall consult certain things or adopt certain interpretive attitudes, regardless of any ambiguity in the statute, before deciding what the words of the statute mean. However, legislators probably would not want to pass such an interpretive direction, since it would increase the chances that courts would depart from the plain meaning of the statute. The problem is not solved by giving courts discretion to follow the interpretive direction regardless of whether the statute is ambiguous. Instead, it makes the effect of the direction even less certain, while leaving courts free to disregard it. For example, Texas law expressly permits reference to an incomplete list of external aids even if the words of a statute are unambiguous.[95]

---

[94] *See, e.g.*, Moody v. Corsentino, 843 P.2d 1355, 1370 (Colo. 1993) (en banc); Estate of Brauch v. Beeck, 181 N.W.2d 132, 134 (Iowa 1970); Schurtz v. BMW of N. Am., Inc., 814 P.2d 1108, 1112 (Utah 1991); Blakey, *supra* note 53, at 289 n.150 ("'Where there is no ambiguity . . . there is no room for construction.'") (quoting United States v. Wiltberger, 18 U.S. (5 Wheat.) 76, 95–96 (1820)); Palm, *supra* note 6, at 170–71 ("If a statute is not ambiguous, courts should accord the language its plain meaning. There is no need for liberal or strict construction when the words of the statute are clear." (footnote omitted)).

[95] TEX. GOV'T CODE ANN. § 311.023.

Yet the Texas Court of Criminal Appeals emphatically declared that courts must decline the invitation and consider external sources only when ambiguity in the text creates an "absolute necessity."[96]

Of course, judges constantly decide whether statutes are ambiguous. Even though this is not unique to interpretive directions, it further limits their impact. If a judge does not want to follow an interpretive direction, for whatever reason, avoiding it is easier than escaping from the substantive text of the statute. Even disregarding deliberate manipulation of the interpretive process, the necessity of first acknowledging ambiguity reduces the predictability and consistency of interpretive directions.

The lack of settled rules for determining the existence of ambiguity increases the judiciary's room to maneuver and avoid legislative directions. First, the point at which a possible meaning creates ambiguity is uncertain.[97] Some might say that the existence of any other possible meaning makes a statute ambiguous, while others would say that there must be at least two meanings that are equally plausible. Second, people disagree about what can be considered initially in deciding whether a statute is ambiguous: they may draw the line at dictionary meanings, the syntactical context, the statute as a whole, the statute's purpose, the social context, or even legislative history. For example, one court considering the meaning of "enterprise" in RICO said that the term could be considered unambiguous only if the legislative history were entirely ignored; it concluded from the history that "enterprise" is not as broad as it would be were it considered alone.[98]

Even though a "clearly expressed legislative intent to the contrary" may override the unambiguous meaning of a statute,[99] an interpretive direction almost surely is not a "clearly expressed

---

[96] Boykin v. State, 818 S.W.2d 782, 785, 786 n.4 (Tex. Crim. App. 1991) (en banc).

[97] *See Turkette*, 452 U.S. at 580 ("[T]here is no errorless test for identifying or recognizing 'plain' or 'unambiguous' language.").

[98] United States v. Anderson, 626 F.2d 1358, 1369–71 (8th Cir. 1980), *cert. denied*, 450 U.S. 912 (1981); *see also* United States v. Altese, 542 F.2d 104, 107–08 (2d Cir. 1976) (Van Graafeiland, J., dissenting) (applying legislative history to resolve ambiguity rather than turning to liberal construction clause), *cert. denied*, 429 U.S. 1039 (1977).

[99] *See Turkette*, 452 U.S. at 580 (quoting Consumer Product Safety Comm'n v. GTE Sylvania, Inc., 447 U.S. 102, 108 (1980)) ("In determining the scope of a statute, we look first to its language. If the statutory language is unambiguous, in the absence of 'a clearly expressed legislative intent to the contrary, that language must ordinarily be regarded as conclusive.'").

legislative intent to the contrary." Many courts have declared that an interpretive clause cannot stretch the meaning of a statute beyond the limits of the text itself.[100]

D. *The Response of Courts to Interpretive Directions*

My objective throughout this discussion has been to explain how courts should respond and probably will respond to interpretive directions in various situations. With this understanding, legislatures can better decide whether to use interpretive directions and can better anticipate and avoid potential problems if they do use interpretive directions. Along with other factors, the legal and practical problems I have described make several general responses possible.

1. Ignoring Interpretive Directions

In the past, courts often ignored interpretive directions entirely.[101] Even today, with well-known, statute-specific clauses like the RICO liberal construction clause, courts may not acknowledge the existence of an interpretive direction.[102] If courts

---

[100] *See, e.g.*, Grider v. Texas Oil & Gas Corp., 868 F.2d 1147, 1150 (10th Cir.) ("[T]he general principle that RICO is to be accorded a liberal interpretation cannot justify expanding section 1962(a) beyond the limits of the subsection's own language."), *cert. denied* 493 U.S. 820 (1989); Haroco, Inc. v. American Nat'l Bank & Trust Co., 747 F.2d 384, 400 (7th Cir. 1984) ("[W]e do not think the general principle of liberal interpretation of RICO can be used to stretch section 1962(c) to reach this situation in the face of the subsection's own limits."), *aff'd*, 473 U.S. 606 (1985); P.M.F. Servs., Inc. v. Grady, 681 F. Supp. 549, 555 (N.D. Ill. 1988) ("In this Court's view, however, there is no reason (including the existence of [the RICO liberal construction clause]) to treat civil RICO as though it falls wholly outside the legal principles that normally apply to statutes and their application.").

[101] *See* Fordham & Leach, *supra* note 47, at 451 (noting cases where liberal construction clauses were ignored and common law strict construction applied); Frank E. Horack, Jr., *Cooperative Action for Improved Statutory Interpretation*, 3 VAND. L. REV. 382, 393 (1950) (observing that courts have paid little attention to general chapters governing statutory interpretation by the courts).

[102] *See, e.g.*, United States v. Computer Sciences Corp., 689 F.2d 1181, 1191 (4th Cir. 1982); United States v. Rubin, 559 F.2d 975, 990–93 (5th Cir. 1977), *vacated and remanded*, 439 U.S. 810 (1978); United States v. Moeller, 402 F. Supp. 49, 58 (D. Conn. 1975) (narrowly construing "enterprise" in RICO); Keeler v. Superior Court, 470 P.2d 617, 624 (Cal. 1970) (en banc) (ignoring statutory abrogation of strict construction clause, declaring that "[i]t is the policy of this state to construe a penal statute as favorably to the defendant as its language and the circumstances of its application may reasonably permit"); *id.* at 632 (Burke, C.J., dissenting) (criticizing majority for ignoring statutory abrogation of strict construction clause, advocating construction according to fair import of text).

do acknowledge its existence, they may still decide not to follow it.[103]

There are several reasons why a court might ignore an interpretive direction. First, the court might be unaware of its existence, or might have forgotten about it. This is especially likely with general interpretive directions that apply to all or a class of statutes, since courts may focus on the statute in question and never think to look at a different title to consult a general interpretive standard.[104]

Even if the court does know about relevant interpretive directions, it may still ignore them. For example, it may conclude that an interpretive direction does not apply because the statute is unambiguous and therefore does not need to be construed.[105] Textualist judges thus may be especially unlikely to apply liberal construction clauses.[106] However, even non-textualist judges may avoid considering interpretive directions if the meaning of otherwise ambiguous words is clear because of context, legislative history, or other interpretive principles and canons.[107] A court also may conclude that other interpretive principles take precedence over an interpretive direction.[108] Or a court may interpret an interpretive direction in a way that limits its application and thus does not require application to the issue before it.[109]

---

[103] *See, e.g.*, United States v. Mandel, 415 F. Supp. 997, 1022 (D. Md. 1976), *rev'd*, 591 F.2d 1347 (4th Cir.), *aff'd by equally divided court en banc*, 602 F.2d 653 (4th Cir. 1979), *cert. denied*, 445 U.S. 961 (1980).

[104] *See* Hall, *supra* note 47, at 755 & nn.38–39 (suggesting that courts may have ignored liberal construction clauses because text writers did not mention them); Horack, *supra* note 101, at 393 (observing that some cases reveal that court and counsel were unaware of interpretive directions from the legislature).

[105] *See* Fordham & Leach, *supra* note 47, at 453 (observing that courts may avoid the effect of a liberal construction clause by invoking the rule that a statute cannot be extended beyond the plain meaning of its words); DaCosta, *supra* note 72, at 72 ("[I]t is possible to find ambiguities where none appear to exist and a court's willingness to search them out usually increases with the severity of the penalty."); *supra* part III.C.

[106] Of course, a textualist might acknowledge ambiguity but, citing a construction clause, choose the favored meaning without referring to legislative history or other external aids.

[107] *See* Kaushal v. State Bank of India, 556 F. Supp. 576, 583 (N.D. Ill. 1983) (denying equitable remedies to private RICO plaintiffs because legislative history and interpretive principles for implying equitable remedies resolved textual ambiguities); *supra* notes 97–98 and accompanying text.

[108] *See, e.g.*, State v. Pena, 683 P.2d 744, 748–49 (Ariz. Ct. App. 1983) (arguing that although statute abolished rule of strict construction in favor of "fair meaning" of the text, the rule of lenity still resolves doubts in favor of criminal defendants); State v. Rackle, 523 P.2d 299 (Haw. 1974) (declaring that penal statutes must be strictly construed in favor of defendant despite statutory abrogation of strict construction); *supra* part III.A.

[109] *See supra* part III.B.

Some judges may not apply interpretive directions because they feel that they are not obligated to apply interpretive directions in all cases where they seem to apply. Judges often pick and choose among available common law interpretive canons, because common law canons are generally considered aids to construction and not mandatory principles that must be followed in all cases where they apply.[110] Courts may feel that interpretive directions from the legislature likewise are merely guidelines for interpretation. If a court decides that other principles of interpretation, as well as policy considerations, favor an outcome that conflicts with an interpretive direction, it is likely to ignore the direction.[111]

2. Considering Interpretive Directions as a Factor

A court that does not always follow interpretive directions may still be influenced by them. At the very least, if a judge is aware that an interpretive direction exists, the judge will probably feel obligated to apply it unless he or she can think of some reason for not doing so. Interpretive clauses therefore may "reduce . . . the degrees of freedom that attend the addressee's interpretive processes."[112] Interpretive clauses also may prevent the use of certain justifications for decisions. For example, a liberal construction clause that explicitly abrogates the common law rule of strict construction may prevent judges from using the common law canon to justify narrow interpretations.[113]

A legislature can never exercise complete control of interpretation since, in the end, statutes are always interpreted by the judiciary. Perhaps the best a legislature can hope for is to give courts "a direction for profitable thinking."[114]

---

[110] *See, e.g.*, Boykin v. State, 818 S.W.2d 782, 785 n.3 (Tex. Crim. App. 1991) (en banc) ("[T]he canons of construction are no more than rules of logic for the interpretation of texts.").

[111] *See* H.J., Inc. v. Northwestern Bell Tel. Co., 492 U.S. 229 (1989) (relying on intent inferred from text and legislative history to hold that the RICO "pattern" requirement demands a relationship among predicate acts, although the statute did not indicate whether anything besides two unconnected acts of racketeering activity was required to constitute a "pattern," and although such a requirement would hinder RICO's remedial purposes).

[112] D'Amato, *supra* note 5, at 565.

[113] *See* Fordham & Leach, *supra* note 47, at 450.

[114] Hall, *supra* note 47, at 770 (quoting James M. Landis, *A Note on "Statutory Interpretation"*, 43 HARV. L. REV. 886, 892–93 (1930)). Professor Hall invoked this phrase in conceding that his proposed law specifying when strict construction is

3. Considering Interpretive Directions as Indications of Legislative Intent

Another way that judges may be influenced by interpretive directions without actually following them is to consider them as evidence of legislative intent.[115] A court may cite a liberal construction clause along with extrinsic interpretive aids as evidence that the legislature must have used ambiguous words in a broad sense. Similarly, a court might cite an interpretive direction forbidding the use of certain extrinsic sources in support of its decision to weigh such sources less heavily, even though the court does not follow the legislature's direction not to consult those sources at all.

In *Russello v. United States*,[116] for example, the Supreme Court did not mention the RICO liberal construction clause when it explained its reasons for deciding that the "interest" that must be forfeited in RICO includes profits from an enterprise, not just interests in the enterprise itself. It analyzed the broad meaning of the general term, compared the clause to subsequent clauses that were not as broad, and reasoned that a narrower interpretation would not impose much of a penalty. Then the Court said: "If it is necessary to turn to the legislative history of the RICO statute, one finds that that history does not reveal, as petitioner would have us hold, a limited congressional intent."[117] The Court then cited the liberal construction clause as evidence that Congress intended a broad application of the statute.[118] Had the Court strictly followed the liberal construction clause, it simply would have said that "interest" could have the broader meaning, which would further RICO's purposes of combating racketeering activity.[119] The Court still might have mentioned the other reasons it

---

appropriate "will not reform the law on this subject," but might at least have some influence on courts if called to their attention. *Id.*

[115] *See* Rutgers Chapter of Delta Upsilon Fraternity v. City of New Brunswick, 28 A.2d 759, 763 (N.J. Sup. Ct. 1942) (declaring that New Jersey statute forbidding implications or presumptions based on arrangement of sections is merely a guide to legislative intent), *aff'd*, 32 A.2d 364 (N.J. 1943); City of Huntington v. State Water Comm'n, 64 S.E.2d 225, 230 (W. Va. 1951) ("Though . . . 'the courts, not the Legislature, exercise final authority in the construction of statutes,' the legislative direction [a liberal construction clause] will be given consideration as indicative of the legislative intent." (quoting State *ex rel.* Holbert v. Robinson, 59 S.E.2d 884, 889 (W. Va. 1950))).

[116] 464 U.S. 16 (1983).

[117] *Id.* at 26 (citation omitted).

[118] *Id.* at 27.

[119] This is true unless perhaps the Court believes that interpretive directions, such as

gave, but they would have been secondary to the interpretive direction.[120]

Some commentators favor this approach. Judge Posner has suggested that judges should watch for and consider signs of legislative intent concerning their freedom in interpreting statutes.[121] Strict or liberal construction clauses, he says, are clues to how a judge should approach interpretation, along with other clues such as the structure of the statute.[122] Of course, this approach is only defensible if the legislature does not have the power to limit judges' interpretive freedom. Although I have concluded that the legislature does have considerable power to direct interpretation, many judges, apparently including Judge Posner, disagree.

4. Following the Interpretive Direction

A final alternative, of course, is for the court to do exactly what the legislature told it to do. Although there are many valid ways to avoid following interpretive directions, a court might feel that interpretive directions are as binding as substantive law, and that it therefore has an obligation to follow the interpretive direction according to its terms. Many courts have done just that.[123]

---

liberal construction clauses, should be applied directly only when all other traditional tools of statutory interpretation have been applied, including reference to signs of legislative intent in the legislative history and the statute itself. *See supra* notes 89–93 and accompanying text.

[120] Earlier, in United States v. Turkette, 452 U.S. 576 (1981), the Court had cited the liberal construction clause in support of its decision. However, the citation of the clause followed its rebuttal of the lower court's reasoning and the Court's own analysis of the statutory language, legislative history, and statutory structure. The Court also said that "[w]ith or without this admonition," it would not interpret "enterprise" to exclude wholly criminal organizations. *Id.* at 587.

[121] Richard A. Posner, *Statutory Interpretation—In the Classroom and in the Courtroom*, 50 U. CHI. L. REV. 800, 818 (1983).

[122] *Id.*; *see also* Senah Elizabeth Green, Note, *Judicial Efforts to Redirect an Errant Statute: Civil RICO and the Misapplication of Vicarious Corporate Liability*, 65 B.U. L. REV. 561, 594–95 (1985) (suggesting that liberal construction clause is nothing more than evidence of legislative intent to encourage broad application of RICO).

[123] *See, e.g.*, United States v. Altese, 542 F.2d 104, 106 (2d Cir. 1976) ("[W]e are obliged to construe the Act liberally."), *cert. denied*, 429 U.S. 1039 (1977); *In re* Marriage of Swink, 807 P.2d 1245, 1247 (Colo. Ct. App. 1991) (citing and following interpretive directions requiring ascertainment of legislative intent and liberal construction of marriage and divorce act); Commonwealth v. Trent, 77 S.W. 390, 393 (Ky. 1903) ("It was within the competency of the Legislature to enact the statute, and it is incumbent on the judiciary to enforce it."); State v. Hemrich, 161 P. 79, 81 (Wash. 1916) (citing liberal construction clause of state prohibition law and declaring that it "must therefore construe the act liberally and inclusively"); *supra* note 48.

IV. CONCLUSION

As this discussion has suggested, most cases in which interpretive directions should apply probably are not strictly decided according to the legislature's directions. Most of the time, courts will either avoid the direction—by concluding that the text is unambiguous, deciding that other interpretive standards should govern, or some other way—or will consider it along with other factors that typically influence judicial decisions. Furthermore, courts that do seem to obey interpretive directions often would have decided the same way even without the interpretive direction, and they simply cite the direction as additional support for their conclusions.[124]

Even though interpretive directions often do not seem to permit it, such judicial flexibility is probably best. This way, courts retain control of the interpretive process, yet often they will give some consideration to the legislature's intent as expressed through interpretive directions. At the same time, this approach helps avoid the inevitable departures from actual legislative preference that would come from mechanical application of interpretive directions. Legislatures cannot possibly foresee every possible ambiguity. If they could, they would be unlikely to agree that every such ambiguity should be resolved in the same way. In the past, legislatures probably have used interpretive directions knowing that courts would not rigidly follow them, but hoping that at least they would clarify the general intent of the statute for the courts.

Legislatures may want to consider using interpretive directions to identify sources for courts to consult in interpreting statutes. Although such directions face many of the same problems as more traditional interpretive provisions, they may help to increase the predictability of judicial interpretation. But, in gen-

---

[124] *See Turkette*, 452 U.S. at 587 (declaring that "[w]ith or without this admonition [the RICO liberal construction clause]," the Court would not interpret "enterprise" to exclude wholly criminal organizations); WAYNE R. LAFAVE & AUSTIN W. SCOTT, CRIMINAL LAW 70 (1972).

> There is something of a dispute among those who like to speculate on the workings of the judicial mind as to whether courts first decide how a defective statute ought to be interpreted and then display whatever canons . . . will make this interpretation look inevitable, or whether the courts actually first use the applicable canons and second reach the result . . . . At all events . . . rules of statutory interpretation sometimes do decide cases.

Id.; *see also* Fordham & Leach, *supra* note 47, at 450 (expressing uncertainty about how often interpretive directions actually change the results of interpretation).

eral, legislatures should not expect much from interpretive directions. They are no more likely to make the interpretive process more rational or predictable than are judicial canons of interpretation. Legislatures can best control the application of the laws by carefully drafting their substance and responding with further legislative action when courts misapply the laws.[125]

Not only would focusing on substantive drafting and responding to judicial mistakes better control the application of the laws, it would also better respect the separation of powers. Even though most interpretive directions probably would survive constitutional challenge, they nevertheless encroach on the judicial function and go beyond the legislature's area of greatest expertise. Both self-interest and broader societal interests thus encourage legislatures to respect the courts' interpretive function and to accommodate and guide it rather than control or restrict it.

[125]This conclusion echoes Frank Horack's conclusion over 40 years ago: "Legislatures can no more solve the problems of interpretation by the enactment of canons . . . than can courts. Only care in the preparation and drafting of the statutes will contribute to the judicial use of legislative materials." Horack, *supra* note 101, at 393.

NOTE

THE NEED FOR A PUBLIC PERFORMANCE RIGHT IN SOUND RECORDINGS

WILLIAM H. O'DOWD*

The Copyright Act currently excludes owners of copyrights in sound recordings, such as musical performers and recording companies, from the right to authorize or be compensated for public performances of their works. All other copyright holders in musical works, such as songwriters, composers, and publishers, have protected public performance rights.

In this Note, Mr. O'Dowd argues that the emergence of digital technology threatens to exacerbate this inequity to the point that holders of copyrights in sound recordings will have little economic incentive to produce new works. He analyzes and rebuts broadcasters' arguments against the provision of a performance right to copyright holders in sound recordings. He then analyzes the market effects of various proposed legislative solutions and proposes that Congress act quickly to amend the law while digital transmission services are still in their early stages.

The lack of a public performance right for owners of copyrights in sound recordings has long been an anomaly in United States copyright law that has disturbed members of the U.S. recording industry. However, with the emergence of digital technology, the failure to protect public performance rights has become more of a threat to the continued economic well-being of the recording industry. Under the current statute, 17 U.S.C. § 101 (1988), copyright owners of sound recordings do not have a public performance right in their copyrighted works. Musical performers and record companies—unlike songwriters, composers, music publishers, and all other owners of a right in a copyrighted musical work—have no rights to authorize or receive remuneration for the broadcast or other public performance of their copyrighted sound recordings.[1] This inconsistency in the

* B.A., Creighton University, 1990; M.A., 1991; Harvard Law School Class of 1994.

[1] 17 U.S.C. § 114(a) (1988) defines the scope of exclusive rights in sound recordings. It currently reads: "The exclusive rights of the owner of copyright in a sound recording are limited to the rights specified by clauses (1), (2), and (3) of section 106 [17 U.S.C. § 106(1)–(3)], and do not include any right of performance under section 106(4) [17 U.S.C. § 106(4)]."

17 U.S.C. § 106 (1988) provides:

> Subject to sections 107 through 118 [17 U.S.C. §§ 107–118], the owner of copyright under this title [17 U.S.C. §§ 101–914] has the exclusive rights to do and to authorize any of the following:
>
> (1) to reproduce the copyrighted work in copies or phonorecords;
>
> (2) to prepare derivative works based upon the copyrighted work;

statute leads to confusing and counterintuitive applications of the law. For example, the owner of a copyright in a sound recording could control the public performance of a music video of the song, as a music video falls under the definition of an audiovisual work rather than a sound recording. However, the copyright owner would have no control over the public performance of that same recorded song broadcast over the radio.[2]

This failure to recognize the substantial contributions of the performers and producers of sound recordings has always been a significant weakness in the copyright law's protective framework. However, the advancement of digital audio transmission services, and the resulting substantial rise in the number and type of public performances of sound recordings, threaten to leave the owners of such copyrights with very little protection for their investment of time and creativity. Sound recordings stored in digital media, such as a compact disc (CD) or computer software, have several advantages over their traditional analog counterparts. They may be transformed easily and rapidly, stored indefinitely, and, most importantly, transmitted and received without any loss of fidelity whatsoever. Thus, a digital audio transmission service, upon the purchase of a CD, may deliver that product electronically to millions of homes via satellite, radio, or cable without any payment to the creators of the sound recording.

The proposed implementation of "Audio on Demand" services in the near future—allowing the consumer to choose music from the entire library of sound recordings currently on compact disc, play the desired selections on her stereo system, and make perfect copies to add to her tangible collection—will only exacerbate the problem. Record sales will drop dramatically. Since copyright owners in sound recordings derive their profits from

---

(3) to distribute copies or phonorecords of the copyrighted work to the public by sale or other transfer of ownership, or by rental, lease, or lending;

(4) in the case of literary, musical, dramatic, and choreographic works, pantomimes, and motion pictures and other audiovisual works, to perform the copyrighted work publicly; and

(5) in the case of literary, musical, dramatic, and choreographic works, pantomimes, and pictorial, graphic, or sculptural works, including the individual images of a motion picture or other audiovisual work, to display the copyrighted work publicly.

[2] *Performance Rights in Sound Recordings: Hearings Before the Subcomm. on Intellectual Property and Judicial Administration of the House Comm. on the Judiciary*, 103d Congress, 1st Sess. (1993) (testimony of Jason S. Berman, President of the Recording Industry Association of America, Inc.) [hereinafter Berman Testimony].

sales, not performances, the economic incentive to invest in the production of sound recordings will disappear. The detrimental effect of these new digital technologies was unforeseen when the copyright statute was passed, so Congress must amend the statute now to avert the harm to owners of sound recording copyrights. The simplest and most equitable solution is to provide such copyright owners with the exclusive right to authorize or prohibit the public performance of their copyrighted works.

I. A Brief History of Sound Recordings and Public Performance Rights in the Copyright Law Through the 1976 Copyright Act

Both the 1891 and 1909 Copyright Acts provided a public performance right to owners of copyrights in nondramatic musical compositions.[3] That is, a copyright owner of a nondramatic musical composition was granted the exclusive right to perform the work publicly for profit and could demand a fee, via a non-exclusive license, for someone else's public performance of the work. However, the public performance of a song was the best promotional tool for the sale of sheet music to that song, and at that time income from sales of sheet music provided the greatest royalties to songwriters. Therefore, there was no economic incentive for the private enforcement of public performance rights.[4] Furthermore, a copyright owner would only have been able to selectively enforce the public performance right before the advent of royalty collection societies, such as the American Society of Composers, Authors and Publishers (ASCAP) and Broadcast Music, Inc. (BMI), that issue "blanket licenses" and effectively monitor public performances of music.[5]

Unfortunately, the drafters of the 1909 Copyright Act did not foresee the new era for communications and broadcasting. New mediums of copyrightable expression quickly rendered the traditional language of statutory law awkward and difficult to ap-

---

[3] 1891 Copyright Act, ch. 565, § 4952, 26 Stat. 1107 (current version at 17 U.S.C. § 101 (1988)); 1909 Copyright Act, ch. 320, § 1(d), 35 Stat. 1075 (current version at 17 U.S.C. § 101 (1988)).

[4] Bernard Korman, *Performance Rights in Music Under Sections 110 and 118 of the 1976 Copyright Act*, 22 N.Y.L. Sch. L. Rev. 521, 523 (1977) (citing *Hearings on S. 6330 and H.R. 19853 Before the House Comm. on Patents and the Senate Comm. on Patents*, 59th Cong., 1st Sess. 25 (1906).

[5] ASCAP was founded in 1914. BMI was organized in 1939. Broadcast Music v. Columbia Broadcasting System 441 U.S. 1, 4–5 (1979).

ply.[6] During congressional hearings prior to the enactment of the 1976 Copyright Act,[7] Herman Finkelstein, the general counsel of ASCAP, pointed to the example of the jukebox.[8] He testified that each year, approximately $500 million in revenue was captured by an estimated 500,000 jukebox machines that played recordings of copyrighted music, yet the copyright owners received no compensation for the commercial exploitation of their work.[9] Owners of jukeboxes benefitted from an exemption for performances in penny-parlors granted by the 1909 Copyright Act, although Congress certainly did not envision the magnitude of revenue to be generated from mechanical performances at the time it drafted the statute.[10]

By the time congressional hearings on the revision of the 1909 Copyright Act began in the 1960's, technological advancement had greatly altered the methods and mediums by which music was distributed and sold. The performances of individual singers and musicians were captured on records and cassettes, which were easily played, stored, and reproduced in the home. Thus, an individual's unique rendition of a particular song could be reproduced and commercially exploited. Congress first recognized a performer's rendition of a song as a separate, copyrightable expression by passing the 1971 Amendment to the 1909 Copyright Act.[11] The 1971 Amendment established a copyright in sound recordings and granted the owner of such a copyright the exclusive right to reproduce the work.[12]

Within a year of its enactment, the amendment was challenged in the District Court for the District of Columbia in the case of *Shaab v. Kleindienst*.[13] A three-judge court unanimously held

---

[6] Jessica Litman, *Copyright Legislation and Technological Change*, 68 OR. L. REV. 275, 277 (1989).

[7] Copyright Act of 1976, Pub. L. No. 94-553, 90 Stat. 2565 (1976) (amended in 1988).

[8] Todd Piccus, *The Small-Business Exemption and the Realities of the Home-Electronic Marketplace*, THE ENTERTAINMENT AND SPORTS LAWYER, Winter 1993, at 8, 10. The evolution of the copyright law culminating in the 1976 Copyright Act was initiated in 1955 when Congress allocated funds for a series of Copyright Office studies of copyright law. The first copyright revision bill was introduced in 1964, and congressional hearings started soon thereafter.

[9] *Copyright Law Revision: Hearings on H.R. 4347, H.R. 5680, H.R. 6831, and H.R. 6835 Before the Subcomm. 3 of the House of Representatives Comm. on the Judiciary*, Ser. No. 8, Pt. 1, 89th Cong., 1st Sess. 178 (1965) (statement of Herman Finkelstein, general counsel, ASCAP).

[10] Piccus, *supra* note 8, at 10.

[11] *See* 17 U.S.C. §§ 1(f), 5(n) (Supp. II 1972) (amending 17 U.S.C. §§ 1, 5 (1970)).

[12] *Id.*

[13] 345 F. Supp. 589 (D.D.C. 1972).

that "[t]echnical advances, unknown and unanticipated in the time of our founding fathers, are the basis for the sound recording industry. The copyright clause of the Constitution must be interpreted broadly to provide protection for this method of fixing creative works in tangible form."[14] The revisions to the 1909 Copyright Act as a whole were completed in 1976 and became effective on January 1, 1978. The 1976 Copyright Act provided the owner of a copyright in a sound recording with the exclusive rights to prepare derivative works based upon the copyrighted work[15] and to distribute the phonorecord to the public.[16] Thus, the 1976 Copyright Act added to the rights already associated with a copyrighted sound recording. However, the Act specifically denies the copyright owner a performance right in sound recordings.[17]

The original version of the 1976 Copyright Act included a provision which would have established a performance royalty right in sound recordings. Congress envisioned mandatory statutory payments to those recording artists whose musical performances were publicly broadcast for commercial purposes.[18] Commercial exploiters of sound recordings, most prominently radio stations, would have been required to make annual flat fee royalty payments calculated on the basis of their gross advertising revenues.[19] These royalty payments would have been collected and distributed to the performers who created the sound recordings as well as to the copyright owners of the sound recordings (generally, the record companies).[20] Payments on each individual song would have been in proportion to the amount of airplay the particular song received.[21]

Furious lobbying efforts by broadcasters against this provision threatened passage of the entire Copyright Act.[22] Eventually, the

---

[14] *Id.* at 590. *See also* Goldstein v. California, 412 U.S. 546, 561 (1973).

[15] 17 U.S.C. § 114(b) (1988).

[16] 17 U.S.C. § 106(3) (1988).

[17] 17 U.S.C. § 114(a) (1988).

[18] *Performance Royalty: Hearings on S. 1111 Before the Subcomm. on Patents, Trademarks, and Copyright of the Senate Comm. on the Judiciary*, 94th Cong., 1st Sess. 1–4 (1975) (statement of Sen. Hugh Scott).

[19] *Id.* at 3–4.

[20] *Id.* at 4.

[21] 17 U.S.C.A. § 115(c) (West Cum. Supp. 1977). In effect, the amount of payments to any individual performer would have been the greater sum between a rate of two and three-quarter cents per broadcast and one-half cent per minute of playing time of that performer's works. *Id.*

[22] Allen Edward Molnar, Comment, *Performance Royalties and Copyright: A Question of "Sound" Policy*, 8 SETON HALL L. REV. 678, 680–81 (1977).

1976 Copyright Act was passed without the clause creating a public performance right in sound recordings. Rather than dismiss the issue altogether, however, Congress set aside the performance royalty question for further review by the Register of Copyrights.[23] In 1978, even prior to the advent of new digital transmission and recording technologies, the Register concluded that a performance right should be granted to copyright owners of sound recordings:

> Broadcasters and other users of recordings have performed them without permission or payment for generations. Users today look upon any requirement that they pay royalties as an unfair imposition in the nature of a "tax." However, any economic burden on the users of recordings for public performance is heavily outweighed . . . by the commercial benefits accruing directly from the use of copyrighted sound recordings . . . . Sound recordings are creative and their unauthorized performance results in both damage and profits. To leave the creators of sound recordings without any protection or compensation for their widespread commercial use can no longer be justified.[24]

However, despite the conclusions reached in the Register's report, Congress took no further action toward amending the Act to create a performance right.

II. The Threat of Digital Technology[25]

Two distinct trends of technological development will ensure that the future of information transit lies in the digital realm. First, digital technicians are constantly creating new applications

---

[23] 17 U.S.C. § 114(d) (1988) provides:
On January 3, 1978, the Register of Copyrights, after consulting with representatives of owners of copyrighted materials, representatives of the broadcasting, recording, motion picture, entertainment industries, and arts organizations, representatives of organized labor and performers of copyrighted materials, shall submit to the Congress a report setting forth recommendations as to whether this section should be amended to provide for performers and copyright owners of copyrighted material any performance rights in such material. The report should describe the status of such rights in foreign countries, the views of major interested parties, and specific legislative or other recommendations, if any.

[24] U.S. Register Of Copyrights, Report on Performance Rights in Sound Recordings, H.R. Doc. No. 15, 95th Cong., 2d Sess. 1063 (1978) [hereinafter Report of Register of Copyrights].

[25] For the purposes of this Article, "digital technology" refers to the use of equipment which can accept and interpret digital data, or information consisting only of discrete numeric values, typically zeros and ones (the binary code).

for digital software to existing media. Digital technology has already had an enormous impact on the recording industry. Sound recordings are created with the aid of computer-enhanced sampling and synthesizing methods, produced (and reproduced) perfectly on compact discs (CDs), and distributed or broadcast distortion-free over cable lines or satellite networks. Second, the percentage of the general population with access to digital technology is growing every day, as evidenced by the boom in personal computer sales.[26]

An even greater number of people would have access to digital technology if it was made available through existing means of telecommunications, namely cable television lines or satellite networks.[27] The infrastructure necessary for the new digital transmission systems is either already established or expected to be in place by the end of this century.[28] Satellite and cable networks are clearly a reality, and their owners and operators are racing to be the first to take advantage of digital advances.[29] Wireless digital broadcasting, or digital audio broadcasting (DAB), would work under the same principle as current analog radio signals, except that digital signals could be received and retransmitted, thereby placing few limits on reaching potential audiences.

The implications of the digital transmission revolution upon the recording industry are enormous. The current effort of the telecommunications and cable industries to create an information superhighway into every home in the country will need to produce programming to make the venture profitable. It is no

---

[26] General Information and Program, World Intellectual Property Organization Worldwide Symposium on the Impact of Digital Technology on Copyright and Neighboring Rights 2 (March 31 to April 2, 1993) (available in Harvard Law School Library).

[27] Nicholas Garnett, Remarks at the World Intellectual Property Organization Worldwide Symposium on the Impact of Digital Technology on Copyright and Neighboring Rights 6 (April 1, 1993) (transcript available in Harvard Law School Library). Garnett further notes that "there are currently two technological factors influencing the potential capacity of cable: fibre optic cable and digital compression." *Id.* The installation of fiber optic cable lines lessens the risk of signal loss or distortion to the fidelity of digital transmissions. Digital compression technology will expand the capacity of cable systems by six- to eightfold.
Id.

[28] *Id.* at 4.

[29] Bruce York, Remarks at the World Intellectual Property Organization Worldwide Symposium on the Impact of Digital Technology on Copyright and Neighboring Rights 6 (April 1, 1993) (transcript available in Harvard Law School Library) (noting that, with the use of satellite transmissions, within the next five years one could "leave by car from New York to Los Angeles and never lose reception of 50 channels of perfect sound").

coincidence that the suitors for Paramount Communications, which owns a wealth of original entertainment programming, are telecommunications and cable companies. Simply put, it does a cable company no good to have a 1,000 channel capacity if it only has programming for 25 channels.

The rapidly expanding capacity of cable systems will ensure that the demand for programming will be greater than the supply for the foreseeable future. However, the promise of interactive or multimedia technology in the near future provides unexplored opportunities for cable operators. Cable operators will be able to let the consumer choose programming from a library of available materials. Conceivably, this library could include all types of entertainment, including books, movies, and music.[30]

With respect to music, the goal of a fully operational, audio-on-demand service is often described as the "celestial jukebox."[31] This cable service would allow the consumer to choose from a library of all recorded music and have her selections digitally transmitted to her home at once, much like the operator of a jukebox which contains every album ever produced. Because the transmission would be in digital form, the consumer could copy the music onto a compact disc or digital audio tape to be played outside the home without any loss in quality.[32] Furthermore, the consumer could choose to have a series of musical selections transmitted at once in scrambled form, stored, and played back or reproduced at some future point in time.[33]

Copyright owners of sound recordings face a unique threat from this evolution of digital transmission. Clearly, these concepts of digital transmission obscure the boundaries between

---

[30] Garnett, *supra* note 27, at 6. Garnett notes:

> Again, it must be stressed that while the technology is still very much advancing, the functions described above are already a reality. At present, TeleCommunications Inc. (TCI), the largest US cable operator, is test marketing the concept of video on demand in Denver. Its test is called "Take One" and allows consumers to choose a selection from a library of 1,000 film titles and have it played almost instantaneously. Likewise Time Warner, the second largest cable operator, has also announced plans to commence experiments with a similar system—called the "Electronic Superhighway"—in Florida by early 1994. *Id.*

[31] Jason S. Berman, Remarks at the World Intellectual Property Organization Worldwide Symposium on the Impact of Digital Technology on Copyright and Neighboring Rights (April 1, 1993) (transcript available in Harvard Law School Library) [hereinafter Berman].

[32] Unlike standard analog recordings, repeated duplications of digital sound recordings do not result in increased static or poorer quality in sound.

[33] Berman, *supra* note 31, at 6.

"broadcasting" and "distribution." Under law, the ability to distribute the music is the exclusive right of the copyright owner of the sound recording. By digitally transmitting music, broadcasters are essentially "distributing" the music. Indeed, a true "celestial jukebox" is the ultimate form of distribution. Yet, because such transmissions are currently regarded as broadcasts, or a form of public performance, no compensation need be made to the creators of the programming for the distribution of their works.[34] Such "broadcasts" result in the loss of sales and deprive copyright owners of sound recordings of the key right created by the copyright laws.

Digital cable audio services in the United States are in their nascent stages. These systems are managed much like a premium movie channel, such as HBO or Showtime. The consumer pays a monthly fee in order to receive "extra" channels of digital audio transmissions.[35] Each channel plays a different style of music (e.g., jazz, classical, blues, heavy metal).[36] Furthermore, the consumer is able to feed the digital transmission directly into her stereo system, thereby allowing the consumer to control the play of the music in all varieties of ways (e.g., the level of bass, the relative volume of speakers, etc.).[37]

These digital cable audio services are directly competing with

---

[34] Nicholas Garnett adds this story:

> An important element of the particular cable digital audio system currently attempting to penetrate the European cable market is that it is fed from a satellite link from a base in Atlanta, Georgia. In initial negotiations with the recording industry in Europe, representatives of this particular system were asked what arrangements had been made with record producers for the supply of material to feed the European system. "None whatever" came the reply, "there is no requirement under relevant US laws to obtain authorisation from the producers for the uplink of their sound recordings; the material is taken from CDs purchased in Tower Records". [sic]

Garnett, *supra* note 27, at 7.

[35] *Id.* at 6–7.

[36] *Id.*

[37] *Id.* at 7. Garnett adds:

> It is worth pausing here to examine the current capability of this system. Assume it offers its subscribers 32 channels of digital recordings, each channel running for 24 hours a day. Assume in addition an average running time for one entire CD programme of 1 hour. This guarantees the consumer access during the course of one day to the equivalent of up to 768 CD programmes way beyond the average existing domestic CD library. Consider further that a major international recording company's catalogue may hold at any one time up to 10,000 CD programmes; every one of these could be transmitted by one cable operator in under two weeks. The entire worldwide inventory of phonograms currently available on CD could be thus delivered by one cable operator in well under 6 months.

Id.

retail record stores for the "sale" of recorded music. These companies currently offer multichannel programming and constant variety in order to dissuade the consumer from purchasing albums.[38] The construction of a "celestial jukebox" would eliminate the need for album purchases altogether. Before the completion of such a system, however, the digital cable audio companies will offer a pay-per-listen service, akin to current pay-per-view services for the distribution of movies.[39] Consumers will pay a rate (undoubtedly below standard retail list price of a new cassette) for the ability to receive a specific album, or selection of albums, on demand.[40]

The threat of private copying from digital transmissions is a very real one to copyright owners of sound recordings. The Office of Technology Assessment (OTA) estimated in 1989 that the recording industry loses approximately one-third of its sales on a yearly basis to unauthorized duplications of sound recordings.[41] The OTA's report further concluded that one billion musical works are copied every year in this country, approximately 439 million of which were made either from radio or television broadcast sources.[42] The Recording Industry Association of America (RIAA) calculates the lost revenues from private copying of sound recordings to be near $1 billion per year.[43] The "celestial jukebox" and other planned digital audio services threaten to exacerbate these losses greatly. Congressman Moorhead stated his support of the proposed Record Rental Amend-

---

[38] One brochure by Digital Cable Radio claims there will be "no need to spend a fortune on a CD library." Berman Testimony, *supra* note 2, at 7–8.

[39] Comments of the Recording Industry Association of America in Response to Notice of Inquiry Concerning Digital Audio Broadcast and Cable Services Before the United States Copyright Office, Library of Congress, In the Matter of Copyright Office Request for Comments Regarding Digital Audio Broadcast and Cable Services 4 (December 14, 1990) [hereinafter RIAA Comments to Copyright Office].

[40] *Id.* at 7. The promotional material of Digital Cable Radio services touts "Album Hours" during which "'an entire CD will be played from beginning to end, uninterrupted. This will be an ideal opportunity for fans to record the work of a favorite artist from the finest quality audio available. The quality will be the same as recording off their own CD player.'" *Id.*

[41] Comments of the Recording Industry Association of America Before the Federal Communications Commission In the Matter of Amendment of the Commission's Rules with regard to the Establishment and Regulation of New Digital Audio Radio Services 5 n.6 (October 11, 1990) [hereinafter RIAA Comments to FCC] (citing U.S. Congress, Office of Technology Assessment, *Copyright and Home Copying: Technology Challenges the Law*, OTA-CIT-422, 154 (Washington, D.C.: U.S. Government Printing Office, Oct. 1989)).

[42] *Id.*

[43] *Id.* at 5.

ment in 1984 with remarks that can also apply to the threat posed by digital transmission services:

> The problems which creators and inventors face today is more than a clash between titan commercial interests. The larger and more difficult problem is the adaptation of old concepts of copyright law, to new and rapidly changing technologies. The problem today is that the public has access like it's never had access before but the creator is not receiving his just compensation. New technologies have brought the concert into the living room but not the box office . . . .[44]

III. The Legal Justification for Providing Economic Incentives Through the Copyright Law

Creators of "cultural" goods (i.e., music, movies, television programs, etc.) that can easily be reproduced need the protection of the copyright law to provide an economic incentive to maximize investment in the creation of those goods. The "[p]rogress of [s]cience and useful [a]rts" is the constitutional mandate underlying the establishment of the copyright law.[45] The Supreme Court has recognized that the best means by which to achieve this goal is by providing economic incentives to the individual creators, saying, "The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in 'Science and useful Arts.'"[46] The public's benefit from the creativity of authors is maximized by the copyright monopoly, which is necessary to stimulate the greatest creativity of authors.[47]

As noted in the previous section, digital home delivery of sound recordings directly competes with the sales of packaged, prerecorded music. Prerecorded music provides the only means by which copyright owners of sound recordings are currently able to seek return on their investment of time, creativity, and money. In essence, the threat posed by the digital broadcasting

---

[44] Quoted in Berman, *supra* note 31, at 6–7.

[45] U.S. Const. art. I, § 8, cl. 8.

[46] Mazer v. Stein, 347 U.S. 201, 219 (1984).

[47] Douglas Reid Werner, *Digital Audio Recording Technology: Challenges to American Copyright Law*, 22 St. Mary's L.J. 455, 458 (1990) (citing Melville Nimmer, 1 Nimmer on Copyright § 103[A] at 1-32 (1990)).

of sound recordings is a classic "free-rider" problem. Record companies, as the typical copyright owners in sound recordings, and broadcasters will become competitors in the business of distributing sound recordings, if the copyright law does not adapt to the advent of digital technology. As a result, record companies will no longer have an incentive to invest in the creation of new sound recordings or to facilite the creative efforts of their artists because there will be no market for their prerecorded music. The broadcasters will have access to the finished product at negligible cost (the price of one CD) and will be able to distribute the goods more cheaply. Thus, broadcasters will take a "free ride" at the expense of the record companies because they will be able to share in the profits of an investment for which they took no risk.

In fact, the economic incentive to invest in the creation of sound recordings will not merely be shifted to another party, but will be eliminated completely. The broadcasters, competing amongst themselves for the delivery of original audio programmings, would face the same problems that confronted the record companies. An individual broadcaster would have only a limited incentive to invest in the development of new sound recordings because all other broadcasters would be able to distribute the final product freely. Thus, the probable end result of denying a public performance right in sound recordings is that investment in the production of music would drop precipitously.

In the new digital landscape, copyright law must regain its focus and provide incentives for initial investment in creative works. Copyright owners must have the absolute right to authorize distribution of their work if they are to enjoy the protection copyright was intended to afford. As digital technology propels traditional "broadcasting" into the sphere of "distribution," it is even more imperative that copyright owners of sound recordings be given a public performance right. Whereas denying a performance right would virtually eliminate the role of the recording industry, granting a performance right would simply "level the playing field" and allow both broadcasters and copyright owners in sound recordings to profit from advancing digital technology.

The recording industry will certainly be in favor of electronic delivery of their goods because such delivery will eliminate manufacturing and distribution costs.[48] Traditionally less profi-

[48] Al Teller, chairman of MCA Music Entertainment Group, foresees "the record store

table genres, such as jazz, may see a revival in investment by record companies because it will be less expensive for the record companies to deliver individual sound recordings.[49] Current cost levels discourage distribution because of low prospects for recouping initial investment, but a reduction in the cost of delivery may make it profitable for record companies to invest in areas of music with lower demand. Thus, digital technology, if controlled by a performance right, has the "capacity of giving rise to a renaissance of musical production."[50]

A. *The International Front*

The international community has responded to the vital necessity of providing performance rights, both with international conventions and with domestic legislation in foreign countries. The United States, by far the world's leader in the creation of sound recordings, is among a very few developed nations that fail to recognize a performance right in sound recordings. Approximately seventy-five nations, including at least nine European Community member states, grant public performance rights in sound recordings.[51] The majority, including Austria, France, Germany, Japan, Spain, Sweden, Taiwan, and the United Kingdom, already have laws in place specifically according rights in the distribution of sound recordings by cable broadcasters to the owners of sound recording copyrights.[52] Furthermore, the World Intellectual Property Organization (WIPO) and others have commenced a policy debate on the implications of digital technology for international copyright.[53] The WIPO has attempted to resolve those critical issues in the context of a new Protocol to the Berne

---

of tomorrow as an electronic store with an unlimited selection that will be measured in bytes, not tonnage. A store where inventory will be downloaded to order from master digital main frames using the same interactive communication technology some are touting now for home use." Al Teller, *Retailing, Digital Realm Can Co-Exist*, BILLBOARD, Apr. 10, 1993, at 6. For an overview of the record business *see also* DONALD PASSMAN, ALL YOU NEED TO KNOW ABOUT THE MUSIC BUSINESS (1991).

[49] Berman, *supra* note 31, at 2.

[50] *Id.*

[51] RIAA Chart of *Sound Recording Performance Rights Around the World* (1993) (on file with *Harvard Journal on Legislation*).

[52] *Id.* Eighteen countries, including Brazil, India, Portugal, Spain, Taiwan, Turkey, and the United Kingdom, provide sound recording copyright owners the exclusive right to authorize or prohibit the distribution of their copyrighted work by cable distributors. *Id.*

[53] An extensive review of the state of performance rights in sound recordings in international conventions is beyond the scope of this paper. However, the proposed

Convention, the largest international copyright treaty, and by a possible new international instrument. In both instances, the WIPO has recommended that an exclusive right to authorize or prohibit the public performance of sound recordings be granted to copyright owners thereof.[54]

The international market is a lucrative one for the U.S. recording industry. The RIAA estimates that forty percent of sales

---

alterations to the Berne and Rome Conventions proposed by the International Bureau of the WIPO, *infra* note 54, are illustrative.

[54] The Report of the Committee of Experts of the International Bureau of the World Intellectual Property Organization on a Possible Instrument on the Protection of the Rights of Performers and Producers of Phonograms (March 12, 1993 for discussion in Geneva, June 28 to July 2, 1993) recommended the following:

> 28. It is proposed that the instrument include the following definitions:
>
> (b) "phonogram" is an exclusively aural fixation of the sounds of a performance or of other sounds or of the digital representation thereof, regardless of the method by which, and the medium in which, the fixation was made; an audiovisual fixation of the sound part thereof (for example, the sound track of a motion picture) is not a phonogram;
>
> (e) "publication" of a phonogram is (i) offering of copies of a phonogram to the public in a reasonable quantity, or (ii) making the sounds fixed in a phonogram or the digital representations thereof available to the public through an electronic retrieval ("digital delivery") system;
>
> (f) "reproduction" of a phonogram is the making of a copy, in whole or in part, and regardless of the means by which, and the medium in which, the copy is made, of a phonogram, including the storing of its contents, even temporarily, in electronic form;
>
> (j) "public performance" of a phonogram is making the sounds fixed in the phonogram, or the digital representations thereof, audible, by means of any device or process, at a place where persons outside the normal circle of a family and its closest social acquaintances are or can be present, irrespective of whether they are or can be present at the same place and at the same time, or at different places and/or different times, and where the sounds made audible can be perceived without the need for communication to the public within the meaning of (i), above. [(i) refers to the definition of "communication to the public"—ed.].
>
> 56. It is proposed that the instrument, subject to the provisions proposed in paragraphs 57 and 58, below, provide for any performer in respect of his performance fixed in a phonogram and for any producer of phonograms in respect of his or its phonogram, the exclusive right to authorize the following acts:
>
> (a) the reproduction of the phonogram . . .
>
> (f) the public performance of the phonogram.
>
> 57. It is proposed that the instrument provide that
>
> (e) it is a matter for national legislation in the countries party to the instrument to provide for the limitation of the rights mentioned in items (e) and (f) of the preceding paragraph to a right to equitable remuneration;
>
> (f) the faculty provided for national legislation under the preceding item does not apply in the case of digital communication to the public of phonograms.

Committee of Experts on a Possible Instrument on the Protection of the Rights of Performers and Producers of Phonograms, Questions Concerning a Possible Instrument on the Protection of the Rights of Performers and Producers of Phonograms, prepared by the International Bureau 8–10, 17 (March 12, 1993) [hereinafter WIPO Recommendation].

revenues from U.S. sound recordings, or approximately $4 billion, are generated outside the country.[55] Although several foreign countries have performance rights in sound recordings and systems for royalty collection, many of these countries follow a principle of reciprocity with respect to performance royalties. That is, within these nations, performance royalties are only distributed to performers whose "home" country also distributes performance royalties.[56] Thus, because the United States copyright laws do not recognize a performance right in sound recordings, American copyright owners are often unable to share in international royalty pools. The RIAA calculates that in 1991, foreign royalty pools distributed public performance royalties in excess of $120 million.[57]

Reciprocity principles are a form of economic protectionism employed in those countries where imports of copyrighted works far exceed exports.[58] This new wave of protectionism has resulted in the abandonment of the principle of "national treatment," which had been a foundation of international copyright law.[59] A country which employs a policy of "national treatment" simply ensures that benefits granted nationals are extended to aliens. However, the few countries in Europe which still apply "national treatment" are contemplating the abandonment of the policy for one of reciprocity.[60]

Ironically, the United States has always been a leader in the promotion of intellectual property rights, including extensive copyright protections for sound recordings,[61] in the international arena.[62] The lack of a performance right in sound recordings in U.S. copyright law restricts the United States government in discussions and negotiations regarding international trade and copyright.[63] Today's negotiators are expected to hold the incongruous position of seeking more copyright protection in the international arena than is provided at home, similar to the awk-

---

[55] RIAA Comments to FCC, *supra* note 41, at 13.
[56] Berman Testimony, *supra* note 2, at 11.
[57] *Id.* at 11–12.
[58] Robert D. Hadl, Remarks at the World Intellectual Property Organization Worldwide Symposium on the Impact of Digital Technology on Copyright and Neighboring Rights 2 (April 1, 1993) (transcript available in Harvard Law School Library).
[59] *Id.*
[60] William I. Hochberg, *Tracking International Tape Royalties*, ENTERTAINMENT LAW AND FINANCE, Apr. 1993, at 3.
[61] RIAA Comments to Copyright Office, *supra* note 39, at 18.
[62] *Id.* at 17.
[63] *Id.*

ward position of U.S. trade negotiators upon the U.S. government's reluctance to accede to the Berne Convention.[64] Furthermore, the absence of a performance right in sound recordings has allowed some of our trading partners to maintain low levels of protection for sound recordings through such means as providing a short term of protection, broad limitations on the exclusivity of rights, and, of course, the application of reciprocity rather than national treatment.[65]

B. *Broadcasters' Response and Arguments of Equity*

The American Bar Association Section on Patent, Trademark & Copyright Law has noted that, in the face of mounting domestic and international pressure upon the United States to provide a performance right to copyright owners of sound recordings, broadcasters have responded with two primary arguments against such proposals. They are: (1) the added exposure resulting from new digital transmission services will prove beneficial to the musical artists, and (2) performance rights royalties would be too expensive for broadcasters to afford.[66]

The first argument, added exposure, may be questioned on several grounds. Broadcasters assert that imposing a performance royalty would be unfair because the recording industry will experience increased record sales, enjoying the promotional benefits of extended airplay that will occur with the advent of digital audio broadcast and cable services.[67] However, these broadcast services will actually have the opposite effect, leading to a decrease in record sales. Digital audio cable services will be in direct competition with retail record stores for the "sale" of recorded music.[68] Since digital transmissions not only "broadcast" music, but also distribute it to consumers, every digital transmission represents satisfaction of consumer demand and will result in the loss of "real" sales of sound recordings.[69]

Furthermore, even if we assumed that increased exposure

---

[64] *Id.* (citing H.R. Rep. No. 609, 100th Cong., 2d Sess. 18–19 (1988); S. Rep. No. 352, 100th Cong., 2d Sess. 4–5 (1988)).

[65] RIAA Comments to Copyright Office, *supra* note 39, at 18.

[66] Michael O. Sutton, ed., *Annual Report*, 1990–1991 A.B.A. SEC. PAT. TRADEMARK & COPYRIGHT L. 187, 188. [hereinafter *ABA Report*].

[67] RIAA Comments to Copyright Office, *supra* note 39, at 8.

[68] *Id.*

[69] Reply Comments of the Recording Industry Association of America Before the U.S. Copyright Office In the Matter of Copyright Office Request for Comments

through digital audio broadcast and cable services would improve record sales, that would not automatically lead to the conclusion that copyright owners should be denied the right to authorize those transmissions. Nor would such an improvement in record sales justify denying copyright owners compensation for the commercial exploitation of their sound recordings by broadcasters.[70] Every copyright owner should have the ability to authorize or prohibit the reproduction, transmission, or distribution of her work, regardless of how well such acts allegedly "expose" the work. Copyright law recognizes this principle and protects the public performances of all other works except for sound recordings.[71] As one commentator noted, "it can be said that the authors of *The Hunt for Red October* or *The Color Purple* were 'benefitted' by the release of the movie versions of those books, but few would argue that those authors should not also be compensated for these visual performances as well."[72]

All broadcasters, including traditional analog radio stations, should have to compensate copyright owners of sound recordings for the commercial exploitation of these recordings. Copyright law already requires broadcasters to compensate composers, authors, and publishers of songs for the public performance of their copyrighted works. It is an anomaly of the copyright law that the performers of the songs, or the copyright owners of the sound recordings, do not receive compensation from those that wish to commercially exploit their work.

Finally, the broadcasters should recognize the inequity and hypocrisy of this "added exposure" argument. The RIAA notes that during congressional hearings prior to enactment of the 1976 Copyright Act, broadcasters "railed against the 'unfairness' of cable retransmission of broadcast programming despite the cable industry's position that they were helping broadcasters by enhancing and enlarging their commercial advertising coverage areas."[73] Robert V. Evans, then Vice-President and General

---

Regarding Digital Audio Broadcast and Cable Services 11 (Jan. 31, 1991) [hereinafter RIAA Reply Comments to Copyright Office].

[70] *Id.* at 8.

[71] *Id.* at 11.

[72] *ABA Report*, *supra* note 66, at 188.

[73] RIAA Reply Comments to Copyright Office, *supra* note 69, at 8.

> In 1976, broadcasters fought for and won an extension of the public performance right, by way of a compulsory license, to compensate for the commercial expropriation of distant signal programming by cable television . . . in recent years, the broadcast industry has gone even further to protect its interests against what it perceives as the "freeloading" of the cable industry. [In 1991,

Counsel of CBS, Inc., testified to Congress that "it is unreasonable and unfair to let [the cable] industry ride on our backs, as it were, to take our product, resell it, and not pay us a dime. That offends my sense of how things ought to work in America."[74] However, these same broadcasters simultaneously lobbied against a public performance right in sound recordings.

The broadcasters' second argument is that because broadcasters already pay royalties to composers and publishers of pre-recorded music, any additional royalty payments would be too costly.[75] As the RIAA points out, the fact that broadcasters are already paying some copyright owners should have no bearing on the issue of whether owners of sound recording rights deserve to be compensated as well.[76] The real problem with this argument is one of distributive equity. Even without the economic incentive rationale, creators of sound recordings deserve to be recognized as owners of a performance right in their copyrighted works as a matter of fairness.

When we listen to a song by a performing artist whom we admire, we are as much listening to it because of that performer's skill and style as because of the skill and style of the musical composer, lyricist, or music publisher (who all receive performance royalties). Indeed, more often than not, the musician, not the composer, determines the level of airplay for a song.[77] There are many examples of "covers" which became hits in their own right. Some recent examples include Rod Stewart's cover of Van Morrison's "Have I Told You Lately,"[78] and Mariah Carey's cover of The Jackson Five's "I'll Be There."[79] It is clear that a performer's contribution to a sound recording is at least

---

> it successfully lobbied Congress to pass legislation] requiring the cable industry to pay for the retransmission of local signals. Under the banner "if carry, must pay," the broadcasting industry has complained through NAB [National Association of Broadcasters] President Eddie Fritts, that "[c]able operators take a broadcaster's signal and in turn sell it to a consumer for a profit and pay nothing to the broadcaster for the use of that signal. On the other hand, cable systems pay all the cable networks a fee for a similar type service. This inequity is something that broadcasters recognize . . . ."

Id. at 8–10.

[74] *Copyright Law Revision: Hearings on H.R. 2223 Before the Subcomm. on Courts, Civil Liberties, and the Administration of Justice of the House Comm. on the Judiciary*, 94th Cong., 1st Sess. 769 (1975).

[75] *ABA Report*, *supra* note 66, at 188.

[76] RIAA Reply Comments to Copyright Office, *supra* note 69, at 7.

[77] *Id.* at 14.

[78] ROD STEWART, ROD STEWART UNPLUGGED (Warner Brothers Records 1993).

[79] MARIAH CAREY, MARIAH CAREY UNPLUGGED (Columbia Records 1992).

as important as that of the composer, lyricist, or publisher, and yet the performer is not currently entitled to compensation from the public exposure of a successful song.[80]

Further inequities of the lack of a performance right in sound recordings result from the fact that the volume of airplay does not necessarily determine the volume of record sales. The works of several musical styles, including those of old blues artists and performers whose works fall under the major radio formats of "Classic Rock" and "Oldies," continue to be profitable for broadcasters despite slow record sales.[81] Performance royalties would provide income for these artists whose works are being commercially exploited by broadcasters.[82] Also, as noted above, a substantial volume of recorded music consists of genres (such as classical, jazz, and gospel) which operate on a marginal economic basis.[83] Performance royalties collected from continuing airplay would provide an additional source of revenues to offset often unimpressive sales returns, and would promote incentives for increased production of such types of music.[84]

A central feature of copyright protection is that copyright owners are the only persons able to authorize the commercial uses of their works. Society at large benefits the most when the copyright owner is granted the necessary economic incentive to invest in the creation of artistic works. A fundamental concept of copyright law has long been that one may not secure commercial advantage through the uncompensated use of another's

---

[80] RIAA Comments to Copyright Office, *supra* note 39, at 13. This conclusion was noted in the dissenting opinion of Judge Learned Hand in Capitol Records, Inc. v. Mercury Records Corp., 221 F.2d 657 (2d Cir. 1955). Judge Hand stated:

> I also believe that the performance or rendition of a "musical composition" is a "Writing" under Article I, § 8, Cl. 8 of the Constitution separate from, and additional to, the "composition" itself. It follows that Congress could grant the performer a copyright upon it, provided it was embodied in a physical form capable of being copied. The propriety of this appears, when we reflect that a musical score in ordinary notation does not determine the entire performance, certainly not when it is sung or played on a stringed or wind instrument . . . in the vast number of renditions, the performer has a wide choice, depending upon his gifts, and this makes his rendition pro tanto quite as original a "composition" as an "arrangement" or "adaptation" of the score itself . . . . Now that it has become possible to capture these contributions of the individual performer upon a physical object that can be made to reproduce them, there should be no doubt that this is within the Copyright Clause of the Constitution.

Id. at 664.

[81] RIAA Comments to Copyright Office, *supra* note 39, at 15.

[82] *Id.*

[83] *Id.* at 15.

[84] Molnar, *supra* note 22, at 708.

product, that is, one may not "appropriat[e] to [oneself] the harvest of those who have sown."[85] It would be at odds with the basic tenets of copyright law to allow broadcasters to use the recording industry's product to generate revenues without payment for that use.

IV. The Need for a Market Solution

Any proposed legislative solution designed to enable the music industry to function amidst electronic delivery systems and digital audio transmission services should be tested against three criteria: (1) Does it provide an economic incentive to invest in sound recordings? (2) Does it provide for an equitable distribution of profits to copyright owners of sound recordings without stifling the progress of digital technologies? (3) Is it flexible enough to adapt to the wide range of possible media through which digital technological advancement may be realized in the future? This last consideration hints at the impossibility of designing legislation which will specifically apply to each possible implementation of digital audio transmissions. Such legislation would need to be framed in general terms to enable copyright owners of sound recordings to protect their interests against all current and future applications of digital technology.[86] Three possible solutions are generally offered by interested parties, only one of which meets the proposed criteria.

First, broadcasters often propose that the best alternative is to not act at all. Their argument is that the harm to the recording industry does not come from the mere "broadcasting" of the music (this is viewed as an inherent promotional good), but rather from the unauthorized taping by individual recipients of the transmission, a problem already addressed by private copying legislation.[87] This line of reasoning fails all three criteria suggested for judging proposed legislative responses. Private copying legislation, standing alone, hardly suffices as an economic incentive to create sound recordings or as a vehicle for the equitable distribution of profits to copyright owners. Such legislation will generate modest royalty income from the sale of blank audio tape and taping equipment, but was never intended

---

[85] International News Serv. v. Associated Press, 248 U.S. 215, 239–40 (1918).
[86] Garnett, *supra* note 27, at 10.
[87] *Id.* at 11.

to compensate rights holders in sound recordings for the lost revenue threatened by digital technology as the primary medium for the delivery of recorded music.[88] Furthermore, reliance upon private copying legislation completely ignores future adaptations of digital technologies. As previously noted, a fully operational "celestial jukebox" would serve as a personal library of CDs, eliminating the need for tangible copies of recorded music. Thus, the need for record companies to have the legal ability to license broadcasting of sound recordings is not addressed by private copying legislation.

A second solution typically offered to the problem of copyright protection of sound recordings in the digital age is the establishment of a statutorily mandated royalty rate similar to that currently enjoyed by songwriters and music publishers. However, this would be an inadequate solution. Any mandatory statutory royalty rate, by definition, would serve as a ceiling on returns for copyright owners in sound recordings from the digital transmission of their works. Such an arbitrary limit to the upside of investment would dampen the economic incentive to invest in the creation of sound recordings.[89]

Furthermore, it is certain that any government-imposed, static royalty rate of today would be obsolete by tomorrow. Each successive generation of digital technology advances transmission services further from simple broadcasting and closer to the electronic distribution of sound recordings. Government should not establish prices in a competitive marketplace merely because the goods will be delivered electronically rather than by some more physical means. A statutorily mandated royalty system would seriously injure the copyright owner's right to oversee commercial exploitation of her work.[90]

---

[88] Berman, *supra* note 31, at 7.

[89] *Id.* at 5.

[90] *Id.* at 4. Berman further states:

> No one would propose that an individual who acquires a compact disc should be able to manufacture copies using that disc subject only to an obligation to pay remuneration based on his profits, yet many legislators have not had so easy a time in arriving at the same solution when that individual, rather than making copies himself, simply makes the original available through a commercial service for others to do the copying. I am referring to record rental and digital transmission, in which it makes little economic difference to the copyright owner whether the copying is being done by the commercial operator or by the end user—the only thing that the copyright owner knows is that he is no longer selling any records and that his income stream has dried up.

Id.

The third solution is the proper one. The copyright law should be amended to ensure that sound recordings cannot be publicly performed or broadcast without the exclusive authorization of the copyright owner in that sound recording. Such a market approach will be able to adapt rapidly to changing technology. Rights holders in sound recordings are best able to determine which digital transmissions are closest to traditional broadcasting and which are substitutes for the distribution of their works.[91] They will then be able to distinguish between various digital audio cable services according to the degree such services compete for sales of their copyrighted works and to account for such distinctions in the terms of the authorization.[92] This approach also negates the broadcasters' "added exposure" argument; to the extent that the argument has any validity, the market will compensate broadcasters for the promotional value of their service.

Furthermore, only an exclusive right to authorize public performances of sound recordings satisfies the first two requirements for effective legislation. An exclusive right maintains the economic incentive to invest in the creation of sound recordings. The copyright owner will be secure in her control of the distribution of the finished product without fear of "free riders." Also, an exclusive right would "level the playing field" between broadcasters and copyright owners of sound recordings during negotiations over how to exploit new digital technologies. In other words, this approach will allow the market to determine the proper distribution of profits. While broadcasters will be able to seek a fair return on their investment, there will be little chance that broadcasters will receive a windfall at the expense of the

---

[91] *Id.* at 8.

[92] *Id.* Berman further asserts:

> Digital transmissions that are in effect electronic deliveries will be viewed in a different manner than digital services that more closely resemble their existing analog counterparts. It may be, for example, that record companies will attempt to recoup investment for production of recorded music vis-a-vis subscription services, while simply seeking agreement from digital over-the-air broadcasters to not publish schedules or play more than a certain number of tracks from a particular recording.

Id.

Indeed, it is this author's opinion that record companies should treat monies received from the authorization of digital transmissions akin to distribution as they would normal sales of pre-recorded music, and provide the performers of the music with the contractually determined royalty percentage of the amount. Monies received from the authorization of digital transmissions akin to analog broadcasts should be split evenly with performers, much like performance royalties between publishers and songwriters.

returns to investment capital risked by copyright owners. Rather, individual copyright owners who have already risked investment in the creation of sound recordings will receive a return based on fair market value. Finally, and most importantly, an exclusive right to authorize public performance of sound recordings will rectify a glaring inadequacy in the copyright law that fails to recognize the sizable contribution which creators of sound recordings make to the cultural advancement of society.

V. Conclusion

The emergence of digital transmission of recorded music has clearly intensified the debate surrounding public performance rights in sound recordings. Sound recordings captured on digital media may be permanently stored and easily reproduced without any loss in sound quality. This fact has transformed digital transmission services into electronic delivery systems more akin to traditional notions of "distribution" than "broadcasting."

A legislative response to the dangers posed to musicians and the U.S. recording industry by new digital technologies must be made soon, while these transmission services are still in their formative stages. Failure to act now may limit Congress' ability to do so in the future, once there is extensive business investment and consumer reliance on the new technologies. Congress should not allow the "celestial jukebox" to march to the same tune as its ancestor, the standard LP jukebox, did earlier in this century, providing its owners with an economic windfall at the expense of the capital investment and entrepreneurial risks of the sound recording industry.

In the immediate future, digital technology threatens the very livelihood of the U.S. recording industry, a multi-billion-dollar economic sector that provides a very favorable balance of trade. Currently, U.S. artists and record companies risk being excluded from annual royalty pools of over $120 million, growing annually, because the U.S. copyright law does not provide performance rights in sound recordings. Yet as grave as these economic consequences are, the strongest argument to provide a public performance right in sound recordings is one of fairness. It is an unjust denial of the contributions of copyright owners of sound recordings that they cannot control the public perform-

ance of their works, when every other owner of a copyrighted work in this country has that right.

The copyright law should be amended to provide copyright owners of sound recordings an exclusive right to authorize the public performance of their copyrighted works.[93] Any other course of action will be inadequate to address the inequities of the current law and the pressing needs of advancing technology.

[93] Several others have also issued a call for the creation of a public performance right in sound recordings. The inequity of the situation was apparent at the time of the passage of the 1978 Copyright Act. Molnar says: "Thus, it can be seen that congressional recognition of a public performance royalty in sound recordings, mandated by copyright policy . . . remains warranted despite the repeated efforts of private interests to defeat such a proposal." Molnar, *supra* note 22, at 711. *See also* REPORT OF REGISTER OF COPYRIGHTS, *supra* note 24. Furthermore, the American Bar Association Section of Patent, Trademark, and Copyright Law declared in their 1990–1991 Annual Report "that the Section of Patent, Trademark and Copyright Law favors in principle legislation which establishes a performance right in sound recordings and that such right should be in addition to, rather than at the expense of, the existing performance right held by other copyright owners." *ABA Report*, *supra* note 66, at 193.

On the international scene, the World Intellectual Property Organization is taking steps to either insert a performance right in sound recordings in the Berne Convention or establish a new instrument altogether. *See* WIPO Recommendation, *supra* note 54. Also, the International Federation of the Phonographic Industry has called for the introduction of an exclusive right to authorize or prohibit the digital diffusion of sound recordings. Garnett, *supra* note 27, at 11.