HARVARD
Journal on Legislation

Volume 38, Number 2 Summer, 2001

Harvard Journal on Legislation
Publications Center
Harvard Law School
1541 Massachusetts Avenue
Cambridge, MA 02138
(617) 495-4400
hlsjol@law.harvard.edu
www.law.harvard.edu/studorgs/jol

The *Harvard Journal on Legislation* is published semiannually by Harvard Law School students. 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.

Editorial 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.

Submissions: The Journal welcomes unsolicited manuscripts. Interested authors interested should send their manuscripts and all related correspondence to the Articles Editor at the above address. Manuscripts should be typed, double-spaced, and submitted in duplicate, with a diskette if possible. Footnotes should comply with *The Bluebook: A Uniform System of Citation* (17th ed. 2000). The Journal does not accept submissions via e-mail.

Subscriptions: Subscription rates for the 2001 volume (Volume 38) are $28.00 domestic; $34.00 foreign surface mail ($36.00 to Canada); $46.00 foreign airmail. Individual issues from the current volume can be obtained for $14.00 (plus $3.00 per copy for foreign mailing). All correspondence concerning subscriptions should be addressed to the Subscriptions Manager, Publications Center, Harvard Law School, 1541 Massachusetts Avenue, Cambridge, MA 02138, (617) 495-3694. E-mail: ssulliva@law.harvard.edu. 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 38 will be honored only until January 2002.

Back Issues: Past volumes are available exclusively from William S. Hein & Co., Inc., 1285 Main St., Buffalo, NY, 14209, 1-800-828-7571, wsheinco@class.org. Contact William S. Hein & Co. for information on prices for bound volumes.

ESSAY

THE MODERN REGULATORY ADMINISTRATIVE STATE: A RESPONSE TO CHANGING CIRCUMSTANCES

JEFFREY E. SHUREN*

In the landmark case Chevron, U.S.A. v. National Resource Defense Council, Inc., *the Supreme Court articulated a deferential standard for reviewing an agency's interpretations of statutory language within that agency's area of concern as long as Congress was silent or ambiguous on the matter and the agency's interpretation was reasonable. Using the Food and Drug Administration as a primary case study, Dr. Shuren contends that one of the main reasons for granting agencies broad judicial deference in the implementation of statutory mandates is that agencies are the governmental entities best equipped to respond to changing circumstances. Dr. Shuren contends that courts should grant sufficient deference to agencies' modifications of prior statutory interpretations in order to ensure adequate agency flexibility to meet new challenges within existing statutory delegations of authority.*

Congress delegates to administrative agencies the authority to implement many important governmental objectives. As agencies have increased in size and complexity, the courts have been forced to address the extent of agencies' authority to interpret broad statutory language. In *Chevron, U.S.A., Inc. v. Natural Resource Defense Council, Inc.*,[1] the Supreme Court articulated a deferential standard for reviewing an agency's interpretations of statutory language within that agency's area of concern, as long as: (1) Congress was silent or ambiguous on the matter; and (2) the agency's interpretation was reasonable.[2] The Court based its deferential standard on two justifications: agency expertise and electoral accountability.[3] On February 27, 2001, a unanimous Supreme Court in *Whitman v. American Trucking Ass'n*[4] reaffirmed its deferential attitude towards agency rulemaking.[5]

---

* Medical Officer, Office of Policy, Planning, and Legislation, U.S. Food and Drug Administration. B.S., Northwestern University, 1985; M.D., Northwestern University Medical School, 1987; J.D., University of Michigan Law School, 1998. The views expressed in this Essay are those of the Author and do not necessarily reflect the opinion of the U.S. Food and Drug Administration.

[1] 467 U.S. 837 (1984).

[2] *Id.* at 842–44.

[3] *Id.* at 865.

[4] 121 S. Ct. 903 (2001).

[5] The Court held, *inter alia*, that the delegation of authority to the Environmental Protection Agency ("EPA") to set national ambient air quality standards at a level "requisite to

This Essay contends that one of the primary reasons for granting agencies broad judicial deference in the implementation of statutory mandates is that agencies are the governmental entities best equipped to respond to changing circumstances. Indeed, the modern basis for regulatory administrative agencies is to provide a more effective mechanism for the federal government to respond to changing conditions. When viewed in this light, an agency's ability to modify its prior legislative interpretations to address changed circumstances becomes a necessary tool to fulfill congressional mandates.

Since *Chevron*, the Supreme Court has been ambiguous in its approach to revised agency statutory constructions. Lower courts have sometimes deferred to revised agency interpretations, but at other times they have accorded less deference to a modified interpretation than to a statutory construction consistent with the agency's original position. This Essay contends that courts should grant sufficient deference to agencies' modifications of prior statutory interpretations in order to ensure adequate agency flexibility to meet new challenges within existing statutory delegations of authority.

The Food and Drug Administration ("FDA") provides a case study that demonstrates how well agencies can respond to crises and develop innovative solutions to emerging problems. Agencies can respond more rapidly than Congress to crises such as AIDS and have the expertise to tailor solutions that are effective and efficient. Through a discussion of actions taken by the FDA to address new conditions confronting its drug approval process and the bases for these actions, this Essay addresses the legislative, judicial, and practical underpinnings of the FDA's approach to altered conditions as an example of the changing circumstances rationale for giving regulatory agencies broad judicial deference. The history of the FDA's evolving interpretations of its enabling statute—the Federal Food, Drug, and Cosmetic Act of 1938 ("FDCA")[6]—and its 1962 amendments[7] demonstrates a continuous process by which the agency defines the standards for safety and effectiveness that it uses to review a new drug for possible market approval and the criteria to meet those standards consistent with its statutory objectives.

In the years since the 1962 amendments, the FDA has developed several standards and approaches with respect to its pre-approval proc-

---

protect public health" under the Clean Air Act, ch. 360, 69 Stat. 322 (1955) (codified in scattered sections of 42 U.S.C.), was not an unconstitutional delegation of legislative power because the Act contained an "intelligible principle" for setting air quality standards and there was no necessity that the Act set precise upper limits for pollutants. 121 S. Ct. at 907.

[6] *See* Federal Food, Drug, and Cosmetic Act of 1938, ch. 675, 52 Stat. 1040 (codified as amended at 21 U.S.C. §§ 301–395 (2000)).

[7] *See* Drug Amendments of 1962, Pub. L. No. 87-781, 76 Stat. 780 (codified as amended in scattered sections of 21 U.S.C.).

ess.[8] These procedures have included accelerated approval,[9] Phase IV testing,[10] and restricted distribution.[11] This Essay contends that such administrative innovations are an appropriate exercise of statutory authority. Congress has delegated broad authority[12] regarding new drug approvals and the FDA has met its statutory objectives[13] by defining requirements for safety, standards of effectiveness, and criteria by which to measure both. Congress, in turn, has responded in several instances by codifying agency action and transforming implicit jurisdiction into explicit authority.

Part I furnishes a brief overview of the rationales previously offered by commentators for the creation of administrative agencies. It then argues that one of the most important rationales for administrative agencies is that they provide the federal government with a means to address evolving and new conditions. Part II describes the FDA's approach to changing conditions within the historical context of the drug approval process. Part III discusses the judiciary's approach to agency statutory

---

[8] Through the pre-approval process, the FDA determines whether to approve a regulated product for distribution and sale in the United States. The standards and tests used to make this decision differ based on the type of product reviewed, such as drugs and devices. *See* 21 U.S.C. § 355(d) (substantial evidence of safety and effectiveness standard for new prescription drugs); 21 U.S.C. § 360e(d)(2) (reasonable assurance of safety and effectiveness standard for certain devices (Class III) that present a potential unreasonable risk of illness or injury).

[9] When a drug is marked for the "fast-track" procedure, the FDA will facilitate the development and expedite the review of such drug if it is intended for the treatment of a serious and life-threatening condition and it demonstrates the potential to address unmet medical needs for such a condition. 21 U.S.C. §§ 301–395.

[10] A company's agreement with the FDA to conduct post-approval research made by the company before or after approval or as a condition of approval is called a "Phase IV commitment." A "Phase IV study," or post-marketing study, is the study performed to meet the company's Phase IV commitment or any other study performed post-approval. *See* CTR. FOR DRUG EVALUATION & RESEARCH, FOOD & DRUG ADMIN., MANUAL OF POLICIES AND PROCEDURES 6010.2: PROCEDURES FOR TRACKING AND REVIEWING PHASE 4 COMMITMENTS 2–3 (1996).

[11] *See infra* Part II.

[12] Federal Food, Drug, and Cosmetic Act of 1938, ch. 675, 52 Stat. 1040 (codified as amended at 21 U.S.C. §§ 301–395 (2000)). The Act itself may provide a degree of interpretive flexibility. *See, e.g.*, Peter Barton Hutt, *Philosophy of Regulation Under the Federal Food, Drug and Cosmetic Act*, 28 FOOD DRUG COSM. L.J. 177, 178 (1973) (characterizing the act as a constitution that requires the FDA to create specific regulations to meet its "fundamental objectives"). *But see* H. Thomas Austern, *Philosophy of Regulation: A Reply to Mr. Hutt*, 28 FOOD DRUG COSM. L.J. 189, 191–92 (1973) (arguing that the act is not an administrative "blank check").

[13] These goals are to ensure that only safe and effective drugs reach the market and that drugs critical to public health are approved promptly. *See* Drug Amendments of 1962, Pub. L. No. 87-781, 76 Stat. 780 (codified as amended in scattered sections of 21 U.S.C.); Federal Food, Drug, and Cosmetic Act of 1938, Food, Drug, and Cosmetic Act of 1938, ch. 675, 52 Stat. 1040 (codified as amended at 21 U.S.C. §§ 301–395 (2000)); Prescription Drug User Fee Act of 1992, Pub. L. No. 102-571, 106 Stat. 4491 (codified in scattered sections of 21 U.S.C.); *see also* United States v. An Article of Drug, Bacto-Unidisk, 394 U.S. 784 (1969) (stating that the primary objective of the FDCA is the protection of the public health).

interpretations, and concludes that the changing circumstances rationale supports deference to shifting administrative statutory constructions when they are made in response to new conditions equal to the deference accorded the original interpretation.

I. Evolving Rationales for Regulatory Administrative Agencies

The executive branch has undergone remarkable growth in size and authority over the past two centuries. Although administrative agencies have been a facet of the federal government since early in its formation, originally there were few agencies and they were limited in scope.[14] For example, during President George Washington's administration, only three executive departments existed: the Department of State, the Department of the Treasury, and the Department of War.[15]

In the second half of the eighteenth century, executive branch agencies gradually evolved from the ministerial servants of a young nation into discretionary policymakers of an industrial age country.[16] By the 1860s, agencies had come into favor as important government bodies to meet the demands of an ever-growing industrial revolution. Clientele-oriented departments, such as the Agriculture and Commerce Departments, were established to promote the interests of specific economic groups through data collection and research.[17] Only twenty years later, the federal government began to form agencies not only to promote industrial power, but also to control it. In 1887, Congress created the Interstate Commerce Commission, which was the first major federal regulatory body.[18] Acting under the Interstate Commerce Act,[19] Congress created an entity with discretionary authority to issue binding decisions and offer flexible solutions to the outrage sparked by corruption in the railroad industry.[20]

During the early progressive movement from 1906 through 1915, the philosophy of government shifted from a contract regime to a regulatory regime.[21] Agencies of this period were created to regulate the economy

---

[14] James Q. Wilson, *The Rise of the Bureaucratic State*, 41 Pub. Interest 77, 78 (1975) [hereinafter *Bureaucratic State*].

[15] Kenneth F. Warren, Administrative Law in the Political System 39 (1997).

[16] The numbers tell much of the story. At the end of the Federalist Period, there were 3000 civilian officers. By 1881, there were 95,000. *Bureaucratic State*, *supra* note 14, at 77.

[17] *Id.* at 87–88.

[18] *Id.* at 94. The Commission was established to regulate railroad rates. Its creation was rooted in the Granger Movement of the 1870s, a period during which incensed farmers demanded relief from the exorbitant rates charged by unregulated railroads. Milton M. Carrow, Background of Administrative Law 7 (1948).

[19] Interstate Commerce Act, ch. 104, 24 Stat. 379 (1887) (codified as amended in scattered sections of 49 U.S.C.).

[20] *Bureaucratic State*, *supra* note 14, at 95.

[21] *See* Marc Allen Eisner, Regulatory Politics in Transition 33–41 (1993).

and correct market failures.[22] The Progressives' reform agenda incorporated their belief that agency expertise could be used to apply scientific and social-scientific knowledge to address the adverse effects of corporate abuses on the rights of consumers, workers, and small businesses.[23] For example, under the Sherman Anti-Trust Act of 1890, Congress relied on economic expertise to craft United States competition policy.[24] Similarly, in the Meat Inspection Act of 1906, Congress relied on scientific expertise to protect the public from unsafe meats.[25]

With the outbreak of World War I, the federal government took greater advantage of administrative agencies, and the number and size of administrative agencies grew rapidly.[26] By 1925, there were almost 500,000 civil servants.[27] However, the power of these agencies was limited by courts' judicial review of agency action. In turn, the courts' checks on agency action largely shaped the role of administrative agencies and many political leaders agreed that agency power should be limited in scope.[28] For example, President Woodrow Wilson recognized administrative agencies as independent bodies capable of powerful action in his New Freedom program; however, he thought they should execute but not formulate policy.[29]

The New Deal revolutionized the role of agencies and solidified their position as active arbiters and decision-makers. President Franklin Delano Roosevelt promoted economic stability by implementing government-supervised industrial self-regulation.[30] The Roosevelt administration relied on agencies for their expertise and independence and increasingly trusted agencies to invest private entities with public authority.[31]

---

[22] *See* STEPHEN BREYER, REGULATION AND ITS REFORM 15–35 (1982) [hereinafter REGULATION AND ITS REFORM] (discussing various rationales behind economic regulatory efforts); *see also* CARROW, *supra* note 18, at 9 (discussing the 1914 creation of the Federal Trade Commission in response to inadequate judicial enforcement of the Sherman Anti-Trust Act of 1890).

[23] EISNER, *supra* note 21, at 36–38. Opposition to party bosses and an overreaching judiciary also contributed to the rise of administrative power in the early 1900s. Stephen Skowronek, *Building a New American State: The Expansion of National Administrative Capacities, 1877–1920*, *in* FOUNDATIONS OF ADMINISTRATIVE LAW 32 (Peter Schuck ed., 1994).

[24] EISNER, *supra* note 21, at 44.

[25] *Id.*

[26] *Id.* at 78.

[27] *Id.* at 77; *see also* CASS SUNSTEIN, AFTER THE RIGHTS REVOLUTION: RECONCEIVING THE REGULATORY STATE at 242–43 (1990) [hereinafter RIGHTS REVOLUTION] (describing the number of regulatory agencies created during different time periods).

[28] WARREN, *supra* note 15, at 39.

[29] EISNER, *supra* note 21, at 39; Woodrow Wilson, *The Study of Administration*, 2 POL. SCI. Q. 220 (1887).

[30] EISNER, *supra* note 21, at 89.

[31] *Id.* at 111. For example, to return the purchasing power of agricultural commodities to pre-World War I levels through price increases, the Agricultural Adjustment Act of 1933, Pub. L. No. 73-10, 48 Stat. 31 (codified as amended in scattered sections of 7 U.S.C.), created the Agricultural Adjustment Administration ("AAA") to control the supply of agricultural goods through contracts limiting the amount of goods produced or the numbers of acres planted. The AAA also taxed processors to fund the recovery program.

The New Deal's distributional programs also fostered reliance on government agencies as shields against acute economic deprivation by making agencies responsible for organizing industry members in the absence of an existing association, granting them public authority, and supervising industry self-regulation.[32]

President Johnson also made extensive use of administrative agencies in his Great Society program of the 1960s, building upon the New Deal ideology of government responsibility.[33] Congress, too, increasingly supported agency autonomy, pursuing social regulation through the establishment of new regulatory programs and agencies that addressed public health and environmental protection issues.[34] For example, Congress established the Environmental Protection Agency in 1970, authorizing the agency to regulate pollutants in water and the atmosphere through the Clean Water Act[35] and Clean Air Act.[36] Congress also created the Occupational Safety and Health Administration to regulate safety in the workplace.[37]

In the late 1960s and early 1970s, a growing number of voices called for reform of regulatory activities.[38] Critics argued that agencies had ignored their mandates to create clear and consistent policy guidelines, and had opted instead to issue oppressive regulations or develop policy through case-by-case adjudication.[39] To address these concerns, Congress began to enact precise guidelines requiring agencies to undertake specific actions.[40] Simultaneously, courts increased the number and extent of judicial controls to limit regulatory excesses.[41]

Because production levels rarely varied based on the type of commodity, the AAA relied on farm associations to implement the program by providing expertise and information to develop production levels and audit compliance by individual processors. EISNER, *supra* note 21, at 92–93. The administration also believed that agency expertise could be employed to determine and measure the objective public interest. STEPHEN G. BREYER & RICHARD B. STEWART, ADMINISTRATIVE LAW AND REGULATORY POLICY 140 (1992).

[32] *See* EISNER, *supra* note 21, at 111–12.

[33] *See* RIGHTS REVOLUTION, *supra* note 27, at 242–44.

[34] *See id.* at 25–31; REGULATION AND ITS REFORM, *supra* note 22, at app. 1 tbl.13.

[35] Clean Water Act, ch. 758 (1948) (codified as amended in scattered sections of 33 U.S.C.).

[36] Clean Air Act, ch. 360, 69 Stat. 322 (1955) (codified in scattered sections in scattered sections of 42 U.S.C.).

[37] RIGHTS REVOLUTION, *supra* note 27, at 26–27.

[38] *Id.* at 29.

[39] *See* PETER H. SCHUCK, FOUNDATIONS OF ADMINISTRATIVE LAW 43 (1994).

[40] RIGHTS REVOLUTION, *supra* note 27, at 29. For example, Congress passed the National Environmental Policy Act, Pub. L. No. 91-190, 83 Stat. 852 (1970) (codified as amended in scattered sections of 42 U.S.C.), which required agencies to consider the impact of their decisions on the environment.

[41] RIGHTS REVOLUTION, *supra* note 27, at 30; *see, e.g.*, Citizens to Preserve Overton Park, Inc. v. Volpe, 401 U.S. 402 (1971) (holding that to determine whether an agency decision was arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with the law, courts "must consider whether the decision was based on a consideration of the relevant factors and whether there has been a clear error of judgment").

Judicial efforts were inadequate, though, because courts suffered from a lack of expertise and accountability and an extremely slow pace in addressing issues.[42] By the 1980s, Congress had become frustrated by the ineffectiveness of judicial restraints on agency authority, and it began deregulating industries and reducing what it had come to view as burdensome agency oversight.[43] Significantly, one of the greatest criticisms of administrative agencies was that they had become so large and unwieldy that they were no longer responsive, expert agencies.[44] Critics contended that excessive regulation impeded market competition and that agencies often failed to distinguish low-cost from high-cost mandates.[45]

Indeed, when President Clinton took office in 1992, agencies still faced considerable criticism for their perceived excesses. The Clinton administration focused on reforming agencies, but highlighted the importance of allowing administrative agencies to fulfill their important public mandates while removing undue burdens on regulated entities and the public.[46] Through the "Reinventing Government" initiative, the executive branch established a regulatory philosophy that favored eliminating redundant regulation and streamlining agency procedures to create more effective government.[47]

Hence, throughout the history of administrative agencies, there have been competing rationales for agency power. For many decades, administrative agencies were viewed as the answer to the expanding economy's negative externalities. Furthermore, they enabled the federal government to adapt to the demanding needs of the economy. In the mid-twentieth century, agencies were viewed as an imperative mediator between the

---

[42] RIGHTS REVOLUTION, *supra* note 27, at 32.

[43] For example, Congress eased restrictions on the airline industry. *See* EISNER, *supra* note 21, at 117.

[44] Agencies may react slowly due to a lack of incentive to respond, the requirements of due process, or the need for political consensus building. *Bureaucratic State*, *supra* note 14, at 98.

[45] Robert Rabin, *Federal Regulation in Historical Perspective*, *in* FOUNDATIONS OF ADMINISTRATIVE LAW 35 (Peter Schuck ed., 1994).

[46] For a description of major problems that plague agencies in their efforts to regulate significant public health risks, see STEPHEN BREYER, BREAKING THE VICIOUS CYCLE: TOWARD EFFECTIVE RISK REGULATION 10–29 (1993).

[47] President Clinton directed federal agencies to issue only those regulations that "are required by law, are necessary to interpret the law, or are made necessary by compelling public need." 5 U.S.C. § 601 (2000). Regulations must be tailored to impose the least burden on society. Each agency must also periodically review its existing significant regulations to determine whether any should be modified or revoked to increase efficiency or reduce the burden imposed by that agency's regulatory program. *See* Exec. Order No. 12,875, 3 C.F.R. 669 (1993), *revoked by* Exec. Order No. 13,083, 3 C.F.R. 13,083 (1998) (prohibiting executive departments and agencies from imposing unnecessary unfunded mandates on state, local, or tribal governments). For a discussion of how the "Reinventing Government" initiative may undermine the appropriate role of administrative law, see generally Jerry L. Mashaw, *The Structure of Government Accountability: Reinventing Government and Regulatory Reform: Studies in the Neglect and Abuse of Administrative Law*, 57 U. PITT. L. REV. 405 (1996).

federal government and Americans in need of assistance. More recently, they have been subject to widespread criticism, and we have yet to fully emerge from this negative portrayal of administrative agencies.

Agency power needs to be reaffirmed as an important asset for the federal government, not a liability to be tamed. Throughout its history, the federal government has designed and depended on agencies to address changing circumstances and new conditions, including industrialization, the Great Depression, and the civil rights movement. In fact, agency creation and expansion of existing agency authority have tended to occur during periods of national crisis or favorable political conditions, when progressive presidents enjoyed majorities in both houses of Congress.[48]

Proponents of regulatory bodies point to agencies as important mechanisms through which to implement legislation, collect information, and provide feedback to the public, industry, and other interested parties. Once established, however, many agencies have provided government responses to changes not originally contemplated by their authorizing legislation. Such actions offer an important rationale for according deference to changes in agency decision-making as the United States enters the twenty-first century. Agencies can respond effectively and efficiently to new situations to which Congress is not equipped to respond due to lack of time, information, or consensus of opinion to enact legislation.[49]

The Food and Drug Administration stands as an illuminating example of why agencies should be afforded broad discretion to respond to changing circumstances. The FDA confronts sensitive issues, where human lives and health are at stake, and thus provides numerous examples of how an agency can respond more efficiently to changing circumstances than Congress. Indeed, when Congress has taken action with respect to issues under the jurisdiction of the FDA, it was often years after the actual crisis or situation arose and, in several instances, subsequently codified the FDA's decisions.[50]

---

[48] WARREN, *supra* note 15, at 82.

[49] CORNELIUS M. KERWIN, RULEMAKING: HOW GOVERNMENT AGENCIES WRITE LAW AND MAKE POLICY 27–30 (1994).

[50] For example, the Food and Drug Administration Modernization Act of 1997, Pub. L. No. 105-115, 111 Stat. 2296 (codified in scattered sections of 21 U.S.C.), codified many of the regulations the FDA had employed for several years. However, this legislation also streamlined several FDA activities to address perceived deficiencies in those operations and reduce regulatory burdens. For example, section 360e of the Act streamlined the process for device sponsors to make changes to certain manufacturing processes by permitting device manufacturers to notify the FDA 30 days before instituting certain types of manufacturing changes instead of submitting a pre-market approval application supplement, unless the FDA finds the notice to be inadequate. 21 U.S.C. § 360e(d). Section 360c also required the FDA, in consultation with the product sponsor, to consider the least burdensome means that would allow appropriate pre-market development and review of a device without unnecessary delays and expense to manufacturers. 21 U.S.C. § 360c(a)(D)(iii).

II. THE FDA AS A CASE STUDY: RESPONSES TO CHANGING CIRCUMSTANCES

The Food and Drug Administration is responsible for regulating products that account for "twenty-five cents of every dollar spent by the American consumer."[51] As mandated by Congress, the FDA's objective is to promote and protect the public health. In the case of prescription drugs, the FDA operates under two congressionally imposed aims: (1) to protect the public health by approving prescription drugs only if they are safe and effective; and (2) to improve the public health by promptly approving safe and effective drugs. The first goal was established by the FDCA of 1938 as well as amendments made to it in 1962.[52] The second goal was first established by amendments made to the FDCA in 1992.[53]

The requirements imposed on drug manufacturers have evolved since passage of the Pure Food and Drug Act of 1906[54] through a series of congressional enactments, innovative FDA regulations, and judicial case law. The history of the FDA's drug approval provisions reveals the alacrity with which an administrative agency—acting as an agent of the legislative and executive branches—can address new or changed conditions.

A. A Brief History of the FDA

The Pure Food and Drug Act of 1906 ("PFDA") constituted the first significant federal law regulating medicinal drugs. The increasing use of chemicals and uncertain practices in food processing, as well as the subsequent publication of Upton Sinclair's, *The Jungle*, sparked the initial public outrage that led to its passage.[55]

---

[51] Elizabeth C. Price, *Teaching the Elephant to Dance: Privatizing the FDA Review Process*, 51 FOOD DRUG COSM. L.J. 651, 651 (1996) (citing PRESIDENT WILLIAM JEFFERSON CLINTON & VICE PRESIDENT ALBERT GORE, NATIONAL PERFORMANCE REVIEW: REINVENTING DRUG & MEDICAL DEVICE REGULATIONS 2 (1995).

[52] *See* Drug Amendments of 1962, Pub. L. No. 87-781, 76 Stat. 780 (1962) (codified as amended in scattered sections of 21 U.S.C.); Federal Food, Drug, and Cosmetic Act of 1938, ch. 675, 52 Stat. 1040 (codified as amended at 21 U.S.C. §§ 301–395 (2000)); *see also* United States v. An Article of Drug, Bacto-Unidisk, 394 U.S. 784 (1969) (stating that the primary objective of the FDCA is protection of the public health).

[53] *See* Prescription Drug User Fee Act of 1992, Pub. L. No. 102-571, 106 Stat. 4491 (codified at 21 U.S.C. § 379g–h (1992)).

[54] Pure Food and Drug Act of 1906, Pub. L. No. 59-384, 34 Stat. 768 (repealed 1938).

[55] *See* Watson B. Miller, *Introduction to the Act*, 1 FOOD DRUG COSM. L.Q. 290, 291, 296 (1946); *see also* James F. Hoge, *The Drug Law*, 1 FOOD DRUG COSM. L.Q. 48, 48 (1946) (quoting President Theodore Roosevelt's 1905 address to Congress, in which he recommended that "a law be enacted to regulate interstate commerce in misbranded and adulterated foods, drinks and drugs[,] . . . protect legitimate manufacture and commerce, and . . . secure the health and welfare of the consuming public. Traffic in foodstuffs which have been debased or adulterated so as to injure health or to deceive purchasers should be forbidden."). For a discussion of the PDFA's legislative history, see generally Charles Wesley Dunn, *Its Legislative History*, 1 FOOD DRUG COSM. L.Q. 297 (1946).

To address the sanitation crisis, the PFDA created the Bureau of Chemistry in the Department of Agriculture, which was the predecessor of today's FDA.[56] The Bureau of Chemistry had authority to intervene against the sale of misbranded or adulterated drugs in interstate commerce.[57] The PFDA did not, however, grant the Bureau of Chemistry authority to require pre-market evidence of drug safety or effectiveness.[58] The Bureau could only prevent the marketing of an ineffective drug if it could demonstrate that the drug failed to work and that the seller had actual knowledge that its claims were false.[59]

Congress amended the PFDA six times between 1906 and its repeal in 1938,[60] but these amendments proved insufficient to redress the Act's limited delegation of power to the Bureau.[61] To a large degree, the PFDA focused on preventing economic fraud rather than protecting the public from deleterious drugs.[62] It took a public health disaster for Congress to create an agency with broad authority to regulate drugs. In what would be called the "Elixir of Sulfanilamide" fiasco, a manufacturer marketed a liquid drug that killed at least seventy-three people in the two months that it was on the market in 1937.[63] The drug manufacturer used diethylene glycol, commonly used in antifreeze, as the solvent in the new drug. It had never tested diethylene glycol to determine whether human beings could use it safely.[64] In response to this disaster, the Secretary of Agriculture asked Congress for legislation that would require testing of new drugs.[65]

Shortly thereafter, Congress approved the Federal Food, Drug, and Cosmetic Act of 1938.[66] The FDCA authorized the Food and Drug Administration to conduct pre-market approval of drugs. Indeed, under the

---

[56] Anna Kelton Wiley, *Its Great Founder*, 1 FOOD DRUG COSM. L.Q. 314, 323 (1946).

[57] Pure Food and Drug Act of 1906.

[58] Richard A. Merrill, *The Architecture of Government Regulation of Medical Products*, 82 VA. L. REV. 1753, 1761 (1996). The Bureau tried to extend its authority by characterizing false therapeutic claims as violative of the Act, but the Supreme Court struck down this effort in *United States v. Johnson*, 211 U.S. 488, 497–98 (1911) (holding that claims of effectiveness were opinions rather than facts, and hence not subject to agency control); PETER TEMIN, TAKING YOUR MEDICINE: DRUG REGULATION IN THE UNITED STATES 126–27 (1980).

[59] Merrill, *supra* note 58, at 1761.

[60] Vincent A. Kleinfeld, *Legislative History of the Federal Food, Drug, and Cosmetic Act*, 1 FOOD DRUG COSM. L.Q. 532, 533 (1946).

[61] *Id.* at 535.

[62] TEMIN, *supra* note 58, at 44. The 1938 Act eliminated the requirement that the government demonstrate the seller's state of mind before a drug could be deemed misbranded. Merrill, *supra* note 58, at 1762.

[63] Carl M. Anderson, *The "New Drug" Section*, 1 FOOD DRUG COSM. L.Q. 71, 72–73 (1946).

[64] *Id.*

[65] *Id.* at 73 (quoting SEC'Y OF AGRIC., REPORT OF THE SECRETARY OF AGRICULTURE ON DEATHS DUE TO ELIXIR SULFANILAMIDE-MASSENGILL, S. Doc. No. 124 (1937)).

[66] Federal Food, Drug, and Cosmetic Act of 1938, ch. 675, 52 Stat. 1040 (codified as amended at 21 U.S.C. §§ 301–395 (2000)).

FDCA a manufacturer of a "new drug"[67] must conduct tests demonstrating that the drug is safe for the use specified on its label.[68] Subsequently, the manufacturer must notify the FDA before it can provide the drug for sale or consumption. However, manufacturers had the right to determine whether their products were new drugs. If a manufacturer determined that its product was not a new drug, but, instead, "generally recognized as safe and effective," it did not have to submit data to the FDA. In such instances, a drug could be marketed without any FDA review.

The FDCA, as originally enacted, delegated authority to confirm a drug's safety to the FDA; however, it did not specify criteria for assessing safety. Congress left such assessments to the judgment of the newly created FDA. Moreover, although Congress did not require manufacturers to demonstrate their drugs' efficacy, the FDA reviewed therapeutic effectiveness based on an implicit grant of authority under the FDCA. Because all drugs pose some risk, the FDA employed a risk-benefit analysis, asserting that a drug could not be considered safe unless there was some health benefit to outweigh the health risk.[69]

In September 1960, a drug company notified the FDA that it intended to market thalidomide as a new drug. Dr. Frances Kelsey, the examiner reviewing the submission, refused to permit marketing of the drug because its manufacturer had failed to provide sufficient evidence of the product's safety.[70] The decision to keep thalidomide off the shelves of American drug stores turned out to be a critical decision. In the interim, phocomelia, a condition wherein infants are born without hands or feet, arose at high rates in Western Europe. In 1961, thalidomide was identified as the cause for the outbreak. The company sponsoring the drug withdrew its notification in March 1962.[71]

Following the thalidomide tragedy in Europe and the FDA's role in preventing a corresponding disaster in the United States, Congress was inspired to give the FDA greater authority over the regulation of new drugs, including the investigation of drugs.[72] In the 1962 Kefauver-Harris

---

[67] The FDCA defined a "new drug" as "any drug the composition of which is such that such drug is not generally recognized, among experts qualified by scientific training and experience to evaluate the safety of drugs, as safe for use under the conditions prescribed, recommended, or suggested in the labeling thereof . . . ." 21 U.S.C. § 321.

[68] These safety provisions have changed little since their enactment. Robert Temple, *Development of Drug Law, Regulations, and Guidance in the United States*, *in* PRINCIPLES OF PHARMACOLOGY: BASIC CONCEPTS AND CLINICAL APPLICATIONS 1643 (Paul L. Munson et al. eds., 1995).

[69] *See Hearings Before a Subcomm. of the House Comm. on Gov't Operations*, 88th Cong. 2 (1964) (statement of George Larrick, Comm'r, U.S. Food and Drug Admin.), *reprinted in* PETER BARTON HUTT & RICHARD A. MERRILL, FOOD AND DRUG LAW 522–24 (2d ed. 1991).

[70] Carol R. Goforth, *A Bad Call: Preemption of State and Local Authority to Regulate Wireless Communication Facilities on the Basis of Radiofrequency Emissions*, 44 N.Y.L. SCH. L. REV. 311, 373 (2001).

[71] *See* TEMIN, *supra* note 58, at 123.

[72] HUTT & MERRILL, *supra* note 69, at 452.

Amendments, also known as the Drug Amendments of 1962, Congress explicitly granted the FDA the authority to review a drug's effectiveness.[73] More importantly, these amendments made effectiveness a separate criterion for approval,[74] and they specified the kind of information manufacturers were required to submit to demonstrate effectiveness: data from "adequate and well-controlled studies."[75]

Under the original FDCA, the FDA engaged in an assessment of relative risk, balancing benefit with risk. Though the FDA still assesses relative risk when determining whether to approve a new drug, under the 1962 Kefauver-Harris Amendments the FDA cannot approve an ineffective drug even if it poses no health risks. The new effectiveness standard marked an important advance from policies mandated by the PFDA of 1906.[76]

The 1962 Kefauver-Harris Amendments also transformed the pre-market notification system created by the FDCA into a pre-market approval system. The notification system permitted a manufacturer to sell its drug 180 days after submitting the product's new drug application if the FDA failed to raise an objection. In contrast, the approval system prohibited a manufacturer from selling its drug until the FDA had affirmed the product's safety and effectiveness.[77] As a result, Congress shifted the burden of proof from the FDA to the manufacturer. Congress also delegated to the FDA the authority to withdraw the approval of a new drug if the FDA determined, based on new information, that the drug was no longer safe or effective for its purported use.[78]

In 1962, therefore, Congress transformed the FDA from an agency with limited powers to one that was to exercise a great deal of discretion over which drugs would be allowed on the American market. Congress had given the FDA an enormous mandate: the FDA would have the final decision as to which drugs were safe and effective enough to reach the market. As the following examples demonstrate, the FDA has aggressively pursued this mandate, modifying its policies and regulations to meet the demands of changing circumstances and conditions. Moreover, Congress has signaled its approval of many of the FDA's activities by passing legislation to codify those procedures. Thus, Congress has stepped aside to allow the FDA to respond first to changing circumstances. Congress retains the ability, however, to step into the debate whenever it wishes and override an agency action.

---

[73] *See* Drug Amendments of 1962, Pub. L. No. 87-781, 76 Stat. 780 (1962) (codified as amended in scattered sections of 21 U.S.C. (2000)).

[74] Congress accordingly expanded the statutory definition of "new drug" to include drugs not generally recognized by experts as safe and effective. *See* 21 U.S.C. § 321.

[75] 21 U.S.C. § 355(d).

[76] TEMIN, *supra* note 58, at 125–26.

[77] Merrill, *supra* note 58, at 1764–65.

[78] *See* 21 U.S.C. § 321.

B. Advancements in Clinical Trials: The FDA Adjusts Its Clinical Testing Requirements

At the same time the Kefauver-Harris Amendments of 1962 gave the FDA new authority, they also created great challenges for the FDA with respect to its mandate to protect the public health. Rather than wait for congressional action to clarify the predicament, the FDA responded to these challenges by modifying FDA policy. The 1962 Kefauver-Harris Amendments contained an investigational new drug ("IND") process that effectively transformed the FDA's role from a reviewer of data to an active participant in the drug development process. The amendments authorized the FDA to supervise the clinical testing of drugs and create standards under which health care professionals could obtain a drug for investigational use before the product received FDA approval.[79] Because a drug cannot be distributed in interstate commerce until approved by the FDA, the amendment required the FDA to issue regulations exempting the use of new drugs by qualified experts solely for investigational use prior to market approval.[80] Under the IND requirements, a manufacturer must submit an investigational plan that includes its research protocols for human subjects before conducting clinical trials.[81] The FDA may stop clinical investigations that pose unreasonable risks or that do not accord with sound scientific procedures.[82]

The amended Act required that the manufacturer demonstrate "substantial evidence" of effectiveness before the FDA could approve the company's drug.[83] Substantial evidence is defined as "evidence consisting of adequate and well-controlled investigations" by qualified experts on the basis of which such experts could determine whether the drug studied has the effect it is represented to have under the "conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling."[84] The FDA later issued a final rule describing the features of an adequate and well-controlled clinical study.[85]

When ascertaining whether a drug causes a particular effect, the possibility exists that the benefits or harms produced by the drug may, in fact, be accounted for by other factors, such as the natural course of the

---

[79] 21 U.S.C. § 355(i).

[80] *Id.*

[81] HENRY G. GRABOWSKI & JOHN M. VERNON, THE REGULATION OF PHARMACEUTICALS: BALANCING THE BENEFITS AND RISKS 4 (1983).

[82] *Id.* For additional changes implemented by the 1962 Kefauver-Harris Drug Amendments, see Alan H. Kaplan, *Fifty Years of Drug Amendments Revisited: In Easy-to-Swallow Capsule Form*, FOOD & DRUG L.J. 179, 181–86 (1995).

[83] 21 U.S.C. § 305(d).

[84] *Id.*

[85] Applications for FDA Approval to Market a New Drug, 21 C.F.R. § 314.50(f)(1) (2001). A U.S. District Court subsequently upheld the rule in *Pharmaceutical Manufacturers Ass'n v. Richardson*, 318 F. Supp. 301, 311 (D. Del. 1970).

disease, other treatment, or investigator or patient expectations.[86] To distinguish between the true effect of the drug and the effects of other factors, clinical trials typically have been designed so that one group of subjects receives the drug under investigation while a second group with similar characteristics—the control group—receives no drug or a placebo—an inert substance.[87]

Based on an emerging consensus among academic clinicians on the essential characteristics of clinical studies,[88] the FDA identified four different types of controls to be used in clinical trials: placebo control,[89] no treatment control,[90] active control,[91] and historical control.[92] The FDA later modified these regulations to add a fifth type, dose comparison concurrent control.[93] The FDA's change in policy reflected a growing experience with clinical trial design and a recognition of the problems resulting from failure to execute adequate dose response studies.[94] In short, the FDA assessed a change in scientific standards and integrated that assessment into its policy. The Third Circuit Court of Appeals upheld the change, recognizing the importance of allowing the FDA discretion to fulfill its mission.[95] The court held that the FDA needed leeway to assess "effectiveness" as the FDCA did not define the term in any detail.[96] Indeed, one of the most important reasons for such deference is that agencies are best equipped to update their policies based on scientific advancements.

---

[86] INTERNATIONAL CONFERENCE ON HARMONISATION, ICH HAROMONISED TRIPARTITE GUIDELINE: CHOICE OF CONTROL GROUP AND RELATED ISSUES IN CLINICAL TRIALS 1, 2 (July 20, 2000) [hereinafter ICH].

[87] Temple, *supra* note 68, at 1651.

[88] Merrill, *supra* note 58, at 1771.

[89] Placebo control indicates that the control group receives a placebo rather than the drug. *See* ICH, *supra* note 86, at 4.

[90] No-treatment control indicates that the control group does not receive the drug, a placebo, or any other treatment for the condition under study. *Id.*

[91] Active control indicates that the control group receives some known effective treatment, usually because it would be harmful to withhold care. Applications for FDA Approval to Market a New Drug, 21 C.F.R. § 314.126(b)(2)(iv) (2001).

[92] Historical control indicates that the control group is a group of patients external to the study who received treatment at an earlier time or in whom the natural course of the conditions was followed. 21 C.F.R. § 314.126(b)(2)(i).

[93] *Id.* §§ 314.126(b)(2)(i), 314.50. Dose comparison concurrent control indicates that the control group receives a randomly determined fixed dose of the drug under investigation. *See* ICH, *supra* note 86, at 4.

[94] *Id.*

[95] Warner-Lambert Co. v. Heckler, 787 F.2d 147, 155 (3d Cir. 1986). Applying the *Chevron* two-step analysis, the court upheld the FDA's decision that effectiveness required a showing of clinical significance. The Supreme Court had previously supported the FDA's reliance on well-established scientific principles defining the term. *See, e.g.*, Weinberger v. Hynson, Wescott & Dunning, Inc., 412 U.S. 609, 617 (1973) (upholding the FDA's "summary judgment" procedure).

[96] *Id.*

C. *Consequences for Generic Drugs*
The FDA Responds with the Abbreviated New Drug Application

A second challenge that the Kefauver-Harris Amendments presented stemmed from their award to the FDA of the authority to approve new drugs. Under the amendments, the FDA's pre-market approval authority extended to all drugs not generally recognized as safe and effective ("GRAS") by qualified experts for the uses described on their labels.[97] The amendments, then, charged the FDA with reviewing all new drug applications approved prior to 1962 and requiring manufacturers of those drugs to submit substantial evidence of effectiveness if their products were not generally recognized as safe.[98] If the FDA found that a drug was not effective, then it could rescind that drug's approval and remove it from the market. This posed a significant problem for the FDA because prior to 1962 approximately 4000 drugs had been approved by the FDA for marketing (drugs covered by a new drug application), but ten times that number had been sold without formal agency approval. The FDA determined that many of these latter drugs (generic drugs) were generally recognized as safe because they contained the same active ingredient as a drug covered by a new drug application ("pioneer" drugs).[99] Administrative precedent had established that once the FDA or the scientific community considered a pioneer drug as safe, any copy of that drug could be marketed without FDA clearance.[100]

While the amendments authorized the FDA to remove drugs from the market that it had approved; they did not give the FDA explicit authority over those generic drugs that it had allowed on the market as mimics of pioneer drugs that had been approved under the pre-Kefauver-Harris Amendment standards.[101] Congress, thus, had left the FDA in a bind. The FDA's mandate was to ensure the safety and effectiveness of drugs, yet there were potentially hundreds of ineffective generic drugs on the market. On the one hand, it had given the FDA a strong mandate to remove unsafe drugs from the market. On the other hand, Congress offered no effective approach to address generic drugs. The FDA might have challenged the manufacturers of generic drugs in court, but the time involved in such litigation would have been prohibitive. Also, a litigation-based approach would have delayed for many years the withdrawal of ineffective products from the market.[102]

---

[97] Drug Amendments of 1962, Pub. L. No. 87-781, 76 Stat. 780 (codified as amended in scattered sections of 21 U.S.C. (2000)).

[98] HUTT & MERRILL, *supra* note 69, at 478.

[99] *Id.*

[100] Kaplan, *supra* note 82, at 182.

[101] Merrill, *supra* note 58, at 1773.

[102] *Id.*

Instead, the FDA created its own procedural solution to the dilemma, adjusting its policies to deal with the consequences of its actions. First, the FDA resolved the problem by viewing all generic drugs as legally bound to their respective pioneer drugs.[103] As a result, if the FDCA required the manufacturer of the pioneer drug to prove effectiveness, the same standard applied to generic copies.[104] In *USV Pharmaceutical Corp. v. Weinberger*,[105] the Supreme Court upheld the FDA's schema. By embracing generic drugs within the regulatory framework for new pioneer drugs, the FDA could assure the public that all marketed prescription drugs would be safe and effective for their intended use.

The FDA's new policy, however, resulted in an obstruction to the approval of generic drugs. Pioneer drug-makers enjoyed extended periods of monopoly following the expiration of their products' patents while generic drug-makers waited for approval. To address this problem, the FDA issued regulations in 1969 that brought generic drugs under the FDCA's safety and effectiveness requirements and reduced generic drug manufacturers' duplicative testing burden.[106] The FDA established the abbreviated new drug application ("ANDA") as a substitute for submitting the full safety and effectiveness data required for a new drug application. Under an ANDA, a copy of a pre-1962 approved drug would receive FDA approval if the copy demonstrated bioequivalence[107] and possessed the same bioavailability[108] as the pioneer drug.[109] The procedures

---

[103] *Id.*

[104] The FDA performed the required re-evaluation of these drugs through a contract with the National Research Council of the National Academy of Sciences. Temple, *supra* note 68, at 1653. The Council required companies marketing drugs and antibiotics that had been introduced before 1962 to submit data supporting their claims of effectiveness. *See* Reports of Information for Drug Effectiveness, 31 Fed. Reg. 9425, 9426 (July 9, 1966); Antibiotic Drugs: Reports of Information for Drug Effectiveness, 31 Fed. Reg. 13014 (Oct. 6, 1966).

[105] 412 U.S. 655, 664 (1973) (holding that generic drugs covered by an NDA were not exempt from the efficacy requirements imposed by the 1962 Drug Amendments).

[106] The regulations appeared to serve what had been the principal goal pursued by Senator Kefauver (D-Tenn.) in his investigations leading to the 1962 amendments: to decrease drug prices through increased competition. HUTT & MERRILL, *supra* note 69, at 576. The FDA's efforts to increase the availability of generic drugs also may have stemmed from the federal government's self-interest in ensuring that there are low-cost drugs available for purchase by Medicaid and Medicare beneficiaries. *Id.*

[107] FDA regulations define bioequivalence as "the absence of a significant difference in the rate and extent to which the active ingredient or active moiety in pharmaceutical equivalents or pharmaceutical alternatives becomes available at the site of drug action when administered" in the same amount and under the same conditions. Bioavailability and Bioequivalence Requirements, 21 C.F.R. § 320.1(e) (2001).

[108] FDA regulations define bioavailability as "the rate and extent to which the active ingredient or active moiety is absorbed from a drug product and becomes available at the site of action." Bioavailability and Bioequivalence Requirements, 21 C.F.R. § 320.1(a).

[109] 21 C.F.R. § 130 (2001); 21 C.F.R. § 320. At the same time, the FDA prohibited generic drug-makers from gaining access to pioneer drug manufacturers' safety and effectiveness data on the grounds that such data constituted trade secrets or confidential business information. 21 C.F.R. § 20 (2001).

did not apply to post-1962 new drugs, which were still required to demonstrate safety and effectiveness independently. In the early 1980s, the FDA implemented a concept known as "paper NDAs," under which copies of post-1962 new drugs could support their safety and effectiveness claims using clinical or pre-clinical published studies.[110]

The FDA's abbreviated new drug application proved controversial. After fifteen years of dispute, though, Congress stepped in and codified the FDA's procedures. Moreover, Congress extended the reach of the ANDA process to include generics for pioneer drugs approved after 1962. The Drug Price Competition and Patent Term Restoration Act of 1984 ("Waxman-Hatch Act")[111] provided that a generic drug need only demonstrate its bioavailability and bioequivalence to the pioneer drug to receive approval.[112] The Waxman-Hatch Act also granted three new statutory protections to pioneer drug-makers to compensate for regulatory delays caused by the 1962 amendments: five-year freedom from generic competition ("market exclusivity") for the pioneer drug, non-patent exclusivity under certain circumstances, and patent extensions up to five years for NDAs.[113]

Regardless of whether Congress anticipated the problems involving generic drugs created by the 1962 Kefauver-Harris Amendments, it had not provided the FDA with guidance on how to resolve them. The FDA responded to the new problems generated by the statutory paradox in a novel way that ensured that unsafe drugs were removed from the market, while safe and effective generic drugs remained available. Far from condemning the FDA for its innovation and initiative, Congress embraced the FDA's efforts by codifying and modifying the FDA's procedures, as well as by establishing new incentives that the agency lacked the authority to create.

---

[110] Kaplan, *supra* note 82, at 189.

[111] Drug Price Competition and Patent Term Restoration Act of 1984, Pub. L. No. 98-417, 98 Stat. 1585 (codified as amended at 35 U.S.C. § 156 (2000)).

[112] The Waxman-Hatch Act maintained the distinction between an ANDA and a paper NDA but imposed uniform requirements. *See id.*

[113] The Act embodies a compromise between pioneer and generic drug manufacturers that was supported by both parts of the industry. Merrill, *supra* note 58, at 1793. Professor Merrill notes, however, that not all major brand name manufacturers supported the Act. *Id.* at 1793 n.125. For a more detailed description of the Act, see generally Kaplan, *supra* note 82, at 189–90; HUTT AND MERRILL, *supra* note 69, at 571–74; Suzan Kucukarslan & Jacqueline Cole, *Patent Extension Under the Drug Price Competition and Patent Term Restoration Act of 1984*, 49 FOOD & DRUG L.J. 511 (1994).

D. Bringing Important Treatments to Market: The FDA Institutes New Mechanisms for Drug Approval

Congress's new effectiveness and safety standards created other significant dilemmas for the FDA. To meet these standards, manufacturers must perform extensive clinical trials. When a manufacturer develops a drug that potentially treats a life-threatening or serious condition for which effective therapy does not exist, such testing may delay access to the medication and leave patients without adequate treatment. Conversely, the FDA's approval of the new medication without sufficient data or safeguards may expose patients unnecessarily to a dangerous or ineffective product. As new public health concerns have arisen and scientific technology and experience have evolved, the FDA has sought to develop creative solutions within the scope of its statutory jurisdiction. To promote and protect the public health, the FDA may expedite the approval of promising medicines while restricting their distribution or requiring that their sponsors perform additional studies.[114] In other instances, a new drug may not qualify for expedited approval, yet additional information about the product would benefit the public health. In such cases, the FDA has sought new means to expand the public's understanding of marketed products.

1. Long-Term Studies—The Case of Levodopa

Levodopa, developed in the 1960s, was the first drug with the potential to treat Parkinson's disease effectively. The FDA understood the dire need for its release to the public, but it had reservations about the long-term use of the drug, which would be inevitable given the chronic nature of the illness. However, because it was not in the public's interest to delay the availability of drugs until such long-term studies could be completed, the FDA explored options to expedite market approval for Levodopa. The FDA modified its regulations in order to provide access to this important medication while adequately assessing the safety and effectiveness of long-term use of products like Levodopa that may be administered on a chronic basis. Under the new scheme, the FDA conditioned approval of Levodopa on the promise that the sponsor would perform a post-marketing (Phase IV) study to assess long-term safety and effectiveness concerns.[115]

---

[114] Of course, these products still must meet the requisite standards for safety and effectiveness under the FDCA. Applications for FDA Approval to Market a New Drug, 21 C.F.R. § 314.500–.520(a)(1) (2001).

[115] 21 C.F.R. § 130.48(a) (1973). The agency based its decision on section 355(e) of the FDCA, which mandated withdrawal of an NDA if the FDA learns that a drug is no longer safe or effective, and section 355(k), which required companies to establish and maintain such records and report to the FDA such information as the agency would need to determine whether an approved NDA should be withdrawn under section 355(e). The FDA has explicitly stated this rationale as the basis to require Phase IV studies for methadone

In the case of Levodopa, the FDA stated that some drugs must be used for long periods of time due to the condition for which they are intended, and that gauging the long-term effects of such drugs would require extensive testing following approval. As a result, the FDA approved Levodopa two to three years earlier than it would have otherwise. In 1974, the FDA issued regulation 21 C.F.R. § 310.303 under which the agency could condition approval on the commitment to perform long-term studies and maintain records of those studies in certain circumstances. The FDA has not employed this rule since 1976. Instead, the agency has sought post-marketing data for certain products through several other Phase IV mechanisms[116] to address a variety of safety and effectiveness concerns in addition to long-term effects.

The FDA faced similar public health choices with respect to drugs that are effective but potentially dangerous if misused. Although the label of an approved drug may specify limitations on how or to whom that product should be administered to ensure its safe use, health care practitioners may administer the drug without adhering to those restrictions. In such cases, the FDA is faced with the dilemma of either requesting that an already approved drug be withdrawn from the market or determining that a new drug effective for its intended use should not be approved because of the risks of misuse. To ensure patients access to important medications while protecting them from exposure to unnecessary harm, the FDA has sometimes required that the distribution of a drug be restricted as a condition of its approval or as a basis for continuing its approval.

In 1972, the FDA promulgated a rule withdrawing approval of all methadone[117] new drug applications due to lack of substantial evidence that methadone was safe and effective under the conditions of use existent at that time. The rule, however, permitted approved hospital pharmacies to dispense methadone for analgesic and antitussive[118] purposes and allowed distribution to certain maintenance treatment programs. The FDA reasoned that under section 355(d) of the FDCA[119] methadone's

---

(1972) and for accelerated approval based on surrogate endpoints. 37 Fed. Reg. 6,940 (Apr. 6, 1972); 37 Fed. Reg. 26,790 (Dec. 15, 1972); 57 Fed. Reg. 13,234 (Apr. 15, 1992); 57 Fed Reg 58,942 (Dec. 11, 1992).

In most other instances, the FDA has requested but not required Phase IV commitments. *But see* Mattison & Barbara W. Richard, *Postapproval Research Requested by the FDA at the Time of NCE Approval, 1970–1984*, 21 DRUG INFO. J. 309 (1987) (in reviewing post-marketing research requested by the FDA for new molecular entities approved between 1970 and 1984, the authors found that the agency's requests commonly included phrases that approval is "contingent" or "conditioned" on a commitment to perform the studies).

[116] *See infra* text accompanying notes 125–152.

[117] Methadone is a long-acting narcotic used to treat addiction to opiate drugs.

[118] Antitussives are cough suppressants.

[119] Section 355(d) prohibits approval of an NDA unless adequate data establishes that the drug is safe for use as stated in its labeling. 21 U.S.C. § 355(d) (2000).

pattern of misuse demonstrated that the product was unsafe for approval unless its distribution was restricted.[120]

In *American Pharmaceutical Ass'n v. Mathews*,[121] the District of Columbia Court of Appeals held that the FDA had exceeded its statutory authority. The court interpreted the word "safe" as requiring the labeling to include the possible risks discovered through drug testing. Methadone, therefore, was safe for its intended use even though the possibility existed that it could be used in an unintended manner.

The FDA asserted that the *American Pharmaceutical Ass'n* case did not mean that it lacked the statutory authority to impose such limitations on use. The FDA noted that the District of Columbia Court of Appeals recognized that restricted use, such as restrictions to a prescription-only basis, are sometimes necessary to ensure that persons who intend to use the drug consistent with its label can do so. The court distinguished the FDA's methadone restrictions, because the FDA promulgated these limitations to control the misuse of methadone by persons who intended to use the drug for non-medical reasons and, therefore, did not implicate the safety issues contemplated by section 355 of the FDCA.

Following the court's reasoning, the FDA persisted in using restricted distribution as a way to ensure that potentially dangerous drugs could reach those patients who truly needed it. In fact, in 1990, the FDA approved Sandoz's new drug, Clozaril, after Sandoz voluntarily agreed to restrict the distribution of Clozaril, an antipsychotic known to cause agranulocytosis.[122] The FDA has stated that its approval of Clozaril was not conditioned on Sandoz's plans to limit distribution to patients who had their white blood cell count monitored weekly. However, the agency indicated that it would consider on a case-by-case basis whether any distribution by the company outside that stated on the label would constitute misbranding of the drug.

In the 1980s, the FDA's ability to react to national emergencies under the 1962 Kefauver-Harris Amendments was truly tested. The onset of the AIDS epidemic challenged the FDA to provide the public with access to important medications while protecting them from exposure to unnecessary harm. By the mid-1980s, newly developed HIV tests became available and demonstrated that many individuals carried the virus who were not yet symptomatic.[123] Potential treatments, however, were a number of years away from approval in the United States.[124] To speed the

---

[120] *Id.*

[121] 530 F.2d 1054 (D.C. Cir. 1976).

[122] Agranulocytosis refers to the lack of production of white blood cells.

[123] John J. Smith, *Science, Politics, and Policy: The Tacrine Debate*, 47 FOOD & DRUG L.J. 511, 513 (1992).

[124] *Id.*

marketing of such products, the FDA responded with several initiatives.[125]

In 1992, the FDA issued a final rule permitting accelerated approval of certain new drugs[126] and biologics,[127] whose safety and effectiveness had not been established through traditional studies.[128] To be eligible, the drug or biologic must treat a serious or life-threatening disease and offer meaningful therapeutic benefit[129] over products currently available on the market.[130] The rule furthermore conditioned approval on: (1) the performance of post-marketing studies of the drug's or biologic's clinical benefits; or (2) the restriction of use or distribution where such restrictions prove necessary for the safe use of the product.

The FDA would review a drug or biological product under its accelerated approval process: (1) when presented with evidence of the product's impact on a surrogate endpoint[131] that is reasonably likely to predict clinical benefit or evidence of a clinical endpoint other than survival or irreversible morbidity; or (2) when a drug proven effective can only be used safely if the agency placed limitations on its use or distribution. In the former case, to compensate for the lack of certainty, the applicant must perform clinical studies to confirm the product's clinical benefit and to determine, if still unknown, the relation of the surrogate endpoint to the clinical benefit or the witnessed clinical benefit to the final outcome.[132]

To counterbalance rapid approval, the FDA could employ an expedited withdrawal procedure. Under this procedure, the FDA could quickly remove a drug approved under the accelerated approval process from the market if post-marketing studies did not demonstrate clinical benefit, the applicant failed to perform post-marketing studies, the restrictions on use or distribution proved insufficient to assure safe use or the applicant violated these restrictions, promotional materials were false

---

[125] *See generally* Sheila R. Shulman & Jeffrey S. Brown, *Food and Drug Administrations's Early Access and Fast-Track Approval Initiatives: How Have They Worked?*, 50 FOOD & DRUG L.J. 503 (1995).

[126] Application for FDA Approval to Market a New Drug, 21 C.F.R. § 314.500–.560 (2001).

[127] 21 C.F.R. § 601.40–.46 (2001).

[128] *See* 21 C.F.R. § 314.500–.560.

[129] Meaningful therapeutic benefits may include the potential for greater efficacy or a lower side-effect profile. Shulman & Brown, *supra* note 125, at 514.

[130] *See* 21 C.F.R. §§ 314, 601.

[131] The accelerated approval rule defined a surrogate endpoint as "a laboratory or physical sign that is used in therapeutic trials as a substitute for a clinically meaningful endpoint that is a direct measure of how a patient feels, functions, or survives and that is expected to predict the effect of the therapy." 57 Fed. Reg. 132,234 (Apr. 15, 1992).

[132] 21 C.F.R. § 314.500. Applicants also had to submit to the FDA all promotional material prior to approval if intended for dissemination within 120 days following approval as well as submit all promotional material following approval if intended for distribution after 120 days post-approval. 21 C.F.R. §§ 314.550, 601.45.

or misleading, or other evidence suggested that the drug was not safe or effective for its prescribed uses.[133]

Critics contend that the FDA does not have the authority to condition prescription drug approval on the performance of Phase IV studies. However, several FDCA provisions, the FDA's long-standing use of Phase IV studies,[134] and later congressional action suggest that prior to 1997, when Congress passed the Food and Drug Administration Modernization Act, the FDA possessed implicit statutory authority to require post-marketing studies for new drugs and biologics[135] and to restrict distribution, at least under certain circumstances.[136]

The FDA's 1992 accelerated approval regulations[137] provided that the FDA could restrict the distribution and use of drugs approved under this process if necessary to maintain the safe and effective use of the drug. The FDA based its decision on the same rationale it invoked in the case of methadone.[138] It also relied on section 352 of the FDCA, which states that a drug is misbranded if it is dangerous to health when used consistent with its label,[139] and section 351, which states that a drug is adulter-

---

[133] 21 C.F.R. §§ 314.530(a), 601.43(a).

[134] In general, such research endeavors to provide new information, confirm existing data, or raise new questions. OFFICE OF INSPECTOR GEN., DEP'T OF HEALTH AND HUMAN SERV., POSTMARKETING STUDIES OF PRESCRIPTION DRUGS 1 (1996). These studies vary widely and may include post-marketing surveillance to ascertain additional safety data or testing to determine dosing, bioavailability, or the effect of the drug in special populations. Mattison, *supra* note 115, at 313–14.

[135] 21 C.F.R. § 314.500–.560. Under section 355(d) of the FDCA, the Secretary of Health and Human Services will approve a new drug if it meets the safety and effectiveness criteria specified in the Act and its implementing regulations. 21 U.S.C. § 355(d) (2000). Section 355(d) describes the criteria for safety and effectiveness. An applicant must provide "substantial evidence" that the drug produces the effect intended. 21 U.S.C. § 355(e). Substantial evidence means "adequate and well-controlled" studies by qualified experts on the basis of which such experts could "fairly and responsibly" conclude that the drug will produce the effect intended. *Id.* The Secretary, however, must withdraw approval of a new drug application if he or she receives new information demonstrating the lack of safety or effectiveness of the drug. 21 U.S.C. § 355(e). To facilitate this endeavor, applicants must create and maintain records and submit reports to the Secretary of information that would permit the Secretary to determine whether a drug should be removed from the market for lack of safety or effectiveness. 21 U.S.C. § 355(k)(1). Under the congressional grant of authority to enact regulations to carry out the efficient enforcement of the Act's provisions, and in an effort to effectuate the legislative mandate to promulgate rules requiring the record keeping and reporting of data necessary to determine whether to withdraw a new drug application approval, the FDA believed it possessed the authority to require additional studies to investigate the clinical impact of a drug approved on the basis of its effect on a surrogate endpoint. 21 U.S.C. § 371(a) (2001); 21 C.F.R. § 314.510. Consistent with this view, the FDA had a long-standing practice of conditioning approval on the completion of post-marketing studies.

[136] The FDA may condition the approval of a biologic on performance of post-marketing studies when the correlation between the surrogate endpoint and the clinical benefit remains uncertain. 21 C.F.R. § 314.500.

[137] 21 C.F.R. § 314.520.

[138] *See supra* text accompanying notes 121–122.

[139] 21 U.S.C. § 352.

ated if the methods or controls used for holding the product do not ensure that the drug is safe.[140]

In 1997, Congress codified the FDA's accelerated approval process as part of the Food and Drug Administration Modernization Act ("FDAMA").[141] The Senate Labor and Human Resource Committee argued that the statutory fast-track system[142] went beyond codifying existing regulations and brought certain categories of drugs under fast-track that the FDA had not previously included.[143]

Although FDAMA did not explicitly authorize the FDA to restrict the distribution and use of drugs approved under the fast-track process, Congress did not prohibit the FDA from imposing such limitations. The Supreme Court has held that congressional inaction supports congressional acceptance of an agency's statutory interpretation.[144] Taking its silence as implicit approval, the FDA invoked its authority under 21 C.F.R. § 314.520 in 1998 to restrict distribution and use of thalidomide as a condition of its approval as a treatment of erythema nodosum leprosum, a severe skin condition secondary to leprosy.[145] The FDA, thus, continues to modify its rules to adapt to changing circumstances.

4. Phase IV Commitments

Phase IV studies are among the strongest tools the FDA has developed to respond to changing circumstances. The FDA requests Phase IV studies in instances where it believes that the drug under review is safe

---

[140] 21 U.S.C. § 351.

[141] Food and Drug Administration Modernization Act, Pub. L. No. 105-115, 111 Stat. 2309 (1997) (codified in scattered sections of 21 U.S.C.).

[142] 21 U.S.C. § 356(b)(2).

[143] *"Fast Track" Law "Goes Beyond" Accelerated Approval Reg–Senate Cmte.*, PINK SHEET, Aug. 10, 1998, at 7.

[144] *See, e.g.*, Young v. Cmty. Nutrition Inst., 476 U.S. 972, 983 (1986)

> But in revisiting § 346, Congress did *not* change the procedures governing unintentionally added substances like aflatoxin. This failure to change the scheme under which the FDA operated is significant, for a 'congressional failure to revise or repeal the agency's interpretation is persuasive evidence that the interpretation is the one intended by Congress.'

(quoting NLRB v. Bell Aerospace, Co., 416 U.S. 267, 275 (1974)); United States v. Rutherford, 442 U.S. 544, 554 (1979) ("Such deference is particularly appropriate where . . . an agency's interpretation involves issues of considerable public controversy, and Congress has not acted to correct any misperception of its statutory objectives."); *accord* Nationsbank of N.C. v. Variable Annuity Life Ins. Co., 513 U.S. 251, 260 (1995) ("The Comptroller has concluded that the federal regime is best served by classifying annuities according to their functional characteristics. Congress has not ruled out that course, *see Chevron*, 467 U.S. at 842; courts, therefore, have not cause to dictate to the Comptroller the state-law constraint VALIC espouses.").

[145] FOOD AND DRUG ADMININSTRATION, FDA APPROVES THALIDOMIDE FOR HANSEN'S DISEASE SIDE EFFECT, IMPOSES UNPRECEDENTED RESTRICTIONS ON DISTRIBUTION (Talk Paper, 1999), *available at* http://www.fda/gov/bbs/topics/ANSERS/ANS0087.html.

and effective and, thus, approvable, yet where FDA staff believe unanswered questions remain regarding information that may need to be addressed in the label to permit the drug to be used or prescribed safely. The FDA requests Phase IV studies when reviewing a drug product not subject to the fast-track approval process.[146] Such investigations provide useful information. For example, the beneficial impact of lipid-lowering drugs on heart disease was demonstrated in Phase IV, not pre-approval, studies.

The FDA seeks to resolve lingering scientific questions through an open dialogue with pharmaceutical companies. The FDA is involved in the design and implementation of drug trials at an early stage in an effort to facilitate the generation of appropriate data to make a reasoned determination whether that drug should be approved. This ongoing process continues over many years and involves bilateral input, feedback, and negotiation between the FDA and the company. Sometimes at the end of this process, questions arise due to unexpected results of Phase III clinical trials.[147]

In other instances, a company may have generated sufficient data for its NDA to be submitted to the FDA for review, however new scientific knowledge and changing views within the scientific community persuade the FDA that safety or effectiveness criteria should be added or modified. The FDA may believe that additional studies to meet these new criteria would provide useful information for prescribers and consumers. Nevertheless, the lack of this data may not prove sufficient to delay the approval of a drug that would benefit the public.[148] Moreover, if the FDA

---

[146] Phase IV studies are commonplace. According to a 1996 Department of Health and Human Services report, the number of Phase IV commitments for new molecular entities has steadily increased. OFFICE OF THE INSPECTOR GEN., *supra* note 134, at 5. New molecular entities contain an active ingredient never before approved in the U.S. CENTER FOR DRUG EVALUATION AND RESEARCH, FOOD AND DRUG ADMINISTRATION, 1999 REPORT TO THE NATION 6 (1999). In the 1970s, 33% of NDAs for new molecular entities carried Phase IV commitments. *Id.* at 5. By the 1990s, the number had grown to a high of 70%, *id.*, with an average of 53% of all NDAs between 1993 and 1998. Personal communication with the Division of Data Management and Services, Center for Drug Evaluation and Research, Food and Drug Administration (Oct. 2, 1998).

[147] For example, when the FDA reviewed Redux for the long-term management of obesity, data was presented demonstrating that high doses of the drug in animals caused persistent neurochemical changes. Meeting of the Food and Drug Administration Endocrinologic and Metabolic Drugs Advisory Committee (Sept. 28–29, 1995), *available at* http://www.fda.gov/ohrms/dockets/ac/95/3107T1.pdf.

Since Redux had been used in Europe for ten years without observed clinical evidence of neurologic impairment in people, the FDA believed that any adverse neurologic effects would be minimal and did not outweigh the benefits of Redux to warrant non-approval. Nevertheless, its original label contained a description of the results of the animal studies as well as a statement that the relevance of the animal findings to humans was unknown. Although a commitment to investigate the significance of the animal data was not a requirement for the approval of Redux, Interneuron and Wyeth-Ayerst agreed to conduct a Phase IV neuropsychological study to answer this question. If the study demonstrated no harmful effect, the label would be changed to provide that information to prescribers.

[148] Pharmaceutical companies may initiate Phase IV studies without an FDA request or

and a drug-maker agree on a trial design for a new drug, the FDA cannot change the agreement after testing begins except in certain limited situations.[149]

In recognition of the FDA's practice of requesting Phase IV studies and the importance of these evaluations, the FDAMA requires a drug-maker who enters into an agreement with the FDA to report to the FDA annually on the study's progress or on the reasons for the sponsor's failure to conduct the study. Each year the FDA publishes a report on the status of these investigations in the Federal Register.[150]

Since the 1962 Kefauver-Harris Amendments, the FDA has modified the approval process to keep pace with medical advancements and address new public health issues. It initiates changes in its policies in order to meet the large and important goal of ensuring that safe and effective drugs reach the market. The FDA has acted consistent with its role as an agent of the legislative and executive branches by responding effectively to changing conditions.

The courts have not been unanimous in their deference to the FDA's policy modifications and changes.[151] Nonetheless, a series of important Supreme Court decisions provide solid legal authorization for agencies to reform and modify their interpretations of their organic statutes.[152] The next Part discusses the importance of judicial deference to agency action and describes the modern judicial approach taken by the Supreme Court towards agency statutory constructions.

III. JUDICIAL APPROACH TO AGENCY STATUTORY INTERPRETATIONS

Federal agencies must operate within their statutory authority, but the jurisdictional boundaries are often unclear. In some circumstances, statutes are written broadly, and the details are left to an administrative agency to flesh out. Since *United States v. Curtiss-Wright Export Co.*,[153] courts have consistently upheld congressional delegations of authority to

---

the FDA may request such evaluations. If the agency makes such a request, it must be cleared by the director of the reviewing division of the Offices of Drug Evaluation in the FDA's Center for Drug Evaluation and Research, unless the drug is a new molecular entity. If it is a new molecular entity, such a request must be cleared by the Director and Center for Drug Evaluation and Research. Regardless, the company enters into discussions about the request prior to making any commitments. In some instances, the FDA has modified or forgone its requests based on discussions with the company.

[149] 21 U.S.C. § 355(b)(4)(c) (2000). Such agreements may not be changed except with the written agreement of the drug-maker or if the director of the appropriate FDA reviewing division decides that "a substantial scientific issue essential to determining the safety or effectiveness of the drug has been identified after the testing has begun." *Id.*

[150] 21 U.S.C. § 356(b).

[151] *See, e.g.*, Am. Pharmaceutical Ass'n v. Mathews, 530 F.2d 1054 (D.C. Cir. 1976) (holding that the FDA had exceeded its statutory authority).

[152] *See supra* note 144.

[153] 299 U.S. 304 (1936).

administrative agencies, provided that Congress established standards to guide the implementation of governmental policies.[154]

During the New Deal era, the Supreme Court granted agency interpretations wide latitude.[155] Although the passage of the Administrative Procedure Act ("APA")[156] in 1946 represented Congress's dissatisfaction with the excesses of administrative power and the need for judicial constraints on administration, the Supreme Court continued to defer to agency interpretations when they were reasonable.[157] For example, the Court wrote, "[a]fter invalidating in 1935 two statutes as excessive delegation, . . . we have upheld, again without deviation, Congress' ability to delegate power under broad standards."[158]

Lacking the time, information, or consensus of opinion to draft definitive legislation,[159] Congress leaves it to the appropriate agency to interpret and implement a statute. Therefore, the crucial issues for the judiciary have been when and to what extent courts should defer to agency interpretations.

The Supreme Court established a deferential standard[160] in 1984 in the landmark *Chevron* case.[161] In *Chevron*, the Supreme Court held that a court must use a two-step process when reviewing an agency's interpre-

---

[154] WARREN, *supra* note 15, at 90.

[155] *See, e.g.*, Gray v. Powell, 314 U.S. 402, 411–12 (1941) (deferring to an agency's interpretation of law, the Court held that "[w]here, as here, a determination has been left to an administrative body, this delegation will be respected and the administrative conclusion left untouched").

[156] 5 U.S.C. §§ 551–559, 701–706 (2000).

[157] *See, e.g.*, Udall v. Tallman, 380 U.S. 1, 16–18 (1965) (holding that the Secretary of the Interior's interpretation of its authority to issue oil and gas leases was reasonable); Mitchell v. Budd, 350 U.S. 473, 480 (1956) (upholding the Wage and Hour Administrator's determination that persons employed at tobacco-bulk plants are covered by the Fair Labor Standards Act's exemption provisions). *See generally* Robert A. Anthony, *Which Agency Interpretations Should Get Judicial Deference?—A Preliminary Inquiry*, 40 ADMIN. L. REV. 121, 122–24 (1988).

[158] Mistretta v. United States, 488 U.S. 361, 373 (1989) (holding that congressional delegation to the Sentencing Commission of the power to promulgate sentencing guidelines for every federal criminal offense did not violate the separation of powers principle or the non-delegation doctrine).

[159] KERWIN, *supra* note 49, at 27–30.

[160] Whether *Chevron* actually increased judicial deference towards agency interpretations is controversial. *See, e.g.*, Richard J. Pierce, Jr., *Chevron and its Aftermath: Judicial Review of Agency Interpretations of Statutory Provisions*, 41 VAND. L. REV. 301, 302 (1988) (arguing that *Chevron* has had a dramatic impact on courts' approach to agency statutory constructions); Russell L. Weaver, *Some Realism About* Chevron, 58 MO. L. REV. 129, 130–32 (1993) (suggesting that *Chevron* did not significantly change the judiciary's approach).

[161] In the first eight years following the *Chevron* decision, the case had been cited in over 2000 federal court opinions. WARREN, *supra* note 15, at 85. For a discussion of *Chevron* as a judicial move towards positivism, see John G. Osborn, *Legal Philosophy and Judicial Review of Agency Statutory Interpretation*, 36 HARV. J. ON LEGIS. 115 (1999). The literature discussing the *Chevron* decision is extensive and conflicting. *See, e.g.*, Sanford N. Caust-Ellenbogen, *Blank Checks: Restoring the Balance of Powers in the Post-*Chevron *Era*, 32 B.C. L. REV. 757 (1991) (suggesting that *Chevron* is unconstitutional).

tation of a statute.[162] First, the court must ask whether "Congress has directly spoken to the precise question at issue,"[163] because both the court and the agency "must give effect to the unambiguously expressed intent of Congress."[164] If the court determines that the statute is silent or ambiguous on the specific issue in question, the court must uphold an agency's reasonable interpretation.[165] The court may not substitute its own judgment for that of the agency.[166] Additionally, the court need not determine that the agency's construction is the only permissible interpretation, nor the one the court itself would have reached if empowered to decide.[167] The Supreme Court observed that, in the past, it had consistently deferred to administrative statutory constructions when they involved "reconciling conflicting policies," and when a "full understanding" of the statutory policy had required the expert knowledge of the agency.[168]

In *Chevron*, the Supreme Court reviewed the Environmental Protection Agency's ("EPA") interpretation of the Clean Air Act Amendments of 1977.[169] The amendments required states that had not achieved the EPA's national air quality standards ("nonattainment states")[170] to establish permit programs for "new or modified stationary sources" of air pollution.[171] To obtain a permit, stringent requirements had to be met.[172] Prior to 1979, the EPA interpreted "source" as any pollution-emitting device in a plant.[173] Therefore, a plant had to obtain a permit any time it sought to add or modify such a device. In 1979, the EPA indicated that it would adopt a "plantwide definition of source."[174] Under this "bubble

---

[162] *Chevron, U.S.A., Inc. v. Natural Resource Defense Council, Inc.*, 467 U.S. 837, 842–44 (1984).

[163] *Id.* at 842.

[164] *Id.* at 843. The Supreme Court indicated that congressional intent should be sought by using the "traditional tools" of statutory construction. *Id.* at 843 n.9.

[165] *Id.* at 843–44. The language in *Chevron* lends itself to opposing interpretations of the two-step test. Under one view, whenever there is any ambiguity in the statutory provision, courts should defer to the agency. *See, e.g.*, Smiley v. Citibank, 517 U.S. 735 (1996). Under the second view, unless the statutory provision is so utterly ambiguous that Congress must have intended to delegate statutory interpretation to the agency, courts should not defer to the agency. *See, e.g.*, INS v. Cardoza-Fonseca, 480 U.S. 421, 449 (1987). Under the latter interpretation, the courts would rarely reach the second *Chevron* step.

[166] *Chevron*, 647 U.S. at 843.

[167] *Id.* at 843 n.11.

[168] *Id.* at 844; *accord* Capital Cities Cable, Inc. v. Crisp, 467 U.S. 691, 699 (1984) (quoting United States v. Shimer, 367 U.S. 374, 383 (1961)).

[169] *Chevron*, 467 U.S. at 837; Clean Air Act Amendments of 1977, Pub. L. No. 95-95, 91 Stat. 685 (1977) (codified as amended at 42 U.S.C.).

[170] Section 109 of the Clean Air Act Amendments of 1970, Pub. L. No. 91-604, 84 Stat. 1679 (1970), mandated that the EPA promulgate National Ambient Air Quality Standards. *See Chevron*, 647 U.S. at 840.

[171] *Id.*

[172] *Id.*

[173] *Id.* at 853.

[174] *Id.* at 855. The EPA promulgated regulations incorporating this interpretation in 1981. *See* 40 C.F.R. § 51.165(a)(1)(i)–(ii) (1999).

concept," a plant could increase the amount of pollution produced by one device as long as the total amount of pollution emitted by the plant did not increase.[175] The D.C. Circuit Court of Appeals held that the plant-wide definition violated the Clean Air Act because it was contrary to the objectives of the non-attainment program.[176]

The Supreme Court reversed. Applying a two-step analysis, the Court concluded that neither the statutory language nor the legislative history spoke to the precise issue of whether the term "source" covered a plant-wide definition.[177] As Congress had not explicitly foreclosed the EPA's interpretation, the Supreme Court next addressed whether the agency's decision was reasonable. The Court deferred to the EPA's definition of a source, finding the "bubble concept" a "permissible construction of the statute."[178] The Court was not dissuaded from its decision by the fact that the agency had changed its interpretation.[179] Instead, it held that to engage in informed rulemaking, the agency needed to revisit the wisdom of its policy on a continuing basis.[180] The fact that the EPA employed different definitions in different circumstances supported the agency's view that the definition was flexible.[181]

The *Chevron* Court based its principle of deference to agency statutory interpretations on two justifications: (1) agency expertise; and (2) electoral accountability.[182] First, agencies possess expertise in specific fields as the result of their specialized personnel and fact-finding capabilities.[183] Agencies, therefore, have greater competence than courts in addressing ambiguities that involve technical and factually complex issues.[184] Second, the resolution of a statutory ambiguity may involve policy making. Since the President is accountable to the public, executive agencies possess greater democratic legitimacy than the judiciary.[185] Al-

---

[175] *Chevron*, 467 U.S. at 840.

[176] *Id.* at 841–42. The appellate court viewed the purpose of the non-attainment program as the improvement of air quality. *Id.*

[177] *Id.* at 859–62.

[178] *Id.* at 866.

[179] A court may accord greater deference if the agency's interpretation is consistent and long-standing, Securities and Exchange Comm'n v. Sloan, 436 U.S. 103, 117 (1978), or if the agency's decision is a "contemporaneous construction" of the statute, Aluminum Co. of Am. v. Central Lincoln Peoples' Util. Dist., 467 U.S. 380, 390 (1984) (citing Udall v. Tallman, 380 U.S. 1, 16 (1965)).

[180] *Chevron*, 467 U.S. at 863–64.

[181] *Id.* at 864.

[182] *Id.* at 865.

[183] *Cf.* Baltimore Gas & Elec. Co. v. Natural Res. Def. Council, Inc., 462 U.S. 87, 103 (1983) (stating that a reviewing court should be most deferential when the agency's construction falls within its special expertise).

[184] It is generally accepted that *Chevron* applies to the review of all agency statutory constructions, whether the agency performed a legislative or a judicial function. *See* Bernard Schwartz, *Administrative Law Cases During 1996*, 49 ADMIN. L. REV. 519, 542 (1997).

[185] *Chevron*, 467 U.S. at 865. This same rationale also applies to independent agencies, which are accountable to Congress.

though courts address questions of law, agencies are better suited to address questions of policy.

As the *Chevron* court explained, "[j]udges are not experts in the field, and are not part of either political branch of the Government. Courts must, in some cases, reconcile competing political interests, but not on the basis of the judges' personal policy preferences."[186] In contrast, because of their expertise and accountability, agencies are particularly well designed for making various kinds of policy decisions.[187] Since statutory interpretations may necessitate an examination of policy and practical knowledge of real world operations, functions better suited to agencies,[188] the Supreme Court held that *reasonable* agency interpretations should be accorded deference.

Alternatively, courts could require Congress to be clear by strictly interpreting statutory language and striking down ambiguous statutes. Such an approach would place the burden on Congress to resolve controversial issues up front rather than defer to agencies. However, this view would hamper the legislature's ability to respond to changing or new conditions. Legislative inertia slows and restricts congressional decisionmaking. It is in part to compensate for this limitation that Congress establishes and relies on federal agencies. Moreover, limiting Congress to drafting very clear legislative mandates, if politically feasible, would drastically narrow the scope and longevity of its statutes. In today's world of rapid change, it would be unrealistic to expect Congress to legislate in response to every new event that cannot be adequately addressed by previously enacted specific statutory language. Instead, Congress can play a more effective role by intervening in the less frequent circumstance where agencies act in a way that is inconsistent with congressional intent.

---

[186] *Id.*

[187] In *Chevron*, the Court explained that:

> an agency to which Congress has delegated policy-making responsibilities may, within the limits of that delegation, properly rely upon the incumbent administration's views of wise policy to inform its judgments. While agencies are not directly accountable to the people, the Chief Executive is, and it is entirely appropriate for this political branch of the Government to make such policy choices—resolving the competing interests which Congress itself either inadvertently did not resolve, or intentionally left to be resolved by the agency charged with the administration of the statute in light of everyday realities . . . . [F]ederal judges—who have no constituency—have a duty to respect legitimate policy choices made by those who do.

Id. at 865–66.

[188] *Cf.* Pension Benefit Guar. Corp. v. LTV Corp., 496 U.S. 633, 651–52 (1990) (holding that "the judgments about the way the real world works . . . are precisely the kind that agencies are better equipped to make than are courts. This practical agency expertise is one of the principal justifications behind *Chevron* deference.").

In *Smiley v. Citibank*,[189] the Supreme Court unanimously elaborated that the courts should grant deference to agencies under *Chevron* because of a presumption that when Congress left a statutory ambiguity, it understood that the agency would resolve the ambiguity and "desired the agency (rather than the courts) to possess whatever degree of discretion the ambiguity allows."[190]

Since at least the Progressive era, Congress has recognized that legislative imprecision is sometimes necessary when addressing complicated regulatory questions.[191] Aside from legislative design, statutory ambiguities may also arise as a result of evolving or new conditions. As Professor Cass Sunstein noted, "*Chevron* is a salutary recognition of a large-scale shift in the allocation of authority within American institutions. It embodies, in those applications, a plausible reconstruction of congressional desires and a sound understanding of the comparative advantages of agencies in administering complex statutes."[192]

While the *Chevron* analysis is deferential to agency interpretations, it also contains several internal protections to avoid legitimizing agency overreach.

A. Chevron *Safeguards*

Although *Chevron* articulates a principle of deference, its two-step analysis is not an abdication of judicial oversight. The Supreme Court has occasionally used the *Chevron* analysis to reject agency interpretations.[193] While it is uncertain whether *Chevron* has had a significant im-

---

[189] 517 U.S. 735 (1996).

[190] *Id.* at 741; *see also* Antonin Scalia, *Judicial Deference to Administrative Interpretations of Law*, 1989 DUKE L.J. 511, 516–17 (1989) (the *Chevron* rationale presumes that, in the case of an ambiguity, Congress intended to delegate authority to the agency).

[191] *See* EISNER, *supra* note 21, at 44–45.

[192] Cass R. Sunstein, *Law and Administration After* Chevron, 90 COLUM. L. REV. 2071, 2077 (1990) [hereinafter *Law and Administration*].

[193] *See, e.g.*, FDA v. Brown & Williamson Tobacco Corp., 529 U.S. 120, 159–60 (2000) (holding that Congress did not grant the FDA the authority to regulate tobacco products); Nat'l Credit Union Admin. v. First Nat'l Bank & Trust Co., 522 U.S. 479, 503 (1998) (finding that the NCUA's interpretation of section 109 of the Federal Credit Union Act was impermissible because it was contrary to unambiguous congressional intent); Maislin Indus., U.S., Inc. v. Primary Steel, Inc., 497 U.S. 116, 130 (1990) (holding that the Interstate Commerce Commission's Negotiated Rates policy was invalid because it was inconsistent with the Supreme Court's previous interpretation of the Interstate Commerce Act); Dole v. United Steelworkers of Am., 494 U.S. 26, 41–42 (1990) (refusing to defer to the Office of Management and Budget's interpretation of the Paperwork Reduction Act of 1980, 44 U.S.C. §§ 3501–3511, because "the statute, as a whole, clearly expresses Congress' intention"); Pittston Coal Group v. Sebbon, 488 U.S. 105, 115 (1988) (holding that Health, Education, and Welfare's interim regulations implementing a provision of the Black Lung Benefits Reform Act of 1977, 30 U.S.C. § 902(f)(2), violated the express meaning of the statute); Bd. of Governors of the Fed. Reserve Sys. v. Dimension Fin. Corp., 474 U.S. 361, 374–75 (1986) (ruling that the Federal Reserve Board's regulation defining "banks" was inconsistent with the language of section 2(c) of the Bank Holding Company Act of 1956, 12 U.S.C. § 1841(c)).

pact on the Supreme Court's practice,[194] lower courts have applied *Chevron* more consistently than the Supreme Court.[195]

Chevron provides several safeguards to prevent administrative abuse of statutory construction. First, an agency can only interpret a statute that it has been delegated the authority to administer.[196] Second, the act must be silent or ambiguous on the specific issue.[197] If Congress explicitly addressed the issue, the congressional mandate must be followed.[198] Third, the agency's interpretation must be reasonable.[199] The authority to address these threshold questions—whether the agency possesses the authority to regulate under the act, whether the statute is silent or ambiguous, and whether the agency's construction is permissible—resides with the courts. Under *Chevron*, the judiciary, therefore, still engages in statutory interpretation as a threshold matter. What *Chevron* changes is whose interpretation should be accorded greater weight when the court determines that the legislative provision is silent or ambiguous on the specific issue in question.

Of course, *Chevron* is not the sole check on agency action; agency interpretations remain subject to congressional oversight. Throughout the

---

[194] *See, e.g.*, Thomas Merrill, *Judicial Deference to Executive Precedent*, 101 YALE L.J. 969, 984 (1992) (finding that the Court often did not invoke *Chevron* or follow its analysis, and, when it did, it only deferred to the agency's interpretation in 59% of the cases as compared to 75% of the cases pre-*Chevron*).

[195] *See* Mark Seidenfeld, *A Syncopated* Chevron*: Emphasizing Reasoned Decisionmaking in Reviewing Agency Interpretations of Statutes*, 73 TEX. L. REV. 83, 94 (1994).

[196] *Chevron*, 467 U.S. at 842; *see also* Adams Fruit Co. v. Barrett, 494 U.S. 638, 649 (1990) (stating that the Court did not have to defer to the Secretary of Labor's interpretation of section 1854 of the Migrant and Seasonal Agricultural Worker Protection Act because Congress had appointed the judiciary as the adjudicator of private causes of action under the Act); West v. Bowen, 879 F.2d 1122, 1137 (3d Cir. 1989) (holding that "[n]o deference is owed an agency's interpretation of another agency's statute"); *cf.* United States v. Rutherford, 442 U.S. 544, 553 (1979) ("[T]he construction of a statute by those charged with its administration is entitled to substantial deference."); CIBA Corp. v. Weinberger, 412 U.S. 640, 643 (1973) ("A decision that FDA lacks authority to determine in its own proceedings the coverage of the Act it administers, subject of course to judicial review, would seriously impair FDA's ability to discharge the responsibilities placed on it by Congress."). Greater deference should be shown when an agency interprets its own regulation rather than a statute. Gen. Carbon Co. v. Occupational Safety & Health Review Comm'n, 860 F.2d 479, 483 (D.C. Cir. 1988) (citing Udall v. Tallman, 380 U.S. 1, 16 (1965)).

[197] *Chevron*, 467 U.S. at 843.

[198] *Cf.* 62 Cases of Jam v. United States, 340 U.S. 593, 600–01 (1951) (overruling an FDA interpretation because it was contrary to the explicit language of the Food, Drug, and Cosmetic Act).

[199] *Chevron*, 467 U.S. at 843–44. In determining whether, under *Chevron*, the agency's interpretation is a "permissible construction of the statute," the court may look at the statutory language, "the legislative history, the agency regulations adopted to implement the statute, and the agency comments made with respect to the regulations." West v. Sullivan, 973 F.2d 179, 185 (3d Cir. 1992); *accord* Skidmore v. Swift, 323 U.S. 134, 140 (1944) ("The weight [accorded an agency's] judgment in a particular case will depend upon the thoroughness evident in its consideration, the validity of its reasoning, its consistency with earlier and later pronouncements, and all those factors which give it power to persuade, if lacking power to control.").

1990s, for example, Congress closely scrutinized the activities of the FDA through hearings[200] and, in response to changing circumstances, modified and broadened the agency's authority.[201]

In effect, congressional oversight represents an additional protection at the second step of the *Chevron* analysis. If Congress disagrees with an agency's statutory construction, even though it is a reasonable interpretation, Congress can override the agency's policy through new legislation.

B. *Judicial Approach to Revised Agency Interpretations*

The Supreme Court's application of *Chevron* to revised agency statutory constructions has been inconsistent.[202] Although most courts have deferred to revised agency interpretations, in several cases courts have accorded less deference to a modified interpretation than to a statutory construction consistent with the agency's original position, required an adequate justification for the change, or incorporated language into the opinion that cast doubt on whether a revised interpretation should be shown deference equal to that granted to the original construction.

---

[200] *See, e.g., Gene Therapy: Promoting Patient Safety: Hearing Before the Subcomm. on Public Health of the Senate Health, Education, Labor, and Pensions Comm.*, 106th Cong. (2000); *FDA Modernization Act: Implementation of the Law: Hearing Before the Senate Health, Education, Labor, and Pensions Comm.*, 106th Cong. (1999); *Dietary Supplement Health and Education Act: Is the FDA Trying to Change the Intent of Congress?: Hearing Before the House Comm. on Gov't Reform*, 106th Cong. (1999); *Improving the Safety of Food Imports: Hearing Before the Permanent Subcomm. on Investigations of the Senate Comm. on Governmental Affairs*, 105th Cong. (1998); *Reauthorization of the Prescription Drug User Fee Act and FDA Reform: Hearing before the Subcomm. on Health and Environment of the House Comm. on Commerce*, 105th Cong. (1997); *FDA Reform Legis.: Hearings on H.R. 3199, 3200, and 3201 Before the Subcomm. on Health and Environment of the House Comm. on Commerce*, 104th Cong. (1996); *FDA User Fees for Prescription Drug Approvals: Hearing Before the Health and the Environment Subcomm. of the House Energy and Commerce Comm.*, 100th Cong. (1992).

[201] In 1992, Congress passed the Prescription Drug User Fee Act of 1992, Pub L. No. 102-571, 106 Stat. 4491 (codified in scattered sections of 21 U.S.C.), permitting the FDA to charge pharmaceutical manufacturers a fee for review of new drug applications. In 1997, Congress enacted the FDA Modernization Act of 1997, Pub. L. No. 105-115, 111 Stat. 2296 (codified in scattered sections of 21 U.S.C. (2000)), which codified past practices of the FDA, granted the agency additional authority in certain areas, and reduced various restrictions the agency had placed on industry. However, sometimes Congress passed legislation to limit the FDA's jurisdiction when it thought the FDA had overstepped its authority. *See, e.g.*, Dietary Supplement Health and Education Act of 1994, Pub. L. No. 103-417, 108 Stat. 4325 (codified as amended in scattered sections of 21 U.S.C.).

[202] *See generally* David M. Gossett, *Chevron, Take Two: Deference to Revised Agency Interpretations of Statutes*, 64 U. CHI. L. REV. 681 (1997). Lack of uniformity in Supreme Court deference under *Chevron*, as a general matter, may stem from differing views on statutory construction held by the justices. A different method of construction could result in a different level of deference. Adam Babich, *Regulatory Reform and the Chevron Doctrine*, 26 ENVTL. L. REP. 10,597, 10,598–99 (1996). In addition, the ideological bent of appellate judges can influence how *Chevron* is applied in a given case. *See* Emerson H. Tiller & Frank B. Cross, *Colloquy: A Modest Proposal for Improving Am. Justice*, 99 COLUM. L. REV. 215, 221–24 (1999). *But see* Patricia M. Wald, *Colloquy: A Response to Tiller and Cross*, 99 COLUM. L. REV. 235 (1999).

The *Chevron* opinion itself suggests that agencies not only can revise their statutory constructions but must continually reconsider their interpretations to remain informed in subsequent rulemaking.[203] However, in a footnote in the majority's opinion of *INS v. Cardoza-Fonseca*,[204] the Supreme Court stated that an agency interpretation that conflicts with a prior interpretation should be accorded less deference than a consistently held construction.[205]

When revisiting this issue during the 1990 term, the Supreme Court made conflicting statements.[206] In the 1993 *Good Samaritan Hospital v. Shalala* case,[207] the Court further complicated the issue. After first stating that an agency can revise its policy,[208] the Court quoted the aforementioned language from *Cardoza-Fonseca*[209] and indicated that "the consistency of an agency's position is a factor in assessing the weight that position is due"[210] and that the weight accorded the agency's interpretation will depend on the facts of the individual case.[211] Similarly, in *Thomas Jefferson University Hospital v. Shalala*,[212] the Supreme Court re-asserted in dictum that an agency's interpretation of a statute or regulation that conflicts with a prior interpretation is entitled to "considerably less deference" than a consistently held agency view.[213]

During its 1995 term, the Supreme Court returned to a more deferential view of revised agency interpretations in *Smiley v. Citibank*.[214] The Court qualified its approach, however, stating that a revised agency con-

---

[203] "An initial agency interpretation is not instantly carved in stone. On the contrary, the agency, to engage in informed rulemaking, must consider varying interpretations and the wisdom of its policy on a continuing basis." *Chevron*, 467 U.S. at 863–64.

[204] 480 U.S. 421 (1987) (rejecting the Board of Immigration Appeals's interpretation of proof of "well-founded fear of persecution" of section 208 of the Refugee Act of 1980, 8 U.S.C. § 1101(a)(42), because it was contrary to the plain language of the statute).

[205] *Id.* at 446 n.30 (citing Watt v. Alaska, 451 U.S. 259, 273 (1981)).

[206] In *Rust v. Sullivan*, 500 U.S. 173 (1991), the Supreme Court bolstered the Chevron ruling, adding that an agency "must be given ample latitude to adapt [its] rules and policies to the demands of changing circumstances." *Id.* at 186–87 (quoting Motor Vehicle Mfrs. Ass'n of U.S. v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 42 (1983) (quotation omitted)). The Court implied, however, that the agency must justify its interpretation with a "reasoned analysis." *Id.* (citing *State Farm*, 463 U.S. at 42). Later that term, the Court indicated that agency constructions inconsistent with previously held views should be given less deference. Pauley v. BethEnergy Mines, Inc., 501 U.S. 680, 706 (1991) (upholding the Secretary of Labor's interim regulations of the Black Lung Benefits Act of 1972, 86 Stat. 150).

[207] 508 U.S. 402 (1993) (deferring to the Secretary of Health and Human Services's interpretation of clause (ii) of section 1861(v)(1)(A) of the Social Security Act, as amended, 42 U.S.C. § 1395x(v)(1)(A)(ii), because it was at least as plausible as competing ones).

[208] *Id.* at 417.

[209] *See supra* text accompanying note 205.

[210] *Good Samaritan Hospital*, 508 U.S. at 417.

[211] *Id.*

[212] 512 U.S. 504 (1994).

[213] *Id.* at 515.

[214] 517 U.S. 735 (1996).

struction may be overruled if it is a "[s]udden and unexplained change, or change that does not take account of legitimate reliance on prior interpretation . . . ."[215]

Despite the inconsistency in the language of its opinions, the Supreme Court has generally deferred to revised agency interpretations, and when it has not, the act of revising has not been dispositive of whether the Court deferred.[216] This Essay contends that the original *Chevron* approach was preferable to the *Cardoza-Fonseca* approach.

Intermediate appellate courts also tend to defer to revised administrative interpretations.[217] Although the large number of cases which have employed the *Chevron* analysis prohibits a complete assessment, David Gossett reviewed all 1995 appellate cases that dealt with revised agency interpretations of statutes.[218] Of the forty-three cases he examined, the courts deferred in twenty-three.[219] In several of these cases, the courts stated that they gave the agency less deference than they would have given to consistent agency positions.[220] In the twenty cases where the courts did not defer, the courts consistently indicated that they accorded the agency less deference because of the change in its interpretation.[221] However, in fourteen of these twenty cases, the courts ruled against the agency on separate grounds that justified not deferring and in two others overturned the agency because it had not provided adequate reasons for the revised interpretation.[222]

This Essay contends that courts should consistently grant deference to agencies' modifications of prior statutory interpretations. Consistent application of this principle—implicit in *Chevron* but inconsistently ap-

---

[215] *Id.* at 742 (citations omitted).

[216] Gossett, *supra* note 202, at 696–97. In *Cardoza-Fonseca*, for example, the Supreme Court did not uphold an agency's modification of an earlier statutory interpretation. However, the Court justified its holding on the basis that the statute at issue was not ambiguous and, therefore, not subject to agency interpretation. 480 U.S. 421, 449 (1987). The Court also rejected a revised administrative construction in *Lechmere, Inc. v. NLRB*, 502 U.S. 527 (1992), but on the basis of stare decisis. The Court held that an agency cannot reinterpret a statute if the Supreme Court has already decided the meaning of the relevant statutory provision. *Id.* at 536–37. Applying the same rationale, the Supreme Court in *Neal v. United States*, 516 U.S. 284, 295–96 (1996), overruled the Sentencing Commission guideline defining an LSD mixture as not including a carrier medium because the Court had previously defined it as including the medium.

[217] Gossett, *supra* note 202, at 695–96.

[218] *Id.*

[219] *Id.* at 697 n.71.

[220] *Id.* at 698.

[221] *Id.* at 698–99.

[222] *Id.* at 699. Some courts will not uphold an agency policy unless the agency justifies its proposed actions. The Supreme Court first applied this "hard look" doctrine in *Motor Vehicle Manufacturers Ass'n v. State Farm Mutual Automotive Ins. Co.*, 463 U.S. 29, 46–57 (1983). Since its decision in *Chevron*, the Court has not used the hard look approach within a *Chevron* analysis except in *Rust v. Sullivan*, 500 U.S. 173, 187 (1991), where the Court implied a justification requirement. *Cf.* Arent v. Shalala, 70 F.3d 610, 616 n.6 (D.C. Cir. 1995) (discussing the occasional overlap between the second step of *Chevron* and *State Farm* review).

plied since then—would ensure agencies the flexibility they need to meet new challenges. Of course, existing statutory delegations form the outer boundary of an agency's power.

C. *Changing Circumstances*

Changing conditions can impose hardships on a regulatory framework. From a practical standpoint, Congress cannot address every new circumstance through statutory amendments. Moreover, excessive statutory rigidity may constrain effective regulation.[223] Therefore, new developments, such as the expansion of technical knowledge and scientific innovations, may require administrative agencies to adapt old regulatory schemes to meet these challenges to remain consistent with congressional objectives.

The Supreme Court has recognized that Congress cannot anticipate all conditions to which a policy may apply.[224] In some instances the Court has construed statutes liberally,[225] acknowledging the authority and expertise of agencies to adopt new policies when confronted with new circumstances.[226]

Although the Supreme Court, in *Maislin Industries, U.S., Inc. v. Primary Steel, Inc.*,[227] ruled that agencies cannot alter an interpretation that the Court has held to be statutorily mandated,[228] agencies can establish new policies in response to changing circumstances.[229] Moreover, the Supreme Court indicated, in a later opinion, that the judiciary should grant "substantial deference" to an agency's change in policy when there are good reasons for the change.[230]

Agencies should be permitted to develop an "evolutional approach" to the statutes they administer. In *Smiley v. Citibank*,[231] the Supreme

---

[223] *See Law and Administration*, *supra* note 190, at 2089.

[224] United States v. Haggar Apparel Co., 526 U.S. 380, 392 (1999) (observing that "[a] statute may be ambiguous, for purposes of *Chevron* analysis, without being inartful or deficient" simply because "Congress need not, and likely cannot, anticipate all circumstances in which a general policy must be given specific effect").

[225] *See, e.g.*, United States v. Article of Drug Bacto-Unidisk, 394 U.S. 784, 798 (1969) (recognizing "[t]he well-accepted principle that remedial legislation such as the Food, Drug, and Cosmetic Act is to be given a liberal construction consistent with the act's overriding purpose to protect the public health . . . .").

[226] *See* NLRB v. J. Weingarten, Inc., 420 U.S. 251, 266 (1975) (stating that Congress entrusted the NLRB with the responsibility to adapt the National Labor Relations Act to "changing patterns of industrial life").

[227] 497 U.S. 116 (1990).

[228] *Id.* at 131.

[229] *Id.* at 134.

[230] Robertson v. Methow Valley Citizens Council, 490 U.S. 332, 355–56 (1989) (holding that the Council on Environmental Quality's new regulation was entitled to substantial deference because the United States Forest Service had a well-reasoned basis for its change in policy).

[231] 517 U.S. 735 (1996).

Court stated that merely contradicting a prior agency position will not prove fatal to an agency interpretation.[232] A change in administrative policy may be unreasonable if it is sudden and unexplained or does not take into account "legitimate reliance on prior interpretation."[233] However, a change in policy alone does not invalidate an agency's decision. Indeed, one reading of *Chevron* is that an ambiguity in statutory interpretation actually serves as a grant of authority to the regulating agency. As the Supreme Court held in *NLRB v. Curtin Matheson Scientific, Inc.*,[234] an agency should be permitted to use an "evolutional approach," continually developing its interpretation of the statute it administers.[235]

Constraining agencies to their prior decisions misconceives administrative decision-making[236] and inappropriately freezes administrative policy choices. As a consequence, agencies might choose not to state final policies and statutory interpretations explicitly in an effort to avoid restricting their ability to implement and govern the statutes they administer effectively. Regulated industries and consumers, therefore, would not receive sufficient guidance from agencies on how to comply with statutory requirements.

New conditions also may render a statute ambiguous. The failure of a statutory provision to remain consistent with popular social norms or the changing legal context can invalidate a statute's underlying factual assumptions, rendering a once clear provision ambiguous.[237] Failure to account for new circumstances when construing a statute could produce outcomes inconsistent with an act's general purposes.[238]

The rationales in *Chevron* that justify judicial deference to initial agency statutory constructions[239]—institutional competence and electoral accountability—also apply when agencies confront changed conditions or establish new policies. Agencies are also better suited than courts to reinterpret statutes consistent with changing circumstances because of their fact-finding abilities, public accountability, and practical knowledge of new developments.[240] In addition, the agency's cumulative experience,

---

[232] *Id.* at 742. *Chevron* offers another good example. In that case, the Supreme Court upheld the EPA's plantwide definition of "source" even though it was contrary to its earlier decision. *Chevron*, 467 U.S. at 863.

[233] *Smiley*, 517 U.S. at 742.

[234] 494 U.S. 775 (1990).

[235] *Id.* at 787 (quoting NLRB v. J. Weingarten, Inc., 420 U.S. 251, 265–66 (1975)).

[236] *Id.*

[237] *See* Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 HARV. L. REV. 405, 494–95 (1989).

[238] For a discussion of the impact of changing circumstances on the interpretation of the Food, Drug, and Cosmetic Act's Delany Clause, which prohibits the sale of food additives that induce cancer, see *id.* at 496–97.

[239] "For a court to do otherwise undermines the essence of *Chevron* deference and makes deference into either a 'doctrine of desperation' or a truly revolutionary doctrine, undermining the Marbury view of the role of the courts in interpreting laws." Gossett, *supra* note 202, at 707 (citation omitted).

[240] *See Law and Administration*, *supra* note 190, at 2088–89, 2102–03.

which results from constant trial and error, provides the agency with unique insight into the statute it administers.[241]

Two conclusions follow. First, an agency should be allowed to modify its policies when confronted with changing circumstances, as long as the agency's interpretation accords with the *Chevron* requirements. In other words, if Congress was silent or ambiguous with regard to a question, and the agency's new interpretation is reasonable, the agency should be permitted to adopt a new approach to new circumstances. Second, courts should grant deference under *Chevron* to a modified agency interpretation equal to that shown to the agency's original construction, as long as the new interpretation is reasonable.[242]

IV. CONCLUSION

By design, the legislature's ability to enact laws to meet evolving and new conditions is hampered by the requirements of political conciliation.[243] The rapid growth of information, the increasing complexity of policy issues, and improved telecommunications have added new hurdles to the traditional functions of government. Public preferences also

---

[241] NLRB v. J. Weingarten, Inc., 420 U.S. 251, 266 (1974) (quoting NLRB v. Seven-Up Co., 344 U.S. 344, 349 (1953)).

[242] *But see Law and Administration*, *supra* note 190, at 2104 ("[N]ew departures should be accorded somewhat less deference than longstanding interpretations, for reasons analogous to those that justify stare decisis in the judicial context.").

[243] The Framers designed the constitutional framework of the federal government to impede legislative action in the absence of widespread consensus. Sufficient uniformity of purpose to enact legislation rarely surfaces on the congressional floor due to Madisonian factionalism, individual legislators' over-responsiveness to constituent needs, and legislators' susceptibility to interest group influence, arising, in part, from a desire to obtain reelection. *See* Harold H. Bruff, *Legislative Formality, Administrative Rationality*, 63 TEX. L. REV. 207, 217–19 (1984) (reviewing the incentives facing interest groups and legislators in the context of lobbying efforts as well as various measures taken to reduce the impact of factionalism in legislative decision-making); GARY C. JACOBSON, THE POLITICS OF CONGRESSIONAL ELECTIONS 188–89 (4th ed. 1997) (discussing changing tactics used by interest groups to lobby congressional members); DAVID R. MAYHEW, CONGRESS: THE ELECTORAL CONNECTION 13–17 (1974) (describing reelection as one goal of congressional members); John T. Tierney, *Organized Interests and the Nation's Capitol*, *in* THE POLITICS OF INTERESTS: INTEREST GROUPS TRANSFORMED 201–20 (Mark P. Petracca ed., 1992) (analyzing the impact of interest groups on congressional decision-making). Agencies, however, may also be subject to interest group influence. *See* GEORGE J. STIGLER, THE CITIZEN AND THE STATE: ESSAYS ON REGULATION 114–41 (1975) (contending that industry obtains regulation and regulation is implemented for industry's benefit).

Arguably, granting deference to executive agency statutory constructions might allow the executive branch to circumvent the legislative process in violation of explicit bicameralism/presentment requirements and the spirit of separation of powers by effectuating executive branch policies through quasi-legislative actions taken by executive agencies. *See, e.g.*, Morrison v. Olson, 487 U.S. 654, 696–707 (Scalia, J., dissenting) (arguing that the Constitution requires an absolute separation of powers, and any acquisition or deprivation of one branch's power by another branch is inherently unconstitutional). However, should the executive use an agency to encroach on legislative terrain, Congress could have the final word by retracting or modifying the agency's authority or nullifying the agency's action.

change over time; they are not externally imposed, but rather are internally generated in response to social pressures, information, legal rules, and consumption patterns.[244] Failure to establish procedures to recognize and respond to these changes undercuts the ability of legislation to accomplish its purposes and goals.

Because of their specialization and expertise, agencies provide the legislative and executive branches with a mechanism for anticipating future economic and social needs and for responding rapidly and effectively to changing conditions without sacrificing political accountability.[245] Agencies also interact continuously with regulated entities and the public as both suppliers of information and collectors of feedback.[246]

Congress and the President can implement policy initiatives through agencies and use them to adapt policies to changing conditions. The history of the FDA's implementation of statutory pre-market drug approval requirements offers a good example of how an administrative agency, as an arm of the legislative and executive branches, addresses new conditions not originally contemplated by the statute's drafters.

To ensure regulatory flexibility, judicial oversight of agency actions must take into account changing circumstances when reviewing agency decisions. At the same time, there must be adequate judicial, legislative, and executive oversight of agency actions, as well as effective procedural requirements, to ensure that agency actions do not stray far afield of congressional delegations of authority and that agencies are attentive to public needs and other stakeholders' views.

The Supreme Court's *Chevron* decision and much of the subsequent judicial evolution of the *Chevron* analysis are consistent with the role of agencies as agents of change. However, the implementation of *Chevron* has been inconsistent. Cases like *INS v. Cardoza-Fonseca*[247] threaten to undercut the ability of agencies to perform the tasks assigned to them by Congress and the President. Courts should grant as broad deference to administrative agencies' modified statutory interpretations as they do to original agency interpretations, so long as the new construction is reasonable.

---

[244] RIGHTS REVOLUTION, *supra* note 27, at 44.

[245] *See* Jerry L. Mashaw, *Prodelegation: Why Administrators Should Make Political Decisions*, 1 J.L. ECON. & ORG. 81, 95–99 (1985) (discussing the use of agencies to respond to voter preferences voiced in presidential elections). See *supra* Part III for the Supreme Court's discussion of agencies' electoral accountability.

[246] *See* Mark Seidenfeld, *A Civic Republican Justification for the Bureaucratic State*, *in* FOUNDATIONS OF ADMINISTRATIVE LAW 25, 25–27 (Peter H. Schuck ed., 1994) (arguing that administrative agencies fulfill the civic republican ideal of deliberative decision-making better than Congress or the courts because agency officials have greater expertise and fewer political pressures than elected officials, and they are not overinsulated from the public as are judges).

[247] 480 U.S. 421 (1987).

Courts often lack the expertise to adjudicate the technical issues involved in administrative law without the benefit of expert testimony and amicus briefs. In contrast, agencies' specialization and expertise give them unique insight into their organic statutes that courts lack.

Courts should also grant agencies deference because agencies are fundamentally more accountable actors. Unlike courts, executive agencies are tied to the democratic process through the election of a President. Additionally, the *Chevron* approach provides two important checks on agency action by requiring fidelity when Congress has spoken and reasonableness when it has not.

Changing circumstances often necessitate an evolving approach. Innovation should not be punished, as doing so creates a disincentive for agencies to pursue the solutions their experience suggests are necessary in the present, merely because previous agency decision-makers had concluded that another approach fit best with a bygone set of circumstances.

ARTICLE

GAME THEORY, LEGISLATION, AND THE MULTIPLE MEANINGS OF EQUALITY

DAVID CRUMP*

Game theory studies encounters (or games) between two or more players in which each has a clearly defined choice of strategies and in which there are well-established payoffs from the potential outcomes. In this Article, Professor Crump examines several different types of games and addresses whether any of them, or game theory in general, has real-world applications in the legislative realm. He concludes that while game theory does not offer many firm answers about legislation, the strategies examined by the discipline help to illuminate many of the problems raised by legislation that affects equality.

Game theory is the study of strategic interactions that can be analyzed by logic.[1] It provides rigorous answers, but that result is possible only because the situations are artificially structured. In this discipline, a game is not the same thing as an amusement, but rather has a technical definition. It is an encounter between two or more players, in which each player has a clearly defined choice of two or more ways of acting (called strategies), with well-defined payoffs from the potential outcomes.[2] By this definition, tic-tac-toe, chess, and paper-scissors-stone are games in the technical sense, but role-playing games such as cowboys and Indians, or those primarily dependent upon physical skills such as hopscotch, are not.[3] Football, as a sixty-minute endeavor, is not a game to the extent that it depends on unquantifiable tests of speed, size, and strength, but many strategic situations in football can be analyzed meaningfully through game theory.[4]

At the same time, games are not limited to frivolous fun. In fact, many ostensibly unrelated disciplines have been illuminated by game theory. For example, marketing strategies in oligopolies can be modeled

---

* Newell H. Blakely Professor of Law, University of Houston. A.B., Harvard College, 1966; J.D., University of Texas, 1969.

[1] *See* JAMES W. FRIEDMAN, GAME THEORY WITH APPLICATIONS TO ECONOMICS 3–4 (2d ed. 1990); MORTON D. DAVIS, GAME THEORY: A NONTECHNICAL INTRODUCTION 6 (rev. ed. 1983); ANDREW M. COLMAN, GAME THEORY AND ITS APPLICATIONS IN THE SOCIAL AND BIOLOGICAL SCIENCES 3–4 (2d ed. 1995).

[2] *See* FRIEDMAN, *supra* note 1, at 3–4; DAVIS, *supra* note 1, at 6; COLMAN, *supra* note 1, at 3–4.

[3] *Cf.* COLMAN, *supra* note 1, at 3 (using different examples to make a similar point).

[4] For example, one could construct a payoff matrix for determining whether to punt on fourth down. As another example, coaches have available a chart that tells whether to kick (one point) or run (two points) after a touchdown, depending on the existing point spread. *See* Ethan J. Skolnick, *Two-Point Probability Chart Works If Play Does*, PALM BEACH POST, Oct. 31, 1997, at 8C.

by games,[5] as can some aspects of military strategy.[6] Similarly, ethical dilemmas, problems of political science, and even the phenomenon of natural selection in biology have been usefully addressed by game theorists.[7]

Can game theory, then, teach us anything about legislation that affects the legal and ethical value that we call "equality"? That is the question this Article seeks to address. The Article begins by examining whether equality has sufficiently definable meanings so that it can be analyzed by logic. This is not a trivial question, and the answer will be inconclusive. The first Part maintains that at the very least, game theory can illuminate ambiguities in our understanding of equality, and it is possible that it can do much more.

The remainder of the Article discusses the application of game theory to legislation affecting equality. Part II begins by explicating a well-known problem called the "truel" (or triple-duel) paradox. It demonstrates the implications the logic of strategy might have for legislative interventions to enhance equality in ambiguous circumstances. More generally, Part II illuminates the usefulness of game theory as a tool for considering equality, as well as its limitations.

Parts III and IV of the Article go on to develop the basic tools of game theory, including payoff matrices for zero-sum games, analyses of two-person mixed-motive games, and solutions to problems of coalition formation, with emphasis on analysis through what are known as Shapley values. In addition, Part V considers what are called maximax, minimax, and maximin strategies, and it applies the analysis to John Rawls's choice of the maximin strategy in his landmark *A Theory of Justice*.[8] In each instance, the Article uses game theory to consider legislative problems involving equality.

The Article concludes by emphasizing that game theory, as a normative discipline distinct from ethics, cannot give us many firm conclusions about legislation affecting equality. In fact, the reader may emerge with a sense that game theory is too stilted and artificial to have many applications to real-world problems, although this conclusion should be resisted. It seems that a small but important set of characteristics of the problem we call equality can be usefully illuminated by game theory, provided it is consulted with appropriate caution. For example, game theory helps to demonstrate that rather than expressing a unitary concept, the word equality has multiple meanings. There are many equalities, and they compete with each other. Strategic considerations must influence a

---

[5] MICHAEL E. PORTER, COMPETITIVE STRATEGY: TECHNIQUES FOR ANALYZING INDUSTRIES AND COMPETITORS 88–89 (1980) (analyzing oligopolistic industry by reference to the Prisoners' Dilemma—a game discussed in Part IV.B.3, *infra*).

[6] *See infra* Part IV.A (using a payoff matrix to model a famous naval battle).

[7] *See* COLMAN, *supra* note 1, at chs. 10–12 (applying game theory to biological, political, and moral questions).

[8] JOHN RAWLS, A THEORY OF JUSTICE (2d ed. 1971).

statutory approach to all of these moral messages, because conflicting strategies inevitably will affect the behavior of individual citizens subject to a legislative norm that mandates equality.

Even apart from its implications for legislation affecting equality, the logic explained in this Article may provide insights to law professors and students about such matters as shareholder agreements, voting rights, negotiation, discovery, the evaluation of settlements, and other problems of legal theory and practice. It can provide insights that are difficult to obtain from the appellate opinion analysis that dominates most learning in law school. Thus, in addition to what it can show us about equality, game theory may help to fill gaps in the traditional method of legal education.

I. Does "Equality" Have Any Meaning That Can Be Explored by Game Theory?: The Semiotic and Metaphysical Questions Underlying the Argument That Equality Is an "Empty" Idea

Game theory presupposes a problem defined so that it is amenable to logical or mathematical analysis.[9] The terms need not be free of ambiguity so long as they still admit of logical treatment.[10] Game theory, in fact, can help to discover concealed ambiguities and non-obvious choices. Once we begin to use it, however, we must confront whether there is meaning in the term equality, to which we shall apply the theory. As this section will show, game theory can assist with this initial inquiry.

A. A Threshold Question: Does Equality Have a Useful Meaning?

The threshold question, then, is whether the word equality matters, or in other words, whether it has any normative or descriptive meaning at all. This may seem a surprising question, but several commentators have argued, with some persuasive force, that it does not. Professor Peter Westen raised the issue in his article, *The Empty Idea of Equality*.[11] As

---

[9] *See supra* note 1 and accompanying text.

[10] *See* Colman, *supra* note 1, at 3–4.

[11] Peter Westen, *The Empty Idea of Equality*, 95 Harv. L. Rev. 537 (1982). Professor Westen's article sparked considerable debate. *See also, e.g.*, Christopher J. Peters, *Equality Revisited*, 110 Harv. L. Rev. 1210, 1211 (1997) (replying to Professor Westen, arguing that the equality principle either reduces to the tautological statement that people who should be treated alike should be treated alike, or, if nontautological, as prescriptive equality, leads to indefensible results or is incoherent); Kent Greenawalt, *"Prescriptive Equality": Two Steps Forward*, 110 Harv. L. Rev. 1265, 1266 (1997) (replying to Professor Peters, arguing that prescriptive equality does have meaningful normative force, in that when one individual is treated better than nonegalitarian justice requires, another person, sufficiently similarly situated, may have an egalitarian claim to the same treatment); Cass R. Sunstein, *On Analogical Reasoning*, 106 Harv. L. Rev. 741, 774 & n.115 (1993) (suggesting, in opposition to Professor Westen, that even if deductive reasoning about equality

the title indicates, the argument is that the very principle of equality is "empty" of content.[12] Westen believes that the very emptiness of the concept, which he expresses in terms of the principle that likes should be treated alike, is precisely the reason that equality has such durable appeal.[13] It is manipulable enough to serve conflicting aims, and every legislator can claim to favor it. Westen explains, however, that if equality is to have any meaning, it must derive from some external determination—some set of principles independent of the word equality—specifying which persons and treatments are alike.[14] He maintains, however, that once these determinations have been made, the concept of equality becomes superfluous.[15] Worse yet, he argues, the word equality confuses the issue. He maintains that we would be better off searching solely for those external values that tell us which people and treatments are alike, and abandoning the rhetoric of equality.[16]

Westen demonstrates convincingly that equality is confusing.[17] There is no single, unitary equality: instead, there are many equalities, and indeed there is an indefinite range of them.[18] Westen also demonstrates that many issues labeled as matters of equality are determined by other values, with formal notions of equality reduced to triviality.[19]

But it does not necessarily follow that we should abandon the rhetoric of equality as an empty concept. First, it is unlikely that legislators, courts, and legal philosophers will do so.[20] Second, although there may

---

can be reduced to meaninglessness, reasoning by analogy may make equality meaningful); Mark Kelman, *Concepts of Discrimination in "General Ability" Job Testing*, 104 HARV. L. REV. 1158, 1230–31 (1991) (concluding that, in the case of job testing, horizontal equality is sometimes of little use); John Stick, *Can Nihilism Be Pragmatic?*, 100 HARV. L. REV. 332, 346–47 & n.56 (1986) (asserting that deductively a rule must specify its range of application to be complete and citing Westen for this model, but concluding that this criticism does not apply to analogic reasoning or reasoning by the gravitational weight of precedent).

[12] *See generally* Westen, *supra* note 11.

[13] *Id.* at 542–47.

[14] *Id.* at 552.

[15] *Id.* at 547–55.

[16] *Id.* at 577–95.

[17] *See id.* at 579–80.

[18] *See id.* at 583.

[19] *See id.* at 560.

[20] For example, in *Brown v. Board of Education*, 347 U.S. 483 (1954), the Court concluded, "Separate educational facilities are inherently unequal." *Id.* at 495. The Court evidently believed it was basing this deduction on formal logic or on the meaning of equality as a concept sufficient in and of itself, because it described the perceived inequality as "inherent." The implication of Westen's point, however, is that the Court logically could not have done so, because it necessarily consulted values extrinsic to the idea of equality, such as the principle against racial discrimination that is central to the purpose of the Fourteenth Amendment. Along these lines, commentators have suggested that it might have been better if the Court had simply applied the historic Fourteenth Amendment, rather than reasoning about equality in the abstract. *Cf.* DAVID CRUMP ET AL., CASES AND MATERIALS ON CONSTITUTIONAL LAW 584 (3d ed. 1998). The point here, however, is that as long as lawgivers persist in the logic of an inherent idea of equality, it may be useful to explore the places this logic will take them.

be instances in which a formal logic of equality will not work, perhaps there are some in which it can.[21] Ambiguity in some senses does not imply the absence of all meaning. Furthermore, Westen's theory depends on deductive logic, or definitional principles of equality, and his thesis is most persuasive in that context; it does not, however, destroy reasoning by analogy as effectively.[22] Analogy, in the form of case comparisons, is a widely used method of reasoning in the law.[23] Equality may have clearer meaning, in other words, if we allow for the possibility of a lobbyist who says, "My client should be treated in such-and-such fashion because her situation is analogous to that of others who have been treated in this fashion."[24] Here, game theory may help us to identify the relevant aspects of a problem upon which to build analogies.[25]

To take a pedestrian-sounding (but actually difficult) example, imagine that a testator's will leaves a parcel of land called Greenacre to "my three children, Alice, Bill, and Carl, to be partitioned by a vote of the majority among them." Thus, any two of the children can impose a division of Greenacre on the third. But the state's Probate Code provides that a class gift (such as one to "my children") "shall be construed to provide equal shares." Thus, if this statute governs, our conclusion changes: any two of the children can act strategically and impose a division of Greenacre on the third, but *only* so long as it can be called "equal." The construction of the testator's three-part gift requires an inquiry into what the legislature meant by "equality," as well as an examination of the different ways that the sudden-rival siblings can coalesce against each other, and thus this seemingly simple task becomes a game treatable by surprisingly complex game theory.[26] Let us assume that there is no legislative history to guide interpretation, or that the legislators

---

[21] For example, Professor Greenawalt finds prescriptive value in the idea that a claimant may be entitled to better treatment if an indistinguishable claimant has received that treatment, even if nonegalitarian justice would not require the treatment for either. Greenawalt, *supra* note 11, at 1265–66. Professor Peters criticizes this reasoning on the ground that it amounts to giving one person a "wrong" benefit because another has received a "wrong" benefit. Peters, *supra* note 11, at 1212. But this criticism presupposes that each theory of justice is either right or wrong. Perhaps egalitarian justice is a useful check upon, or adjustment to, nonegalitarian theories of distribution.

[22] *See* Sunstein, *supra* note 11, at 774 & n.115 (arguing in response to Professor Westen that the emptiness postulate does not hold if the lawgiver uses analogy rather than deduction).

[23] *Id.* at 741.

[24] This analysis combines the responses of Professors Greenawalt and Sunstein. *See supra* notes 21, 22.

[25] *Cf. infra* Part II (using a familiar game called the "truel" paradox to demonstrate that two apparently equivalent government strategies for securing equality, founded on different assumptions about the players, lead in non-obvious ways to different results: success in one case, failure in the other).

[26] A similar encounter, called the "stockholders' problem," has been the subject of extensive analysis by game theorists. If we assume that a disgruntled sibling can contest the division in court (a realistic assumption), the strategies of the players become still more complex. *See infra* Part III.

specifically agreed that there was to be no legislative history and that the unadorned statutory text should govern.[27]

If Bill and Carl agree to vote together so that they each get half of Greenacre, and Alice receives nothing, observers will likely reach a consensus that, unless something else is going on, the division violates the legislative intent expressed in the word "equal." This conclusion depends upon a formal concept of equality (at least if the result is based only upon the text of the statute). Specifically, the formal concept here is a mathematical construct (which is to say, a game theory construct) of equality as congruence or exchangeability, and it implies a distribution of one-third to each. We have, of course, assumed that the three devisees are "alike" enough to warrant this congruent treatment, and Professor Westen's point is that this "alikeness" issue always is ambiguous. Here, however, the ambiguity in the alikeness of the three seems trivial.

But equality, Professor Westen might say, still remains empty. Suppose Bill and Carl's lopsided decision can be defended; perhaps, for example, Alice is dead, and the law of the jurisdiction is that her gift lapses. In this event, however, the stage is set for arguments by analogy in favor of enacting an anti-lapse statute.[28] Although extrinsic values such as family integrity, inter-generational distributive justice, or support of minor children might aid the debate, so might formal inquiry into what is meant by equality, which is the concept the legislature had targeted. Alternatively, imagine that a probate judge awards all of Greenacre to Alice (and nothing to Bill and Carl) on the ground that Alice is a woman. (Assume that the jurisdiction is thoroughly sexist and regards Alice as the only competent devisee.) Again, extrinsic values such as a principle for or against gender discrimination may enter into the argument for corrective legislation, but so might a focus upon the original statutory language, together with inquiry into the formal meaning of equality.

As the remainder of this Article shows, game theory can assist this inquiry. In fact, the Probate Code example that we have just considered is itself a rudimentary application of game theory, although our analysis here has been superficial. We shall explore the nuances of the Probate Code example in a later section, where it will be re-named the "stock-

---

[27] For example, the Texas venue statute was vigorously negotiated over a substantial time by plaintiff and defendant representatives under the guidance of a state supreme court justice, and it passed with the agreement that there would be no legislative history. Telephone Interview with William V. Dorsaneo III, Professor of Law, Southern Methodist University (Mar. 29, 2001); *accord* Telephone Interview with Kent Caperton, former Texas State Senator (Apr. 2, 2001) (Senator Caperton was a sponsor of the legislation). Professor Dorsaneo reports that the negotiations were contentious, both parties recognized that they contained strategic statements, and they considered it best if the text rather than partisan statements controlled. Professor Dorsaneo monitored the legislative process because he had the task of drafting court rules conforming to the statute.

[28] *See, e.g.*, DEL. CODE ANN. tit. 12, § 2313 (1995).

holders' problem."[29] This well-studied game models the fundamental dilemma of constitutional democracy: how a regime governed by competitive, strategic voting by majority rule, in which players seek their own advantage, can be made compatible with a judicially imposed requirement of equal protection.

B. The Two Layers of the Question: Semiotics and Metaphysics

Actually, the threshold problem—whether equality has any meaning—is twofold: there is a semiotic issue which overlaps with a metaphysical question. Semiotics is the study of symbols and of the accuracy with which they convey concepts.[30] Metaphysics is an inquiry into the true nature of things, or in other words, into how closely concepts correspond to reality.[31] On hearing the word "bird," for example, one person may visualize a cardinal, another a robin. To the extent they understand the communication differently, they are victims of a semiotic difficulty. A similar slippage in transmission occurs if the word is equality. Then, the two listeners also must deal with the reality that birds include not only these more prototypical examples, cardinals and robins, but also specimens that stretch the category, such as penguins, emus, and perhaps even pterodactyls. Here, the two thinkers confront a metaphysical problem: how closely does the concept of a bird correspond to any external reality? As for equality, it too is malleable: it can be framed to include a given set of features or to include others that are antithetical. Game theory cannot solve these problems, and in fact, in some of its uses, it wishes them away by insisting upon rules sufficiently well-defined to allow solution by mathematical logic. But as the Probate Code example demonstrates, the theory can help to explore the difficulties.[32]

Perhaps these semiotic and metaphysical difficulties can be overcome by a demonstration that they do not make words and concepts such as bird or equality completely empty of meaning, even though they do make them ambiguous. Again, an example may be useful, although this one does not qualify as a game because the precise point is that ambiguity controls the answers. In one of the wonderful police stories by Joseph Wambaugh, a rookie and a veteran officer encounter a victim who explains that the perpetrator has just left the scene and is wearing a red

---

[29] *See infra* Part III.

[30] *See* RANDOM HOUSE WEBSTER'S COLLEGE DICTIONARY 1219 (1996) (defining semiotics as "the study of signs and symbols as elements of communicative behavior").

[31] *See id.* at 852 (defining metaphysics as "the branch of philosophy that treats of first principles, includes ontology and cosmology, and is intimately connected with epistemology"). Ontology, in turn, is the study of "the nature of existence or being as such." *Id.* at 946.

[32] *See supra* Part I.A.

shirt.[33] After driving several blocks, the veteran focuses upon a suspect wearing a brown shirt.[34] When the suspect proves to be the perpetrator, the rookie asks how the veteran knew that the complainant meant to describe this suspect's brown-rather-than-red shirt.[35] The veteran replies that "[i]t was a color that could be *called* red."[36]

Thus, Wambaugh wraps together the problem of semiotics, the enigma of metaphysics, and the fallacy of assuming that either or both of these inquiries implies complete indeterminacy. It is true that some people, in using the word red, may mean crimson, magenta, maroon, brown, or orange. They thus create a problem of semiotics. It also is true that metaphysically, there is no such thing as a precise concept of red because there is a range of wavelengths that stretches across various colors, including some that are more or less red. Perhaps, then, one can argue that if equality is empty of meaning, red is just as empty. The veteran officer overcame these semiotic and metaphysical difficulties, however, and furnished proof that red, as symbol and concept, is not completely indeterminate.

Equality, as an abstraction, may be vulnerable to greater semiotic and metaphysical ambiguity than the color red. But the proposition that it therefore is empty is more difficult to argue, as the Probate Code example helps to show. Examination of the concept through game theory will demonstrate that indeed, equality is ambiguous,[37] but to the extent that equality is not completely empty of meaning, game theory may also help us to investigate it. Furthermore, game theory may help us to identify what external values a legislature consults to imbue it with meaning.[38] The remainder of this Article is devoted to these aims and to preserving the caveat that the contribution made by game theory requires skeptical evaluation.

II. THE TRUEL PARADOX: USING GAME THEORY TO CONSIDER LEGISLATIVE EFFORTS TO ENHANCE EQUALITY

This Part explores a familiar paradox identified by game theory. The game in question was not designed to analyze equality, but it is useful for that purpose. It shows how hidden assumptions about equality and ine-

---

[33] JOSEPH WAMBAUGH, THE NEW CENTURIONS 75 (1970).

[34] *Id.* at 76.

[35] *Id.* at 79.

[36] *Id.* at 80 (emphasis added).

[37] *Cf. infra* Part III (dealing with different game theory solutions to a political problem of distributive justice called the "shareholders' problem" and demonstrating that there are many defensible conceptions of equality, including non-obvious ones).

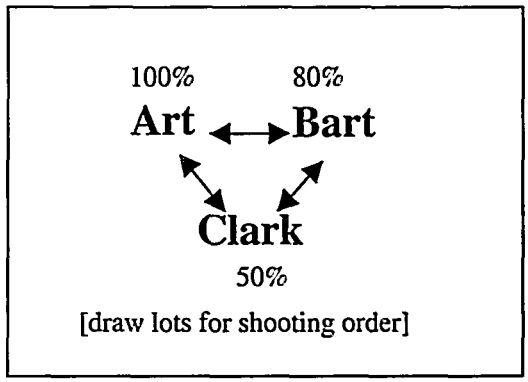
[38] *Cf. infra* Part IV (using game theory to identify the assumptions underlying different visions of equality).

quality defeat policy expectations, and it points to non-obvious ways in which a legislature can succeed or fail in its interventions.

A. Sketching the Truel Problem

The so-called "truel" problem results in a delightfully counter-intuitive paradox. It was first analyzed by game theorist Martin Shubik in 1954.[39] Imagine that Art, Bart, and Clark have agreed to conduct a triangular pistol duel or "truel." They will draw lots to determine the shooting order, and they will continue firing until only one is left alive. This odd encounter, of a kind only game theorists could love, is diagramed in Figure 1.

FIGURE 1
THE "TRUEL" PARADOX



But the three contestants are not equal in their marksmanship. Art is the best shot, and from the distance contemplated here, he hits his intended target one hundred percent of the time. Every hit is fatal. Bart is the next best: he hits his target eighty percent of the time. Then there is Clark, who is the worst shot of the three: he hits his target only fifty percent of the time.

---

[39] Martin Shubik, *Does the Fittest Necessarily Survive?*, *in* READINGS IN GAME THEORY AND POLITICAL BEHAVIOR 43–46 (Martin Shubik ed., 1954); *see also* COLMAN, *supra* note 1, at 273–74 (discussing and solving the problem and attributing it to Shubik).

Imagine further that the state legislature is concerned about the inequality among Art, Bart, and Clark. This particular legislature approves of duels (and even truels) but insists that they be conducted so that each contestant has an equivalent chance of survival. It demands a level playing field, to use a common metaphor. The leadership therefore decides that the state must intervene to even the scales for Clark. There are several forms this intervention could take. For example, the legislature might require a scattering device on Art's and Bart's firearms so that they, like Clark, become only fifty percent marksmen. Or perhaps the legislature could provide Clark with an unseen, partial bulletproof vest to reduce the likelihood of fatality to Clark from Art's and Bart's weapons to fifty percent. Superficially, these two approaches look equivalent, but they have very different implications in terms of equality.[40]

Fundamentally, there is a problem with the government's attempted leveling of the playing field in this case. The basic theory of legislative intervention in the truel problem is fallacious. Specifically, the lawmakers have failed to ask the question that should be asked first: "Who has the best chance of survival from this truel?" And the answer, paradoxically, is that Clark, the worst shot of the three, has the best odds of emerging alive![41]

The reason for this strange conclusion is based on the insight that the three players will act strategically. Art's and Bart's best strategies are to ignore Clark and shoot at each other. Clark's strategy, if he shoots first, is to fire deliberately to miss and thus to force Art and Bart to fight a duel between themselves.[42] The odds are that one of them will be killed, and then Clark will have the next shot. With this insight realized, we can use probability theory to compute the likelihood of survival for each of the three. If their chances were equal, Art, Bart, and Clark would each have a probability of survival of 1/3, or 0.333. Under the conditions of this truel, however, the probability of Art's survival is less than 1/3: it is 3/10, or .3000.[43] For Bart, the probability is a miserable 8/45 or 0.1778, much less

---

[40] *See infra* Part II.B.

[41] *See* Shubik, *supra* note 39, at 43–46; COLMAN, *supra* note 1, at 273–74.

[42] *See* Shubik, *supra* note 39, at 43–46; COLMAN, *supra* note 1, at 273–74.

[43] The probability of both of two independent events occurring is the product of their individual probabilities: $P_1 \times P_2$. The probability of either one of two mutually exclusive events occurring is the sum of their individual probabilities: $P_1 + P_2$.

Remember that Clark will not shoot to kill until after either Art or Bart is killed. We therefore begin computing Art's chances by figuring the probability that he will kill Bart, with Clark not playing. There is a fifty percent probability that Art will get the first shot in this duel, in which event he will survive. There also is a fifty percent probability that Bart will get the first shot, in which event Bart has a 4/5 likelihood of killing Art, so the probability of Art's surviving is $(1/5)(1/2) = 1/10$. Therefore, the probability of Art's surviving the initial contest with Bart is $1/2 + 1/10 = 6/10 = 3/5$.

Once he survives the duel with Bart, Art must face Clark, who gets the first shot. His probability of survival is 1/2. Therefore, Art's probability of surviving both the initial contest with Bart and the shot by Clark is $(1/2)(3/5) = 3/10 = 0.300$. *See* COLMAN, *supra* note 1, at 273 (explaining the mathematics).

than 1/3.[44] Clark's probability, on the other hand, is 47/90 or 0.5222.[45] Again, it is the worst shot of the three, Clark, who has the best odds. In fact, he has better than even odds—a greater than fifty percent chance—of surviving.

It follows that the legislature's intervention is counterproductive. It is counterproductive, that is, *if* the intervention succeeds in actually helping Clark. If, for example, the government provides Clark with a partial bulletproof vest, and it does so in secret so that Art and Bart do not know to change their strategies, the government will aggravate the existing imbalance. It will tilt further the already uneven playing field by giving advantages to the player who already has more than his share. On the other hand, if the government "helps" Clark in a different way, such as by placing devices on Art's and Bart's guns to reduce the accuracy of each to exactly fifty percent, Art and Bart presumably will know to adjust their strategies, and the change will not help Clark at all. Instead, it will reduce Clark's odds of survival. Accidentally and in spite of its aims, the legislature might, in this latter case, blunder into providing more equal probabilities of survival to all three.

B. Does the Truel Paradox Provide Any Insight into Problems of Equality?

What are the implications, if any, of this problem for the issue of equality? It is unlikely that any real-world situation conforms to the truel scenario precisely.[46] Rarely, one might imagine, do duelists know the exact percentages of their rivals' accuracy as marksmen. Nor do they often

---

[44] If Art gets the first shot in the contest with Bart, Bart will not survive; the probability of Bart even getting a shot, therefore, is 1/2. If he does, he has a 4/5 chance of surviving the opening encounter, and thus the probability of his survival is (1/2)(4/5) = 4/10 = 2/5.

But then Bart faces Clark, and since neither is a perfect marksman, there is an infinite series of potential exchanges. In the first, the probability of Bart's surviving Clark's shot is 1/2, and he then has a 4/5 chance of killing Clark, for a probability of (1/2)(4/5) = 4/10. If Clark misses, Bart misses, and Clark gets a second shot, the probability of his killing Clark is (1/2)(1/5)(1/2)(4/5) = 4/100. If we take the computation into the next round, the probability of Bart's surviving and killing Clark is (1/2)(1/5)(1/2)(1/5)(1/2)(4/5) = 4/1000. Adding these mutually exclusive outcomes, 4/10 + 4/100 + 4/1000, we obtain 4/9. This is the probability of Bart's surviving the duel with Clark, if he has survived the first duel with Art.

Since we earlier found that the probability of Bart's surviving the duel with Art was 2/5, Bart's overall probability of survival is (2/5)(4/9) = 8/45 = 0.1778. *Id.* at 273–74.

[45] This figure is easier to obtain. We subtract the combined probabilities of Art's and Bart's survival from 1. Thus, 1–[(3/10) + (8/45)] = 90/90–(27 + 16)/90 = 47/90. *Id.* at 274.

[46] The key word, however, is "precisely." The truel game probably models many kinds of real-world scenarios in an approximate way. Both Shubik and Colman, for example, have used it to illuminate natural selection and to show why the catch phrase "survival of the fittest" may be misleading. *See* Shubik, *supra* note 39, at 43–46; COLMAN, *supra* note 1, at 273–74.

fight three-way contests. And if they did, it seems unlikely that a legislature would intervene to level the playing field, as opposed to prohibiting the activity altogether.

But if one regards the truel problem as a metaphor for competition among multiple actors or for the division of scarce resources, perhaps one can learn something useful about the nature of equality from this problem in game theory, or at least use it to generate questions or raise cautionary considerations. First and most obviously, the truel paradox shows that policymakers need to understand whom they really are blessing with advantages, or saddling with disadvantages, before they attempt to intervene. This analysis requires a detailed survey of the current situation to determine who has real advantages or disadvantages.

Recent efforts at welfare reform (specifically, changes in the availability of Aid to Families with Dependent Children)[47] may reflect efforts to better address this first question. Admittedly, these efforts are controversial,[48] and to address the question is not to answer it. But the alleged reforms do reflect an initial consideration of whether old-fashioned federal mandates of long-term benefits result in enhancing equality, or whether they hurt precisely those whom they are intended to help. Is it possible that mothers will be capable of assuming a greater approximation of equality if the duration of benefits is limited and if some of the freed resources are expended to improve their work prospects? Is it conceivable that poor children who have working mothers will obtain a more equal status if some resources are expended on child care? Can we hope that absentee fathers, who are the real "Clarks" in this case, will become less plentiful as a result of women's strategies in response to these alleged reforms (or that government will use some remaining resources to ensure that they will pay their fair share)? Game theory does not tell us the answers, but the truel problem reminds us that it is important to ask the questions.

Another important idea the truel problem presents is that people will respond in strategic ways to legislation, including laws intended to enhance equality. Human beings sometimes act uncooperatively in an effort to enhance their own self-interest rather than tailoring their conduct to maximize the government's goals. Just as Clark, in the truel problem, confounds the expectations of policymakers by uncooperatively firing deliberately to miss, thereby enhancing his own position, disadvantaging his rivals, and aggravating the inequality that the legislature seeks to redress, so real people can be expected to choose courses of action, in re-

---

[47] *See, e.g.*, 42 U.S.C. §§ 604–619 (1994 & Supp. IV 1998); TEX. HUM. RES. CODE ANN. § 44.032 (Vernon 1990 & Supp. 2000).

[48] *See* Dana Milbank, *Welfare-Overhaul Legislation, Devised By Republicans, Is Signed by President*, WALL ST. J., Aug. 23, 1996, at A3; Christina Binkley, *Welfare Law Has Region Scrambling*, WALL ST. J., Aug 21, 1996, at S1.

sponse to governmental policies, that will benefit themselves at the expense of those very policy goals.[49]

In fact, if Clark is rational, he will attempt to aggravate the legislature's mistaken belief that he needs help. Clark will be in the forefront of legislative activity, hiring lobbyists and buying advertisements to communicate the impression that he is the victim of disadvantageous treatment, when in reality he has the greatest chance for survival. Human beings are strategic and may confront government strategically when it acts in its distributive role.[50]

A related idea is that apparently similar interventions may have widely different results. Thus, for example, the government would aggravate the inequality in the truel problem if it provided Clark secretly with an unseen bulletproof vest that warded off fifty percent of Art's shots.[51] The legislature would succeed more closely in leveling the playing field if it required scattering devices on Art's and Bart's guns (although if it did so with the intent of helping Clark, it would achieve the goal of equality by accident).[52] The point is that every option needs to be considered in light of its implications for strategy by those affected.

What do these issues mean in practical terms? Imagine an affirmative action program that operates by granting preferences to minority-owned businesses, defined as those in which members of specified minorities hold more than a certain percentage of shares.[53] This kind of statute likely will achieve its goal at least to some extent, in that it will enhance minority participation in industries in which it historically has

---

[49] For an analysis of this phenomenon in the context of scarce necessities, see generally GUIDO CALABRESI & PHILIP BOBBITT, TRAGIC CHOICES (1978). Calabresi and Bobbitt consider means of allocating the right to have children, access to organ transplants, and other "tragic choices," as well as the strategies of human beings who wish to maximize their chances in response to the allocation policies.

[50] *See, e.g., id.* at 121 (depicting potential strategies of people denied scarce resources); *id.* at 103–17 (discussing defects in wealth-neutral markets as a means of allocation and potential strategies used to exploit those defects).

[51] The vest would need to be modified if Clark faced Bart instead, and Bart would need a vest in his initial encounter with Art. The vest idea, then, would be clumsy. Theoretically, however, it could be done, and perhaps the concept is no more artificial than the truel problem itself.

[52] This is so because Art and Bart presumably would be aware of the government's action. If the government succeeded in equalizing the marksmanship of all players at fifty percent, the rational strategy would change. It then would be a matter of indifference whom the first shooter targeted, and each player would have a probability of survival of 1/3. Another way to look at the puzzle, however, is to realize that Clark's shoot-to-miss strategy still works in this modified game, since it always would be advantageous to have the first shot after one adversary has been killed. Perhaps the rational strategy of all three, then, would be to fire in the air. This strategy, however, would defeat the purpose of the game.

[53] *Cf.* Fullilove v. Klutznick, 448 U.S. 448 (1980) (upholding the Federal Public Works Employment Act of 1977, Pub. L. 95-28, 91 Stat. 116, which generally required a set-aside for minority business enterprises of ten percent of federal funds used in public works projects).

been low.[54] The difficulty, however, is that it also will precipitate strategic behavior that benefits individuals at the expense of program goals. For example, some individuals can be expected to reposition themselves as minority candidates when they really do not fit the classes intended to be protected,[55] and wealthy individuals of multiple backgrounds may be motivated to create pass-through businesses that perform no actual work so as to defeat the program goal of enhancing long-run participation.[56] It might be better, not only from a constitutional standpoint[57] but also from a legislative one, to use a different mechanism for the government's intervention, such as considering whether the business or its owners historically have been disadvantaged in their efforts to enter the industry.[58]

A very different example that illustrates the same point is provided by statutes that make mothers and fathers equal in seeking custody of their children upon divorce.[59] Historically, mothers sometimes have been preferred even when fathers were better parents, unless the mothers notoriously were unfit.[60] Modern laws endeavor to increase equality. They allow a good father to obtain custody over the claims of a not-as-good mother, serving the principle of evenhandedness as well as the best interest of children. But it would be a mistake to leave it at that, without considering the strategic behavior that modern laws precipitate. Fathers who do not want or expect custody, but who do want to reduce child support and to influence property division, sometimes plead for custody.[61] The risk-averse mother who calculates that the father's odds of obtaining custody are only one in five, but who is unwilling to take a twenty percent chance of losing her children, may see her best strategy as capitulation,[62] and thus legislative efforts to assure equality may result, tragically, in greater inequality.

---

[54] The act upheld in *Fullilove v. Klutznick* was premised on this idea. *See id.* at 459.

[55] *Cf.* Metro Broad., Inc. v. FCC, 497 U.S. 547, 633 n.1 (1990) (Kennedy, J., dissenting) (describing a categorization of an individual as "Hispanic" that depended on tracing the individual's family history all the way back to 1492, the year of Christopher Columbus's first landing). Another possibility is that of a Caucasian individual recently arrived from South Africa who checks the box for African American status. In *Fullilove*, Justice Stewart's dissent argues that "our statute books will once again have to contain laws that reflect the odious practice of delineating the qualities that make one person a negro and another white." 448 U.S. at 531 (Stewart, J., dissenting).

[56] *Cf. Fullilove*, 448 U.S. at 488.

[57] *See* Adarand Constructors, Inc. v. Pena, 515 U.S. 200 (1995) (invalidating a federal preference that presumed members of racial minorities and women to be eligible). Analysis of the constitutional issue is beyond the scope of this Article.

[58] *Cf. id.* at 238 (suggesting this solution).

[59] *See, e.g.*, ARIZ. REV. STAT. § 25-403(B) (Supp. 2001); CAL. FAM. CODE § 3040(a)(1) (West 1994 & Supp. 2001); TEX. FAM. CODE ANN. § 153.003 (Vernon 1996).

[60] *See* MARY ANN MASON, FROM FATHER'S PROPERTY TO CHILDREN'S RIGHTS: THE HISTORY OF CHILD CUSTODY IN THE UNITED STATES 123 (1994).

[61] *See* Saul Levmore, *Joint Custody and Strategic Behavior*, 73 IND. L.J. 429, 432 (1998).

[62] *See infra* note 292 and accompanying text (concluding that maximin strategy, or in other words concentrating solely on avoiding the worst result, may be mothers' best option

The corrective for this cynical "game" is not to turn back the clock or to give an automatic preference based on gender. Rather, it may be to adjust the game by a sanction that penalizes ill-conceived contests of custody. We must again sound the cautionary note, however, that this solution too will prompt strategic responses that might counteract its purpose,[63] but still, the adjustment may enhance equality. Thus, game theory might help to remind policymakers that legislating the equality of the genders, although desirable, will have side effects that should be considered.

Yet a deeper implication of the truel paradox is that equality is not a unitary concept. There are many theories of equality. Perhaps, for example, one can argue that the truel among Art, Bart, and Clark is, in fact, already equal, in the sense that each of the three participants has an equal chance that is influenced only by the skill and strategy of each. This phenomenon of many equalities may be better illustrated, however, by the next Part of this Article, which describes the "stockholders' game" and considers it in light of the theory of coalition formation in three-person encounters in an effort to tease out the multiple meanings of equality.

III. THE STOCKHOLDERS' GAME: COALITION FORMATION, DISTRIBUTIONS, AND THE MULTIPLE MEANINGS OF EQUALITY

A. *Describing the Stockholders' Game*

The stockholders' game is played among three owners of voting stock in a close corporation.[64] Let us call them Andi, Brandi, and Candi. Andi owns fifty percent of the shares, Brandi owns thirty percent, and Candi owns twenty percent. Since every action taken by the stockholders requires the vote of a majority of the shares, no action can be taken unless a voting coalition forms, and a successful coalition must include Andi, the fifty percent stockholder, although Andi cannot act alone.

The inciting event for the game is a windfall received by the corporation. Unexpectedly, the corporation has taken in thirty bars of gold. Andi, Brandi, and Candi must exercise their votes to distribute these thirty gold bars. Assume one additional key fact: the corporation will be dissolved immediately afterward. This plot twist means that the stockholders will have no reason to engage in vote trading, because there are

---

here).

[63] For example, custody seekers might use the threat of such a sanction as a weapon to discourage meritorious custody claims. In addition, the imposition of such a sanction would require the custodial parent to litigate the matter and thus to forgo the ostensibly rational strategy of capitulation, and therefore it may not prove effective.

[64] This framing of the game resembles that in COLMAN, *supra* note 1, at 165–66, although we use different numbers here to make the mathematics clearer. DAVIS, *supra* note 1, at 163–67, sets out several equivalent games with different framing.

no future votes to trade. Finally, we assume that Andi, Brandi, and Candi are motivated solely by informed, rational self-interest.

A game theorist would use these facts to set out what is called the "characteristic function" of the game.[65] The characteristic function simply maps the payoff from each possible relevant division by a voting coalition. In particular, it seeks to discover "imputations,"[66] or results of coalitions that conceivably could reflect the rational self-interest of their members. The first and simplest imputation is one that includes no one. This null set is symbolized by "Ø." Since a proposal that is voted for by no one in this situation cannot produce any action, it cannot distribute any gold bars, and game theorists say that the "value" of this imputation, the null set, is zero. In technical notation, we would write: $v(Ø) = 0$.[67]

Similarly, a vote by Andi alone, Brandi alone, or Candi alone cannot deliver any gold bars. Our notation is: $v(A) = 0$, $v(B) = 0$, $v(C) = 0$. So now, we look to imputations that contain more than one player. We quickly strike pay dirt, because the imputation composed of Andi and Brandi can muster a majority vote, enough to distribute all thirty gold bars. We write, $v(AB) = 30$. If we then analyze all of the other imputations, up to and including the "grand coalition" that is composed of Andi, Brandi, and Candi all together, we will produce the composite notation shown in Figure 2.[68] This is the characteristic function of the stockholders' game, showing the values or payoffs of all its included imputations.

FIGURE 2

$v(ø) = 0$
$v(A) = 0, v(B) = 0, v(C) = 0$
$v(AB) = 30, v(AC) = 30, v(BC) = 0$
$v(ABC) = 30.$

For a game theorist, the characteristic function would be the first step toward predicting coalition formation.[69] The game theorist would

---

[65] JOHN VON NEUMANN & OSKAR MORGENSTERN, THEORY OF GAMES AND ECONOMIC BEHAVIOR 240 (3d ed. 1953); *see also* COLMAN, *supra* note 1, at 163–64; DAVIS, *supra* note 1, at 179–81.

[66] An imputation is a "rational division of the payoff in which the players each receive at least as much as they would by acting independently, and they receive jointly as much as they would by acting as a grand coalition (nothing is wasted)." COLMAN, *supra* note 1, at 167. More technically, it is "a payoff vector that gives each player at least as much as he can guarantee himself and gives all players together v(N) [i.e., the value of the grand coalition]." FRIEDMAN, *supra* note 1, at 246.

[67] *See* COLMAN, *supra* note 1, at 165–66.

[68] *Cf. id.* (setting out an analogous characteristic function with different numbers).

[69] *Id.* at 163–66.

assume that all the players will follow the rules of the game faithfully, but that outside this constraint, all will act according to strategies motivated solely by rational self-interest.[70] We shall return to the game theorist's view in a moment, and we shall consider such concepts as Shapley values and minimum winning coalition theory[71]—tools of strategic analysis. First, however, we shall consider legislating about the concept of equality in the context of the stockholders' problem. The question is, "What on earth do we mean if we speak of an 'equal' distribution of the thirty bars of gold?"

B. Two Concepts of Equality in the Stockholders' Game: Individual Equality and By-the-Shares Equality

One way that Andi, Brandi, and Candi can achieve a kind of equality is to distribute equal numbers of gold bars to each. This distribution means that each of the three will receive ten gold bars. Notationally, we can write A = 10, B = 10, C = 10. Let us call this "individual equality."[72] It is diagramed in Figure 3.

But Andi would probably complain about a distribution that reflects individual equality: "It's not fair. I invested wisely in the beginning by buying fifty percent of the stock, and now I'm being treated the same as somebody else who didn't invest as wisely." In other words, Andi will complain about this distribution on equitable grounds. Moreover, Andi may say, "I will withhold my vote from an individual-equality distribution." Thus, there is great likelihood that no coalition based on individual equality will stay together in this situation. An individual equality solution, then, is unstable. Indeed, a game theorist would say that, for purposes of the stockholders' problem, an individual equality solution is neither "equitable" nor "stable."[73]

---

[70] *Id.* at 6–8.

[71] *See infra* Part II.D–.E.

[72] Calabresi and Bobbitt refer to this distribution as "formal" or "naive" egalitarianism, which "treats everyone alike." CALABRESI & BOBBITT, *supra* note 49, at 24–26. They suggest that this simplistic view of equality is unworkable in the more difficult circumstances of allocation of scarce goods, in which a more sophisticated model can achieve "a degree of egalitarianism much greater than its formal structure would suggest." *Id.* at 179.

[73] Stability has a technical definition, based on the concept of "dominance." One imputation is said to dominate another if there is a potential coalition that prefers it. COLMAN, *supra* note 1, at 168–69. Here, for example, Andi can form a coalition with Candi (or Brandi) that the two partners will find superior, simply by both taking fifteen bars. This Shapley solution also is said by game theorists to reflect the advantage of equity. *See infra* notes 77–80 and accompanying text.

FIGURE 3

Individual equality:	A = 10 B = 10 C = 10
By-the-shares equality:	A = 15 B = 9 C = 6

Another approach the stockholders could take is to distribute the gold bars according to the respective percentages of shares that each of them holds. This distribution would give fifteen bars to Andi, nine to Brandi, and six to Candi. Notationally, A = 15, B = 9, C = 6, as is indicated in Figure 3. Let us call this solution "by-the-shares" equality.[74]

Andi is more likely to regard this by-the-shares division as equitable. As for Brandi and Candi, both have reason to prefer individual equality, but neither has the votes to achieve that result, and being rational, both recognize this fact. Both may believe that the distribution is fair. By-the-shares equality is closer to an "equitable and stable solution," in game theory terminology.

C. A Third Possibility: Efficient-Use Equality

It would be a mistake, however, to regard these two possibilities as exhausting the universe of distributions that might be called equal. Individual equality and by-the-shares equality are only two of many possible equalities. There are, of course, bewildering numbers of possible distributions, ranging all the way from giving all thirty bars to Andi, who has the most shares, to giving all thirty to Candi, who has the least. Is it possible that this latter distribution, giving all thirty bars to Candi (who has only twenty percent of the shares), could be called equal?

Let us suppose that Candi is destitute, but she needs a lifesaving medical treatment that will cost exactly thirty gold bars. Andi and Brandi are multi-billionaires, both wealthier than Bill Gates. Candi argues, "A regime of equality would take into account the distributee's needs. The utility to me of the thirtieth gold bar, even after I have received twenty-

[74] Calabresi and Bobbitt refer to "laissez faire egalitarianism," which "requires equal opportunity to move within categories which may be properly the basis for discrimination, but begins in a context of variegated starting points, for instance, existing wealth distributions." CALABRESI & BOBBITT, *supra* note 49, at 25. Here, the starting point or existing wealth distribution is defined by the allocation of shares.

nine others, is greater than the utility of just one gold bar to either of you. If we adopt this rule of distribution, and if there is a parallel universe in which either of you is poor and moribund, you will get the benefit of this marginal-utility rule, and you will get all thirty bars. In this sense, my proposal treats you equally." We can call this distribution "efficient use equality,"[75] and it is diagramed in Figure 4. It posits that conceivably there can be a concept of equality that distributes all of a given resource to the ostensibly least entitled, while distributing nothing to those at least superficially more entitled. Sometimes, legislative choices about scarce necessities follow this approach. A regime of regulations governing organ donations, for example, may favor recipients who are most healthy, or (conversely) most likely to die quickly, instead of benefiting those most "deserving," most able to pay for the goods, or who have been on the waiting list the longest.[76]

FIGURE 4

Efficient use equality:	A = 0 B = 0 C = 30

And as we shall see, game theorists can put yet another spin on the concept. Is it unequal to take into account the voting strength of Andi, Brandi, and Candi? Perhaps not. The issue may depend upon whether we consider "equality of opportunity" or "equality of result." Thus far, we have considered only notions of equality that ignore foresight, strategy, and opportunity, and that focus on result. A game theorist might view the issues differently.

D. Using Game Theory: A Shapley Value Solution to the Stockholders' Game

A completely different way to approach the distribution of the thirty gold bars in the stockholders' problem is to consult game theory. The pioneering work of Lloyd Shapley points to one way in which voting

---

[75] Calabresi and Bobbitt refer to "corrected" egalitarianism, which accepts the general premise of formal egalitarianism that discrimination is proper as long as likes are treated alike, but "corrects the operation of this premise by rejecting it whenever methods applying it happen to produce results which correlate the permissible category of discrimination—health, for example—with an impermissible one, such as wealth or race." *Id.* at 25.

[76] *See, e.g.*, 42 C.F.R. § 121.8 (1999) (directing the Organ Procurement and Transplantation Network Board of Directors to establish allocation policies consistent with certain stated goals).

coalitions might result in a distribution that game theorists would consider relatively equitable and stable.[77] "Shapley values," as they are called, measure the contribution that each player can make to a potential winning coalition, or majority vote, in this case.[78]

Fundamentally, Shapley values signify the percentage of winning coalitions in which each player is "pivotal," or in which that player completes a majority.[79] There are, in fact, six ways in which winning coalitions can form in the stockholders' game. The six orders are shown in Figure 5. As this chart shows, Brandi is pivotal in one case, Candi is pivotal in one case, and Andi is pivotal in four of the six combinations. Each player's "Shapley value," as it is called, corresponds to that player's proportion of pivots. Thus, the Shapley value for Andi is 4/6, or 2/3, or 0.667. The comparable Shapley values for Brandi and Candi are each only 1/6, or 0.167.[80]

FIGURE 5

A then B; B is pivotal
A then C; C is pivotal
B then A; A is pivotal C then A; A is pivotal B then C then A; A is pivotal C then B then A; A is pivotal

A Shapley value solution, then, would distribute the thirty gold bars in the stockholders' problem by giving 2/3 to Andi and 1/6 each to Brandi and Candi. This division would mean that Andi would receive twenty of the thirty gold bars, while Brandi and Candi would receive five each. Notationally, A = 20, B = 5, and C = 5, as is shown in Figure 6. The division does not reflect individual equality, nor does it reflect by-the-shares equality; rather, it reflects equality as measured by the contribution of each player to the solution to the problem.

---

[77] *See* L.S. Shapley, *A Value for n-Person Games*, *in* 2 CONTRIBUTIONS TO THE THEORY OF GAMES 307–17 (H.W. Kuhn & A.W. Tucker eds., 1953).

[78] *See id.*; COLMAN, *supra* note 1, at 172–73 (discussing Shapley's solution); DAVIS, *supra* note 1, at 204–12 (discussing Shapley's solution).

[79] *See* COLMAN, *supra* note 1, at 173–74; DAVIS, *supra* note 1, at 210–12.

[80] COLMAN, *supra* note 1, at 173.

FIGURE 6

Shapley value equality:	A = 20 B = 5 C = 5

On its face, it seems obvious that this Shapley value solution is not equal. Andi has only half the stock, but receives the lion's share, 2/3 of the gold. Brandi has thirty percent, but she receives the same amount as Candi, who has only twenty percent.

A game theorist, however, might disagree with these criticisms and regard the Shapley solution as more reflective of equality than other solutions. In fact, game theorists call the Shapley solution "equitable," meaning that it reflects the contribution of each player to a solution, and it is relatively stable, meaning that it is likely to be arrived at by players who agree to it.[81] In real-world terms, perhaps one can say that simpler notions of equality, including individual equality and by-the-shares equality, are unequal, in that each artificially cobbles together a solution by ignoring the relative importance of actual votes. Arguably, those solutions are anti-democratic.[82] Further, one might defend the Shapley solution by pointing out that each of the stockholders must have known the consequences of this arrangement when she bought her shares. It is unlikely that Candi, for example, failed to realize that she was a minority shareholder and that this status would have an effect on the results of votes.

If these considerations are not enough, let us hypothesize that Andi, to obtain her fifty percent share, paid a higher per-share price, or undertook other sacrifices or work in order to gain her percentage. Such an extra payment is well-known in the world of business and finance: it is called a "control premium."[83] A person pays an additional price, in some circumstances, for a large block of shares precisely because it provides the prospect of a return greater than the absolute number of shares would

---

[81] *See id.* at 172–74. A Shapley solution, compared to some other relatively stable solutions, "has the virtue of equity [in] that the players' relative shares of the joint payoff are proportional to the average contributions that they provide in process of coalition formation." *Id.* at 174. It is "the best-known and most widely used solution concept for n-person cooperative games." *Id.* at 172. There are, however, solutions that are more stable, and thus a Shapley solution is only "relatively" stable.

[82] But then again, virtually every intervention designed to assure equality can be assailed as anti-democratic, since it counteracts majority will. *Cf.* CRUMP ET AL., *supra* note 20, at 10–12 (collecting commentary on the allegedly anti-democratic nature of judicial protection of minorities).

[83] *See* BLACK'S LAW DICTIONARY 1200 (7th ed. 1999).

predict.[84] Given this consideration, perhaps it is unequal for Andi, whose shares cost more, to receive the same amount per share as Brandi or Candi.

Shapley's theory is not limited to three-person games or to the division of thirty gold bars. Instead, it is general, and it applies to any number of players and to an infinite variety of decisions. The equation for the Shapley value V_I for any player, I, in the general case is as follows:[85]

$$V_I = \sum_{i=1}^{n} \frac{[(s_i - 1)!\,(n-s_i)!]}{n!} \bullet D(S_i),$$

where n is the total number of players, S_i is the number of players in the ith coalition, and $D(S_i)$ stands for the difference in value of the ith coalition with, and without, player I (or if I is not in S_i, then $D(S_i) = 0$).

Furthermore, Shapley's theory is applicable to a wide variety of real-world situations, although it is like every game theory model in that it provides only a map, and a limited one at that, rather than a picture of the entire territory. For example, using assumptions and approximations, it is possible to compute Shapley values for legislative power in its distribution among the President, Senators, and members of Congress. Shapley and Martin Shubik assigned power indices of 350 to the President, 9 to each Senator, and 2 to each member of Congress.[86] Obviously, these numbers are prototypes only, in that they do not take into account such factors as seniority, committee chairmanships, or even persuasive personality.[87] Nevertheless, they provide a fascinating insight into the relative strengths of the players on the national legislative scene.

---

[84] *Cf.* DAVIS, *supra* note 1, at 210 ("[T]he power of a coalition [is] *not* simply proportional to its size; a stockholder with forty percent of the outstanding stock, for example, would actually have about two-thirds of the power if the other sixty percent of the stock was divided equally among the other six hundred stockholders.").

[85] For slightly differing but equivalent notations, see COLMAN, *supra* note 1, at 173; DAVIS, *supra* note 1, at 207; FRIEDMAN, *supra* note 1, at 267. The Shapley value for player I is conventionally denominated as $N_I(v)$, but this Article uses a simpler symbol, V_I.

[86] L.S. Shapley & Martin Shubik, *A Method for Evaluating the Distribution of Power in a Committee System*, 48 AM. POL. SCI. REV. 787, 789 (1954). Shapley and Shubik assigned power indices of five to the Senate as a whole, five to the House of Representatives, and two to the Presidency. *Id.* at 789. They based their calculations on the fact that to pass legislation, it takes a majority of the Senate and House of Representatives with the President or two-thirds of each without the President. *See id.* Shapley and Shubik arrived at these numbers in 1954, when there were ninety-six Senators. *See id.* at 788.

[87] *Cf. id.* at 791 ("It would be foolish to expect to be able to catch all the subtle shades and nuances of custom and procedure that are to be found in most real decision-making bodies.").

E. Minimum Winning Coalition Theory and Minimum Winning Resource Theory

Minimum winning coalition theory undermines the argument that Shapley values represent equality, or at least tends to reduce one's faith in a solution based upon contribution to a majority. Minimum winning coalition theory, as its name implies, concludes that the smallest possible voting coalition will form and carry the day.[88] Coalition members motivated by rational self-interest will not unnecessarily add others to their number with whom they would have to share, because the addition would only reduce their own distributions.[89] Thus, instead of the grand coalition forming among Andi, Brandi, and Candi, it is conceivable that Andi will approach Candi alone and say, "I need only one of you to complete a majority. If you vote with me to distribute the gold bars, we will cut Brandi out completely, and you and I will each get Brandi's 'share' of the gold." Thus, for example, Andi and Candi might agree to begin with a Shapley solution—A = 20, B = 5, and C = 5—but to change it, by dividing the five bars that otherwise would go to Brandi, with Andi taking three and Candi taking two. Thus, one reaches the solution, A = 23, B = 0, C = 7, as shown in Figure 7.[90]

FIGURE 7

Minimum winning *coalition* theory:	A = 23 B = 0 C = 7
Minimum winning *resource* theory:	A = 0 B = 15 C = 15

The question arises, if a Shapley value distribution based upon the contribution of each player to a solution is legitimately to be called equal, then why is a solution reflecting a minimum winning coalition not also equal? It, too, reflects contributions to a solution; the solution simply is

---

[88] *See generally* WILLIAM H. RIKER, THE THEORY OF POLITICAL COALITIONS (1962) (discussing theories that predict coalition formation); *see also* COLMAN, *supra* note 1, at 179 (attributing this theory to Riker).

[89] *See* COLMAN, *supra* note 1, at 179.

[90] This, however, is only one possible division between the two coalition members. They conceivably could divide the thirty bars by taking fifteen each. But this solution is unstable, because Andi then would be tempted to bargain with Brandi for a share larger than fifteen.

more lopsided. Yet, the notion of two stockholders colluding to freeze out a third, so that they can take that person's allocation to their own sole advantage, does not sound like equality. It sounds like the raw exercise of power, in denial of equality. Thus, many people who might be persuaded to accept a Shapley value solution might reject a concept of equality based upon minimum winning coalition theory by proposing legislation that would outlaw it. Paradoxically, by taking this position, it is arguable that they also undermine the legitimacy of the Shapley value solution they have been persuaded to accept as equal, because both distributions are based on voting power and coalition formation in the same way.

Is there a way to defend the result of minimum winning coalition theory, on the ground that it provides a kind of equality? Imagine that Andi and Candi have good extrinsic reasons to freeze out Brandi. Hypothesize, for example, that Brandi has just murdered the plant supervisor. Or, in the alternative, consider that Brandi has decided to register as a Republican—something that Andi and Candi, both committed Democrats, find nearly as repugnant as murder. Is it equal, then, for Andi and Candi to agree to a division in which A = 23, B = 0, and C = 7, as a means of disapproving Brandi's act of murder? Perhaps it can be argued that a regime is equal when it regards pathological actions such as Brandi's intentional homicide as forfeiting the right to a different model of equality, making minimum winning coalition theory equal in this case. This solution seems less palatable if it is based upon Brandi's choice of party affiliation (although game theory cannot explain why).

But we are not yet through exploring different kinds of distributions. There also could be a distribution based on "minimum winning *resource* theory."[91] To explore this possibility, we need to change the players' stock percentages. Assume, now, that Andi holds forty percent, Brandi thirty percent, and Candi thirty percent of the stock. This alteration changes the entire dynamic, because Andi no longer is required for a winning coalition. In fact, any two of the three can form a majority. Minimum winning resource theory predicts not only that the smallest possible coalition will form, consisting of only two out of the three stockholders, but that it will be the coalition that contains the fewest number of votes that still are enough for a majority.[92] In other words, Brandi and Candi will coalesce with their sixty percent of the votes, and they will take fifteen bars each. Thus, A = 0, B = 15, C = 15, as is shown in Figure 7. Minimum win-

---

[91] *See* W.A. Gamson, *A Theory of Coalition Formation*, 26 AM. SOC. REV. 565 (1961) (discussing the theory); *see also* COLMAN, *supra* note 1, at 180 (attributing the theory to Gamson).

[92] Specifically, the theory assumes that each coalition member rationally will demand a share corresponding to that player's relative voting strength. *Cf. supra* note 66 (explaining this demand in the context of the stockholders' problem). For this reason, a rational player seeking to form a coalition will seek out the weakest possible set of partners that can form a winning combination.

ning *coalition* theory, in other words, predicts the smallest number of coalition members, whereas minimum winning *resource* theory predicts the smallest possible majority vote.

Paradoxically, minimum winning resource theory thus suggests that it is the player holding the largest number of votes, the biggest gorilla in the pack, that is in the weakest position. In real-world terms, one can imagine Brandi and Candi, both fearful because each has less than the largest number of votes, as being prone to coalesce because they are in the same "weaker" position. This power inversion sometimes is referred to as the "chairman's paradox."[93] It explains why, in some situations, the United States is unable to form coalitions with weaker nations, coalitions that would seem to make economic and political sense.[94]

There are additional methods of analyzing coalition formation. Game theory provides much more complicated formulations,[95] but we would pass the point of diminishing returns to explore the more difficult concepts that are necessary for other coalition-formation theories.

But before departing from Shapley values, minimum winning coalition theory, and minimum winning resource theory, it is worth observing what happens among real experimental subjects who play games like the stockholders' problem. Psychologists have performed such experiments, and while the results are mixed, they tend to show that some subjects, in situations comparable to the stockholders' problem, tend to avoid either individual equality or by-the-shares equality. They also fail to conform exactly to Shapley, minimum winning coalition, or minimum winning resource solutions. Experimental subjects tend, in other words, to follow neither simplistic theories of equality nor more complex theories about coalitional power. The distributions they negotiate tend to fall somewhere between simple equality and pure coalitional power solutions.[96] Perhaps they reflect a compromise between an instinct favoring individual equality and the stability that comes from wider disparities in distribution.

---

[93] COLMAN, *supra* note 1, at 265–66.

[94] *Id.*

[95] *Cf.* COLMAN, *supra* note 1, at 167–69 (discussing solutions based on the "core," or set of imputations that provide each participating player a greater payoff than the grand coalition would, and the "stable set," or set of imputations exhibiting stability); *id.* at 174–77 (developing further solutions based on "kernel," "nucleolus," and "least core"); FRIEDMAN, *supra* note 1, at 247–66 (discussing the same concepts with symbolic notation).

[96] *See generally* COLMAN, *supra* note 1, at 180–85 (discussing psychological literature). "[T]he evidence consistently shows that subjects tend to reach agreements somewhere between this parity split [i.e., that predicted by minimum resource theory] and an [individually] equal division of the payoff." *Id.* at 184.

F. Implications of the Stockholders' Problem for Notions of Equality and Tradeoffs with Other Values

The Fourth of July speeches that politicians give often praise equality and freedom as though they represented the same thing.[97] Furthermore, they regard equality as a unitary value. Although general statements in favor of freedom and equality may be useful, the stockholders' problem provides a reminder that the world is not that simple.

First, the stockholders' game shows us that equality is not, in fact, a unitary value. There are many equalities,[98] and something can be said for each of them.[99] For this reason, it is difficult to reject any of these conceptions uniformly and in all cases. When confronted with a demand for equal treatment, it is appropriate for a legislator to ask, "Exactly which of the many conceptions of equality do you mean?"

A corollary to this observation is that legislative interventions to enhance equality must decide among competing visions of equality. If, for example, a legislature decides to prevent individuals from arriving at solutions based upon Shapley values, minimum winning coalition theory, or minimum winning resource theory, it must disapprove these solutions on the ground that they are unequal (or on some other ground). This statutory regime may reflect highly debatable theories of equity and stability. Furthermore, the legislature must favor another theory or class of theories of equality by labeling the preferred approach as legitimate.

A third insight from the stockholders' game, related to the first two, is that every legislative intervention to secure equality results in some decrease in freedom. Often, we may decide that the gain is worth the loss, but nevertheless this factor must be considered. Thus, in the case of the stockholders' problem, imagine that Andi and Candi vote in favor of a Shapley solution: A = 20, B = 5, C = 5. Brandi, however, is dissatisfied, dissents from the vote, and files suit. Imagine further that a state statute supports Brandi's position, because it provides that a divi-

---

[97] For example, President Bill Clinton's 1998 Independence Day message included the following:

> We are all heirs to the rights articulated in our Constitution and reaffirmed by courageous men and women of every generation who have struggled to secure justice and *equality* for all. We are all forever indebted to the millions of Americans in uniform who have shed their blood to defend our *freedom* and preserve our values across America and around the globe.

Clinton Stresses 'Common Heritage as Americans,' WASH. TIMES, July 4, 1998, at A6 (emphasis added).

[98] As Calabresi and Bobbitt put it, "no society adheres wholly to one conception of equality." CALABRESI & BOBBITT, *supra* note 48, at 25–26.

[99] *See id.* (reflecting ambivalence across societies about different models of equality); *see also supra* Part II.A–.E of this Article (justifying different and inconsistent allocations as arguably all equal).

sion according to individual equality, A = 10, B = 10, C = 10, is the only lawful alternative. For better or for worse, the legislature has curtailed the freedom of Andi, Brandi, and Candi to manage their own affairs. Furthermore, it has interfered with the freedom of future shareholders who might wish to structure their businesses as Andi, Brandi, and Candi have attempted.

Similarly, a law against housing discrimination prevents a large apartment building owner from discriminating against a prospective tenant on racial grounds.[100] Society recognizes that this law curtails the freedom of the landlord to do business with those he chooses, but we regard the gain in equality as overwhelming the loss in freedom. Thus, when the government intervenes to ensure equality, it must be because we as a society have determined that the gain in equality exceeds the loss in freedom. Significantly, the housing discrimination law contains an exemption for small landlords who take in only a few tenants and who reside on the premises.[101] This landlord is free to discriminate on racial grounds, however foolish we may think such an action. The small impact of this discrimination on the housing market, the interest of a resident in controlling the interior of his or her home, and similar factors, may lead us as a society to believe that it is more important to allow this individual to act autonomously, even with foolish freedom, than to mandate equality.

Other aspects of the clash between equality and freedom can be seen in decisions concerning California's Unruh Civil Rights Act ("Unruh Act").[102] The Unruh Act is a broad prohibition of discrimination and a mandate of equal treatment in any business situation.[103] The Act is not limited to traditional categories of invidious discrimination: it is open-ended, prohibiting discrimination generally. In one celebrated case, a religiously motivated landlord refused to rent to an unmarried couple.[104] They sued, in part, on the theory that it was a violation of the Unruh Act to treat them differently from a married couple.[105] Although the appellate court held the Unruh Act inapplicable, upholding the authority of the landlord to exercise the freedom of her religious convictions and to ex-

---

[100] *See* 42 U.S.C. § 3604(a) (1994).

[101] *See id.* § 3603(b)(2) (exempting "rooms or units in dwellings containing living quarters occupied or intended to be occupied by no more than four families living independently of each other, if the owner actually maintains and occupies one of such living quarters as his residence").

[102] CAL. CIV. CODE § 51 (West Supp. 2001) (providing that "all persons within the jurisdiction of this state are free and equal, and no matter what their sex, race, color, religion, ancestry, national origin, disability, or medical condition are entitled to the full and equal accommodations, advantages, facilities, privileges, or services in all business establishments of every kind whatsoever").

[103] *See* Bd. of Dirs. of Rotary Int'l v. Rotary Club, 481 U.S. 537 (1987) (upholding against freedom of association attack the application of the Unruh Civil Rights Act to prevent a private club from excluding women).

[104] Smith v. Fair Employment in Hous. Comm'n, 913 P.2d 909 (Cal. 1996).

[105] *Id.* at 913.

clude the unmarried couple, the California Supreme Court denied the landlord the liberty to act consistently with her perceived religious duty.[106] Again, every intervention to secure equality, whether well- or ill-conceived, implicates some curtailment of the freedom or autonomy of those whom it impacts.

Interestingly, these conclusions fit comfortably with a great deal of equal protection and due process analysis in the decisions of the United States Supreme Court. The prospect that many different equalities might be defensible counsels caution in intervention, particularly to the extent that what we are doing is fine-tuning. A law such as the Unruh Act arguably violates this principle. To put the matter another way, the government should be conservative (conservative with a small "c," rather than politically conservative) in mandating a particular version of equality. In fact, the Supreme Court's decisions square with this conclusion, in that the usual application of the Equal Protection Clause presumes constitutionality.[107] Upholding a legislative act requires only a "rational basis" for concluding that the classification will serve a "legitimate" governmental interest.[108] Any reasonable defense of a distribution of rights will suffice in most economic and social situations. The Supreme Court followed a tortuous pathway of blind alleys, discredited decisions, and reversals before settling upon the rational basis test.[109] But perhaps game theory explains why it reached this destination: there are many equalities, and all of them are potentially defensible in the sense that there is no principled basis for definitively rejecting any of them.

G. The Case for Legislative Intervention: Game Theory and Equality

It would be a mistake, however, to conclude that this minimalist approach to equality is all that game theory can teach us. Should a legislature accept theoretical coalitional contributions such as Shapley values as indicators of equality? If so, should it accept minimum winning coalition theory, in which two shareholders can agree to freeze out a third for no reason other than their own advantage? Game theory demonstrates that

---

[106] *Id.* at 918.

[107] *See, e.g.*, United States v. Carolene Prods. Co., 304 U.S. 144, 152 n.4 (1938). This famous "footnote 4" sets out the *Carolene Products* formula, which both recognizes the presumption of constitutionality and suggests situations in which it might be overcome. *See id.*

[108] *See, e.g.*, FCC v. Beach Communications, Inc., 508 U.S. 307 (1993) (upholding a distinction between a cable television operator serving multiple customers and an operator serving buildings under common ownership and concluding that, under rational basis review, any plausible or conceivable basis will suffice).

[109] *See, e.g.*, Lochner v. New York, 198 U.S. 45 (1905) (striking down a state law limiting work hours of bakery workers as a violation of liberty of contract); *see also* CRUMP ET AL., *supra* note 20, at 311–27 (collecting due process cases).

rational self-interest is strategic. It depends upon power in a way that sometimes seems inconsistent with equality.

A reasonable legislator might conclude that minimum winning coalition solutions, such as that in which Andi and Candi coalesce to freeze out Brandi, do not represent an ethically acceptable vision of equality in the general case. And, since the Shapley value solution also depends upon voting power, perhaps some legislators would reject it too. The difficulty with this conclusion is that one easily can imagine factors that would make the Shapley solution, or even the minimum winning coalition solution, seem equitable in a given case.

At some point, society simply must cut the Gordian knot and pronounce some distributions, such as those based on minimum winning coalition theory, unequal.[110] In the absence of this intervention, there cannot be any protection of minority shareholders. Indeed, there cannot be protection of equality in any circumstances. The difficulty is in recognizing precisely where this intervention is justified in the absence of principles that would enable us to distinguish the cases by logic.[111]

How does a legislature best implement these episodic protections of equality, given that there is an indefinite range of different kinds of equality, when it lacks principled means for distinguishing them? First, as the previous section indicates with regard to courts, by avoiding fine-tuning. We must not attempt to intervene in every instance when intimations of an unequal distribution mildly disturb us. But second, we must have reliable, hard-edged rules that tell us to change the game when the outcome depends too much, in our unprincipled but necessary judgment, on power rather than equity. In a later section, we shall examine a game theory analysis of a game called "Chicken," and that analysis will reinforce the conclusion that sharp, clear rules, as opposed to vague ones, may be necessary for protecting equality.[112]

This conclusion is reflected in the decisions of the Supreme Court concerning the constitutionality of statutes. Heightened scrutiny is triggered by regulations that impinge upon suspect classes or fundamental rights.[113] Exactly what constitutes a suspect class or a fundamental right continues to elude us.[114] Courts have recognized special classifications in

---

[110] "[T]he outcome of the tragic choice depends mainly on its relationship to a particular culture's notion of when it is right to accord some [people] a good and let others suffer, and sometimes die, without it . . . ." CALABRESI & BOBBITT, *supra* note 49, at 25.

[111] "[B]ecause [every] conception [of equality] is held in tension with other, antagonistic conceptions of equality, the resolution gained is temporary and the city is under siege almost as soon as it has been subdued." *Id.* at 24.

[112] *See infra* Part IV.B.2.

[113] *See* CRUMP ET AL., *supra* note 20, at 565–66.

[114] *See generally* David Crump, *How Do the Courts* Really *Discover Unenumerated Fundamental Rights?: Cataloguing the Methods of Judicial Alchemy*, 19 HARV. J.L. & PUB. POL'Y 795 (1996).

the case of illegitimate children,[115] for example, but not in the case of disabled people,[116] and they categorize abortion as a right deserving of heightened protection,[117] but not education.[118] This complexity comes from the indeterminacy of equality. But even with that fuzziness, we have drawn certain bright lines; racial classifications, for example, survive only if they are "narrowly" targeted to serve "compelling" state interests.[119] This strict scrutiny is as close to an across-the-board prohibition as the law typically comes, and it arguably is necessary to counteract the indeterminacy of equality. Game theory, perhaps, helps us to understand why.

At some point, if equality is to be a value that will be advanced by society through government, our legislatures and Courts will have to decide that beyond a certain point, allowing individuals to exploit strategic advantages to the disadvantage of others is unequal, even if some theory of equality can be advanced to justify it. There always will be an element of circularity to this decision. We may label potential or prohibited distinctions as "arbitrary," "irrational," "discriminatory," or "invidious," but what we really mean by these terms is that our society has made the judgment that the actions or strategies they represent should be prohibited.[120]

Without this sort of legislative judgment, minority shareholders, such as Candi in the stockholders' problem, could obtain no protection for their investments. Nor could other citizens insist upon equality in any other circumstance. The paradox, then, is that there are many equalities, including distributions that direct the lion's share to ostensibly undeserving persons, and many of them can be defended at least in some manner, even those that at first appear oppressive.[121] But if we are to retain anything resembling the value of equality at all, in the face of claims that seem to be based upon raw power but that actually can be defended as "equal" by one theory or another, there must be an enforcement of

---

[115] *E.g.*, Reed v. Campbell, 476 U.S. 852 (1986) (applying middle-tier scrutiny to categorization based on illegitimacy).

[116] *E.g.*, City of Cleburne v. Cleburne Living Ctr., 473 U.S. 432 (1985) (applying lowest-tier scrutiny to a classification based on mental retardation). The ambivalence of this holding, however, is demonstrated by the Court's failure to find a rational basis among the various plausible possibilities offered by the city. One can argue that *Cleburne* actually applies a de facto kind of heightened scrutiny.

[117] *E.g.*, Planned Parenthood v. Casey, 505 U.S. 833 (1992) (declining to adopt the usual standard of strict scrutiny of laws impinging on fundamental rights, but nevertheless implementing a version of heightened scrutiny by protecting the abortion right against "undue burden").

[118] *E.g.*, San Antonio Indep. Sch. Dist. v. Rodriguez, 411 U.S. 1 (1973).

[119] *E.g.*, Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 227 (1995).

[120] *See* Crump, *supra* note 114, at 795 (cataloguing methods for recognizing fundamental rights used in actual decisions of the Supreme Court and comparing them to "judicial alchemy").

[121] *See supra* Part II.A–.F.

norms that reject some of these proffered equalities and label them as, actually, unequal.

IV. Insights into Equality from Two-Person Zero-Sum and Mixed-Motive Games

In this Part, we consider two-person games, examining the insights into legislation affecting equality that might be gained from them. "Zero-sum" games are those in which everything gained by one player must have been lost by the other.[122] Chess, where there must be a loser if there is to be a winner, and poker, in which every amount won by one player must have come from another, are examples of zero-sum games. Mixed-motive games, on the other hand, are those in which there is an element of competition, but in which cooperative strategies may enlarge the pie that the players share.[123]

Two-person zero-sum games may seem to have relatively little to offer by way of insights into distributive equality. These are purely competitive games. Nevertheless, we begin this discussion with zero-sum games for two reasons. First, a regulatory regime providing equal opportunity or procedural justice can be modeled by zero-sum games. In this sense, equality corresponds to what is called a "value of the game" equal to zero.[124] Second, and more importantly, zero-sum games are sound vehicles for developing some basic tools of game theory, particularly the devices known as "payoff matrices."[125] Later, we shall use these tools for other games that have more to do with legislative problems involving equality.

A. *Payoff Matrices for Two-Person Zero-Sum Games: Developing the Tools of Game Theory*

An example of a zero-sum game is based on a famous naval engagement called the "Battle of the Bismarck Sea."[126] During the middle years of World War II, Allied intelligence learned that the Japanese intended to sail a major convoy around the island of New Britain to supply Japanese forces in New Guinea.[127] The Allied Commander, General George Kenney, had to decide whether to send his reconnaissance aircraft north or south of the island, and similarly, the Japanese admiral, Hitoshi Imamura, had to decide whether to sail north or south.[128]

---

[122] *See* Davis, *supra* note 1, at 14.
[123] *See* Colman, *supra* note 1, at 100.
[124] *See infra* Part IV.A.
[125] *See infra* Part IV.A.
[126] *See* Davis, *supra* note 1, at 13–14.
[127] *Id.* at 13.
[128] *Id.*

The resulting encounter, the Battle of the Bismarck Sea, has been analyzed frequently both by naval historians and by game theorists.[129] The strategy considerations are ideal for illustrating what are called "payoff matrices."[130] Kenney knew that if Imamura sailed north, north-stationed reconnaissance would detect the Japanese convoy in time to allow for two days of Allied bombing. However, if Kenney stationed his reconnaissance to the south, it would detect the north-sailing fleet in time for only one day of bombing. On the other hand, north-stationed reconnaissance still would detect the fleet in time for two days of bombing if Imamura sailed south, but if Kenney stationed reconnaissance to the south, the Allies could bomb a south-sailing Japanese convoy for a full three days.[131] Every advantage gained by one commander reflected a congruent loss by his adversary. Thus, Kenney and Imamura faced each other in a zero-sum game.

A payoff matrix is a visual display of strategy combinations and their resulting payoffs. Two players, each with two strategies, will produce a two-by-two matrix with four payoff squares. Figure 8 is a payoff matrix for the Battle of the Bismarck Sea. Customarily, Player 1 (Kenney, here) is listed to the left, and his strategy is shown by the horizontal rows; Player 2's (Imamura's) strategies are reflected by the vertical columns.[132] Thus, if we want to see the outcome or payoff when Kenney searches north and Imamura sails north, we look to Row 1 (horizontal) and Column 1 (vertical) to see the payoff, which is two days of bombing. In other words, the matrix shows us that if Kenney searches north and Imamura sails north, the payoff for each is two days of bombing.

How, then, does game theory offer strategic solutions? It might seem that, without knowing the opponent's intentions, neither side would have a clear strategy in the Battle of the Bismarck Sea.[133] Paradoxically, this is not true. Even under conditions in which players act secretly, there is a single rational strategy for both parties in this game.[134] The outcome, in the real Battle of the Bismarck Sea, actually reflected these logical strategies. Kenney sent his reconnaissance to the north, Imamura sailed north, and the Allies bombed the Japanese fleet continuously for two days. The battle resulted in the worst Japanese naval loss up to that time: twenty-two ships and fifteen thousand troops.[135]

---

[129] *See, e.g.*, COLMAN, *supra* note 1, at 54–57 (analyzing the same game, with references to analyses by others).

[130] *See* DAVIS, *supra* note 1, at 13 (constructing a payoff matrix equivalent to the one shown here).

[131] *Id.*

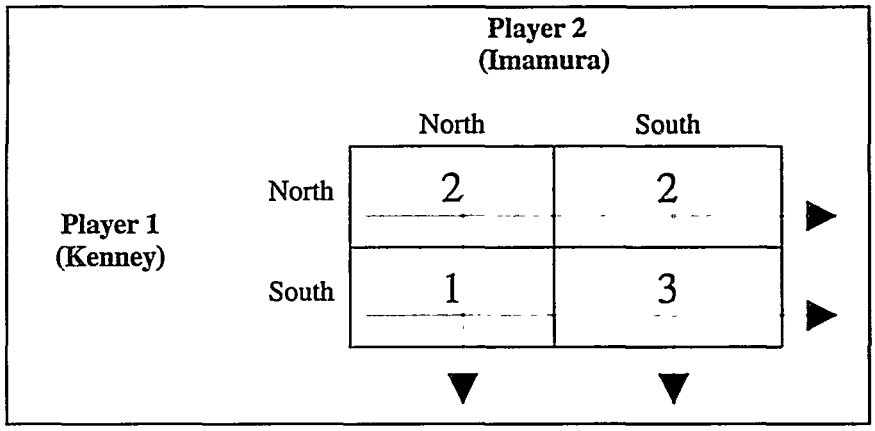
[132] *See id.*

[133] *Id.* at 13–14.

[134] *See id.*

[135] *See* COLMAN, *supra* note 1, at 58.

FIGURE 8
PAYOFF MATRIX FOR THE BATTLE OF THE BISMARCK SEA



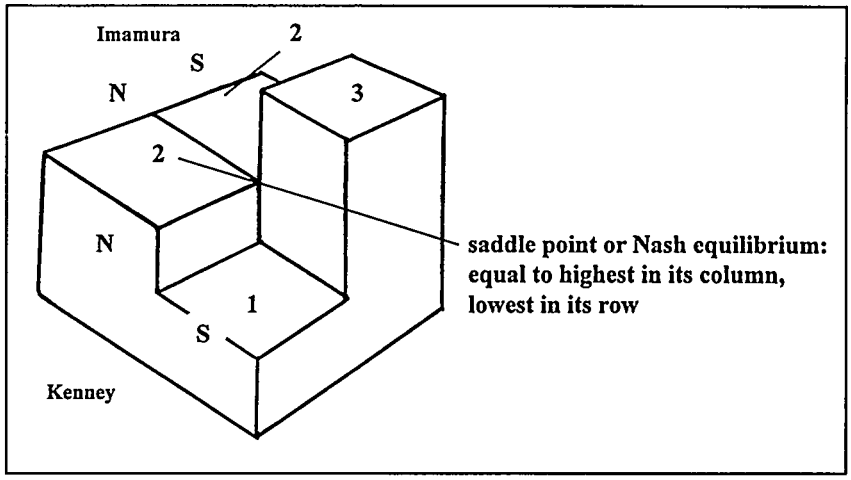
This outcome was dictated by rational strategies on the parts of both players. The payoff matrix for this game contains what is called a "saddle point" or "Nash equilibrium."[136] In other words, there is a single payoff square that is dictated by the logical choices of both parties. The structure of the payoff matrix, with higher and lower payoffs encompassing this square to make it resemble a saddle, is what dictates the outcome. We say, therefore, that the "value of the game" is two days of bombing.[137] We proceed, now, to examine saddle points or Nash equilibria.

In the Battle of the Bismarck Sea, the strategies that minimize the worst outcome for each player coincide in the upper left box of the payoff matrix, the north-north box. This point is called a "saddle point" because it is equal to the highest point in one direction, its column, and equal to the lowest one in another, its row—just as a saddle sits at the highest point between the horse's flanks and at the lowest point between its neck and hind.[138] A saddle point is also called a "Nash equilibrium," after John Nash, who contributed importantly to the theory.[139] Figure 9 shows this concept with a three-dimensional projection, demonstrating visually why this is a saddle point: equal to the highest in its column and to the lowest in its row.[140]

---

[136] *Id.* at 58–60.
[137] *See* DAVIS, *supra* note 1, at 19.
[138] *See* COLMAN, *supra* note 1, at 59.
[139] *Id.* at 58–59.
[140] This depiction of the concept is original with the author of this Article.

FIGURE 9
THREE-DIMENSIONAL PROJECTION OF THE PAYOFF MATRIX, SHOWING SADDLE POINT VISUALLY



Game theory in this situation specifies what is called a "minimax" strategy: the minimization of opportunity loss against an equally rational opponent. We shall explore this minimax strategy in a later section.[141] For now, let us examine each player's options and the outcomes associated with them. To find each player's optimal strategy, we first consider Imamura's options. Imamura wants low numbers, or fewer days of bombing. His north boxes are (2, 1) and his south boxes are (2, 3). If he sails north, Imamura will suffer either one or two days of bombing, depending on Kenney's choice; if Imamura sails south, he will suffer either two days or three, depending on the same choice. Thus, the north column is in all instances either equal to or better than the south column for Imamura. In game theory terminology, therefore, we say that the north column "dominates" the south one for Imamura and that south is "inadmissible."[142] Imamura must sail north; any other strategy is irrational.

For Kenney, neither column is dominant in the same way, but the minimax strategy of minimizing opportunity loss indicates that Kenney's correct move is north. This is so because if Kenney contemplates a rational Imamura, he knows that Imamura must choose north, the dominant strategy, since south for Imamura is inadmissible; therefore, Kenney must choose north, which is dominant (where he will have two days of bomb-

[141] *See infra* Part V.B.3.
[142] *See* COLMAN, *supra* note 1, at 69; DAVIS, *supra* note 1, at 20.

ing) against south (where he would have only one day of bombing, which is worse for him and therefore inadmissible).[143]

Whenever there is a saddle point, or Nash equilibrium, it determines the dominant strategy.[144] We say that it fixes the "value of the game," or the outcome of reciprocal rational strategies. Here, the value of the game is two days of bombing.[145] Notice that this analysis assumes a rational opponent. It does not consider strategies that might precipitate errors by the adversary, because game theory solves problems by the application of logic. Of course, there are many applications of psychological strategies, but they are distinct from game theory, which cannot predict how real human beings will make or attempt to precipitate moves that might contradict logic.[146] Nevertheless, game theory is a valuable tool. During World War II, the thrust of American military strategy was based on enemy capabilities, not on efforts to guess enemy intentions.[147] This approach makes analysis of the zero-sum game even more relevant.

What are the implications of a zero-sum game for legislation affecting equality? By definition, a zero-sum game does not always produce equal outcomes among real players, unless it is relatively trivial.[148] Equality of opportunity, however, can be modeled by zero-sum games. It corresponds to a value of the game equal to zero because zero means that each player has the same opportunity to come out ahead, depending solely upon the value of the strategy adopted. A value of zero, in other words, corresponds to the metaphor of a level playing field. The game of paper-scissors-stone has a value of zero;[149] the Battle of the Bismarck Sea does not.

Perhaps the most important reason for considering these questions of equality here is to expose the tools of game theory, particularly payoff

---

[143] *See* COLMAN, *supra* note 1, at 57–59.

[144] There also are games in which payoff matrices have no saddle points. In such a situation, the rational player adopts what is known as a "mixed strategy." *See* DAVIS, *supra* note 1, at 29. The player contemplates the rational opponent, computes frequencies for each move that maximize the payoff over many games, and then uses a randomizing device to determine each move by chance with a probability corresponding to the maximizing frequency. *See id.* at 28–31.

[145] *See* COLMAN, *supra* note 1, at 58.

[146] *See id.* at 59.

[147] *Id.* at 57.

[148] For example, complex games such as chess do not always produce predictable outcomes among masters, even though we know that there is a theoretically optimal strategy that would produce either a draw, a win for white, or (conceivably) a win for black. (John Nash received the Nobel Prize in Economics for 1994 in part for proving that every finite game has at least one equilibrium point or solution in pure or mixed strategies, even though we do not know what it is for games as complex as chess. *See* COLMAN, *supra* note 1, at 14, 186.) On the other hand, simpler games such as tic-tac-toe, in which most adult players know how to force a draw, have more obvious solutions.

[149] The rational (and relatively simple) strategy of a paper-scissors-stone player is to select randomly among the three possible moves with equal probability. This strategy gives each player an equal chance. The value of the game is zero.

matrices. In the section that follows, we use this device to explore mixed-motive games, which have more sophisticated implications for equality.

B. Two-Person Mixed-Motive Games

In order to understand mixed-motive games, we must recognize that a given payoff may be different for different players.[150] This result can happen either because the absolute payoffs are different or because the players derive different utilities from them.[151] For example, a given combination of moves may result in a payoff of one for Player 1 but may give a bigger payoff, let us say two, to Player 2. This structure can lead to matrices that reward cooperation to a degree, within an overall regime of rational self-interest.

There are four mixed-motive two-person games that have received the most attention from game theorists. These four situations have colorful, well-recognized names that reflect their traditional descriptions: "Battle of the Sexes," "Chicken," "The Prisoners' Dilemma," and "Leader."[152] All four of these games result from arrangements of differential pairs of payoffs valued from one, the least desirable, to four, the most desirable, with the pairs contrived to favor mixed competitive and cooperative strategies.[153]

1. The Battle of the Sexes (or Marital Cooperation): A Game That Shows Gain to Both Players from the Sacrifice of Equality

Figure 10 is a payoff matrix for a game with differential payoffs. This particular game classically has been known as "Battle of the Sexes."[154] The matrix, unlike the one for the Battle of the Bismarck Sea, shows independent payoffs for each player in each square. For example, if Player 1 chooses row Y and Player 2 chooses column A, the payoff square in the lower left applies, and the payoff is (3, 4): three to Player 1, and four to Player 2.

---

[150] *See* DAVIS, *supra* note 1, at 82.

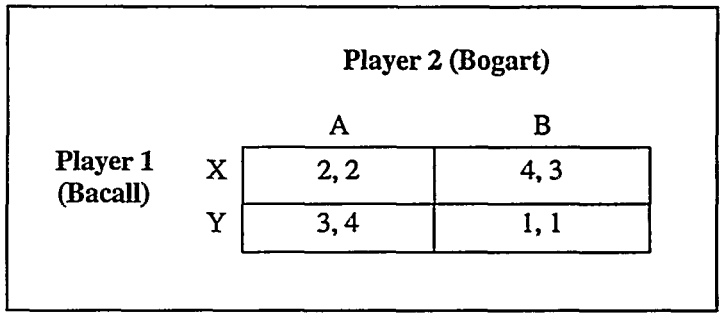
[151] *See id.* at 57–65.

[152] *See* COLMAN, *supra* note 1, at 108–21. It is possible to reframe Battle of the Sexes as "Marital Cooperation," Chicken as "Stand Up for Your Rights," and the Prisoners' Dilemma as "Arms Race." *See infra* Part III.B.1–.3 (explaining the reframings and demonstrating how the different stories alter attitudes of players).

[153] *See* COLMAN, *supra* note 1, at 118–21.

[154] *See* DAVIS, *supra* note 1, at 88, 101–02.

FIGURE 10
BATTLE OF THE SEXES (OR, MARITAL COOPERATION)



Although this game classically has been called Battle of the Sexes, it might as well be called "Marital Cooperation." This is because the payoff matrix has a further difference from that for the Battle of the Bismarck Sea, in that it presents a mixed-motive or partly cooperative situation rather than a zero-sum game.[155] The classical formulation of this game, developed during the 1950s, used a somewhat sexist set of assumptions: Player 1 was the wife, who wanted to attend the ballet, while Player 2 was the husband, whose choice was to go to a boxing match.[156] Here, we shall assume that the wife (Bacall) wants to watch the movie *Gone With the Wind* (her choice "X"), while the husband (Bogart), wants to watch *Casablanca* (his choice "A"). Bacall prefers not to watch *Casablanca*, let us imagine, because in that film Bogart was paired with Ingrid Bergman. The substitution of two romantic war movies makes the game less sexist, but it does not change the analysis.

The couple has two VCR's, and each can retire to his or her own separate entertainment, but this separation produces less-than-optimal utilities of (2, 2), because the husband and wife would prefer to watch together. On the other hand, if the husband sacrifices and watches *Gone With the Wind* with his wife (her choice "X," his choice "B"), the payoffs are (4, 3) (four to the wife and three to the husband), and the husband gets to be a hero, because both are better off. If both watch *Casablanca* (her choice "Y," his choice "A"), the payoffs are reversed (3, 4), and the wife is the one to make the sacrifice. These optimal results occur, that is, unless both sacrifice by making their unfavored choices, leaving the husband alone watching *Gone With the Wind* (which he dislikes) while the

[155] *See supra* note 132 and accompanying text (analyzing payoff structures).
[156] The classical description is attributed to R. Duncan Luce and Howard Raiffa. *See* DAVIS, *supra* note 1, at 88.

wife, also alone, watches *Casablanca* (which she dislikes). This worst arrangement results in a payoff of (1,1).[157]

Battle of the Sexes provides an interesting insight into value conflicts. The game presents a situation in which strict equality must be sacrificed if either player is to obtain maximum satisfaction. Equality is attainable, but ironically is not preferable for either party to an outcome that places one party in a less satisfactory position than the other.[158] The best payoffs for an equal resolution are (2, 2). Higher payoffs result from inequality: (4, 3) or (3, 4).

Thus, the interesting inference from Battle of the Sexes is that, at least in some situations, if one party is willing to suffer inequality in the form of inferior satisfaction, both parties may emerge with higher payoffs, including the one who has sacrificed. To put it another way, sometimes equality conflicts with other important values. Sometimes, human beings may consider themselves better off by not insisting on equality. A statute that mandates equality in such a situation would be counterproductive. Below, after developing other mixed-motive games, we shall consider some concrete examples.

2. Chicken (or Stand Up for Your Rights): A Mixed-Motive Game That Demonstrates the Power of a Reckless Bully To Raise the Cost of Equality

In the game called "Chicken," two cars race toward each other. The first driver to turn is a chicken. If neither has a sufficient instinct for self-preservation, the cars collide head-on, killing both.

Game theorists have devoted significant attention to the game of Chicken.[159] Figure 11 is a payoff matrix for this game. It superficially resembles the matrix for Battle of the Sexes, but there is a perverse difference. Cooperative, socially acceptable behavior does not lead to maximum utility for the individual.[160] After all, this perversity is exactly the point of the real-life game of Chicken. Whimsically, the figure labels the two combatants as "James" (Player 1) and "Dean" (Player 2), after the movie icon who popularized the framing of this nasty entertainment in the movie Rebel Without a Cause.[161]

---

[157] *See id.*

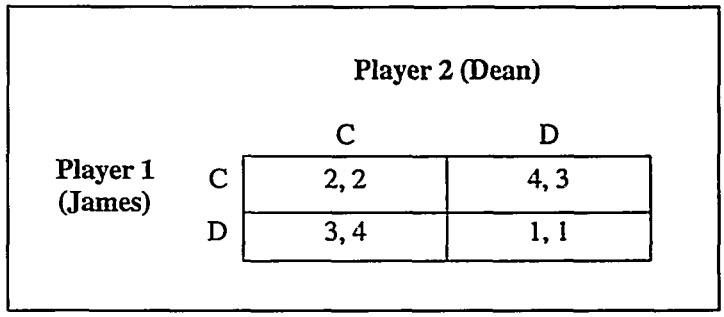
[158] *See* COLMAN, *supra* note 1, at 110–11 (analyzing the trade-off and the strategy of sacrifice).

[159] *See* COLMAN, *supra* note 1, at 111–15.

[160] *Id.*

[161] *Id.*

FIGURE 11
CHICKEN (OR, STAND UP FOR YOUR RIGHTS)



As the figure shows, there is a payoff of (3, 3) if both drivers cooperate by turning away. In this case, both are chicken, but both are alive, and there is less embarrassment since neither has been beaten by the other. But if one driver turns while the other continues straight ahead, the driver who turned is a chicken, and the other driver has proved his manhood or, if one prefers, his prowess as a suicidal misfit. Thus, if James defects (by driving straight ahead) while Dean cooperates (by turning), the payoff is (4, 2). The losing player (Dean) is chicken and alive, but has been beaten and humiliated, while the winning player (James) publicly is confirmed as mean, tough, and macho. This is so even though objective observers might consider loser Dean's behavior more rational, commendable, and socially acceptable than that of winner James; in fact, this incentive toward depravity is the defining characteristic of the game of Chicken. If the choices are reversed, so is the payoff: (2, 4). And finally, the worst possible outcome is for neither player to be chicken, so that the result is a head-on collision, killing both, with a payoff of (1, 1).

The game of Chicken exhibits a number of unusual characteristics. For one, irrationality is power; one might even say that irrationality is rational.[162] Thus, it has been observed that cantankerousness, anger, and recklessness can be successful strategies in negotiation.[163] So can a "lock-in" or commitment strategy, by which a player ostentatiously binds himself or herself to a potentially self-destructive course of action,[164] which encourages the other player to give in to avoid the same disastrous result for both. It follows that the mere appearance of cantankerousness, anger, or irrationality, or the credible communication of a "lock-in" or commit-

---

[162] *Id.*

[163] *See* DAVID CRUMP ET AL., CASES AND MATERIALS ON CIVIL PROCEDURE 767 (3d ed. 1998) (discussing irrationality as a negotiating tactic).

[164] *Id.* at 769.

ment strategy, can be effective. Soviet leader Nikita Khrushchev visited the United Nations during the heyday of the Cold War, where in addition to his belligerent rhetoric, he interrupted a delegate's speech by taking off his shoe and banging it on the table.[165] The Soviet Union was a nuclear power, but Khrushchev's action implied that it was headed by a maniac who did not care about consequences. This strategy works quite well in the game of Chicken, provided only that the exploiter is skillful at communicating it convincingly to another player who is rational.

The commitment strategy, together with the rationality of giving in to the exploiter, has been extensively explored by game theorists. Herman Kahn gives the following description of the technique:

> The "skillful" [Chicken] player may get into the car quite drunk, throwing whiskey bottles out the window to make it clear to everybody just how drunk he is. He wears very dark glasses so that it is obvious that he cannot see much, if anything. As soon as the car reaches high speed, he takes the steering wheel and throws it out the window. If his opponent is watching, he has won. If his opponent is not watching, he has a problem; likewise if both players try this strategy. [166]

Game theorist Andrew M. Colman summarizes the paradoxical advantage that is available to the ostentatiously reckless Chicken player by observing, "people tend to give a wide berth to a lunatic."[167]

Another paradoxical aspect of Chicken is that success does not reinforce cooperative behavior. In fact, the exploiter's strategy is reinforced by his or her success as a bully.[168] Iterated Chicken games, in which the players repeat a sequence of identical games ten or a hundred times in a row, demonstrate the unfortunate advantage that a clear history of reckless domination can bring to the uncooperative party.[169] The domineering player gains confidence that strengthens the strategy, practice makes for greater skill, and a reckless reputation makes opponents more cautious and therefore more likely to defer to the suicidal misfit. It has this effect, that is, until two "skillful" Chicken players face each other, whereupon their deliberate, confident strategies may kill them both.

Perhaps it is unwise to call this game "Chicken" because framing it in this way creates psychological images that may be misleading. The game actually models a variety of real-world circumstances. The Cuban missile crisis, in which President Kennedy risked nuclear war but caused

---

[165] ROY MEDVEDEV, KHRUSHCHEV 154 (Brian Pearce trans., Anchor Press 1983) (1982).

[166] HERMAN KAHN, ON ESCALATION 11 (1965).

[167] COLMAN, *supra* note 1, at 113.

[168] *Id.* at 112–13.

[169] *Id.*

the Soviets to blink and to remove ballistic missiles from Cuba, has been modeled by the same payoff matrix as Chicken.[170] If it is desirable to make the exploiter's role sound less pejorative or even laudable, we can re-frame the game as "Stand Up for Your Rights."[171]

For example, two litigants both claim an otherwise valueless sentimental prize or a matter of principle. They have the option of proceeding to trial, but trial will cost much more than any verdict could be worth. Nevertheless, both players would like to "stand up for their rights," particularly if they believe that belligerent posturing might cause the opponent to throw in the towel. Although this game defeats many of the purposes of our legal system, and although court encouragement of settlements has increased, this game is played daily at our courthouses. Sometimes it results in the worst possible outcome, that of an actual trial, which has the same payoff as a collision between Chicken players. If the stakes are small, the expense of trial will exceed the spoils of victory for the players, making losers of both.

The implications of this game for equality are unpleasant. In some situations, a determined, reckless bully can succeed in obtaining more than his or her fair share. Cooperative players earn only a reputation for cowardice, while exploiters see their own reputations aggrandized and their strategies rewarded by the very act of exploitation. Thus, for example, the Rambo-like lawyer, who pushes in obtaining discovery, hinders his opponent by refusing discovery, uses delay and expense as a weapon, and credibly communicates a lack of concern about the possibility of sanctions by the court accompanied by an indignant threat to file sanctions against his opponent instead, sometimes succeeds at the strategy.[172] This particularly is true in situations in which there is a limited estate, such as in bankruptcy matters. Bankruptcy lawyers are familiar with the phenomenon of the domineering lawyer who succeeds in getting more than an equal share.[173] Likewise, situations where one litigant has a war chest and the other has barely enough to litigate, so that there is a credible threat that the weaker player cannot fund the litigation without great sacrifice, or situations in which the cost of a loss is ostensibly so disproportionate that the stronger player can communicate indifference about the outcome, similarly are vulnerable to the Chicken strategy.

The response of some judges, and even some attorneys, to this phenomenon may be to recite the old proverb that "what goes around, comes

---

[170] *See* COLMAN, *supra* note 1, at 114.

[171] This reframing is original with the author of this Article.

[172] *See infra* note 220 and accompanying text.

[173] Interview with Raymond T. Nimmer, Bankruptcy Attorney, Weil, Gotschal & Manges, and Professor of Law, University of Houston, in Houston, Tex. (Feb. 7, 2000) (during a presentation of this Article by the author, confirming that the strategy "works every time").

around."[174] In other words, these judges or lawyers hope that the law of the jungle ultimately will bring exploiters to heel with a dose of their own medicine. This hope reflects an unfortunate failure to understand the game that is being played when the payoff matrix fits the Chicken model. It particularly fails to take account of iterated Chicken strategies and experiments in which bullies succeed. The fact is, to the extent that what "goes around" does "come around," the implications for equality are disastrous, because players who historically have cooperated are disadvantaged even as they are reinforced in that strategy, while exploiters are reinforced and rewarded by their Rambo-like tactics. Game theory predicts this result until and unless two Rambo-type exploiters meet each other in the same case. But it does not seem obvious to conclude that both Rambo-like lawyers in this situation will lose enough to be deterred, compared to what they have gained over the course of many other encounters that paired them with cooperative players. Furthermore, it seems unlikely that the goes-around-comes-around mantra can give any comfort to cooperative litigants who have been exploited in the past, and will be exploited in the future, by these same Rambos.

Instead, the lesson of the Chicken game is that enforcing rules clear enough to discourage exploitative behaviors, while rejecting the goes-around-comes-around philosophy, for example by making clear when courts will impose sanctions for Rambo tactics, enhances the likelihood of obtaining distributive and procedural equality. Mushy legislation that leads to difficulty (and therefore to relative absence) of enforcement creates a game that enables a skillful Chicken player to obtain more than his or her share.[175] In some instances, clear rules may be more advantageous

---

[174] This maxim could be used by a judge to avoid dealing with a procedural dispute in which one attorney exploits another. Furthermore, there is a natural human tendency to justify nonaction by saying, "what goes around comes around." Telephone Interview with David Hitter, United States District Judge, Southern District of Texas (Mar. 29, 2001). Judge Hittner adds that a judge usually should avoid this tendency. The "what goes around comes around" strategy may be workable in some cases, but not when the Chicken payoff matrix controls.

[175] There may be good reasons for judges to remain reluctant to impose sanctions on would-be Rambo lawyers. Determining "who started it" is always difficult, intrusive, and controversial. Without extensive hearings into the fine details of squabbles, judges may be prone to make mistakes by sanctioning victims. The effort diverts the court, the lawyers, and the parties from the initial issue, which is resolution of the dispute between the parties. It usually imposes costs on lay clients, who often are innocent. And perhaps most importantly, it compromises the independence and impartiality of a judge who hopes to continue presiding effectively over the case.

Moreover, the mushiness of rules that provide for sanctions makes enforcement situations more difficult to identify. "Because both sides have contributed materially to the protraction of this discovery dispute—and therefore of this opinion—Rule 37(a)(4) is best served by letting the expenses rest where they have fallen. Consequently the Motion for an award of fees and other expenses is denied."
Bd. of Educ. v. Admiral Heating, 104 F.R.D. 23, 37 (N.D. Ill. 1984). By leaving losses where they lie, however, the judge may play into the hands of the Rambo-like lawyer who is a skillful Chicken player.

than those that allow for thoughtfulness, deliberation, and discretion to take account of all cases.[176] In a way, the game of Chicken provides a counterpoint to the apparent indeterminacy of equality that seems inferable from the stockholders' problem. In our earlier discussion of that problem, we determined that although there are many potentially defensible models of equality, and few absolute principles for choosing among them, the legislature must cut the Gordian knot, choose among the possible models of equality, and enforce its choice with a hard edge.[177] Game theorists' analyses of the game of Chicken dramatically demonstrate when such a clear definition can be important.

3. The Prisoners' Dilemma (or Arms Race): The Efficiency of Cooperative Equality

The Simple Form of the Prisoners' Dilemma. The Prisoners' Dilemma probably is the most widely analyzed game in all of game theory,[178] largely because it models so many (and such interesting) real-world circumstances in which people act both competitively and cooperatively. The traditional framing of the Prisoners' Dilemma is as follows.[179] Two arrested suspects, apparently both guilty, have been separated in two rooms for interrogation. Let us call the two suspects "Gravano" and "Gotti." The payoffs are ranked ordinally from 1 (the worst outcome) to 4 (the best outcome). Each player can either confess and implicate the other—in game theory terminology, "defect" (D)—or remain silent—"cooperate" (C). Both suspects know that if neither confesses, each will obtain a relatively light but not minimal sentence. The payoff is (3, 3). If, however, Player 1 confesses but Player 2 does not, Player 1 will be sentenced to the lightest possible term of years, while Player 2 will be sentenced to the maximum. The payoff is (4, 1), in favor of the turncoat.[180] If Player 2 confesses but Player 1 does not, the payoffs are reversed: (1, 4). If both confess, both will receive moderately severe sentences, for a payoff of (2, 2). These payoffs are presented in Figure 12.

---

[176] Indeed, game theory suggests that thoughtful open-mindedness can be a fault, at least in some mixed-motive games. *See* DAVIS, *supra* note 1, at 92, 99–100 (asserting that an irrevocable commitment and reduction of options can be strategic).

[177] *See supra* note 110 and accompanying text.

[178] *See* COLMAN, *supra* note 1, at 115.

[179] *See* DAVIS, *supra* note 1, at 108–09 (setting forth equivalent matrix, but framed, instead, in years sentenced); COLMAN, *supra* note 1, at 115 (using ordinal payoffs, ranked from one to four, like the matrix here).

[180] This outcome, incidentally, resembles the sentences imposed on Sammy "the Bull" Gravano and John Gotti in real life, after Gravano testified against Gotti. *See* Paul Leavitt, *5-Year Prison Term for Mafia Turncoat*, USA TODAY, Sept. 27, 1994, at 3A (reporting that Gravano was sentenced to five years in prison, while Gotti received a life sentence without possibility of parole).

FIGURE 12
THE PRISONERS' DILEMMA

		Player II (Gotti)	
		Cooperate	Defect
Player I (Gravano)	Cooperate	**2, 2**	**4, 3**
	Defect	**3, 4**	**1, 1**

Thus, no matter what Player 1 does, it is rational for Player 2 to confess, or defect. If Player 1 cooperates, Player 2 gets a payoff of four by defecting and three by cooperating. If Player 1 defects, Player 2 gets a payoff of two by defecting and one by cooperating. In either case, Player 2 is better off defecting. The dilemma is that if both players defect, they will suffer the lowest possible combined outcome of (2, 2) for a total of four, but if they both cooperate, they will achieve the highest possible combined payoff of (3, 3) for a total of six.[181]

The Prisoners' Dilemma can be reframed as "Arms Race." Figure 13 depicts possible strategies of the United States and the Soviet Union during the Cold War.[182] Cooperation in an arms limitation treaty would have produced maximum combined welfare (3, 3), but mutual fear of an imbalance in which either the United States would have a strategic advantage (4, 1), or the Soviet Union would (1, 4), prompted both Congress and the Kremlin to defect, resulting in an arms race that reflected the lowest possible welfare combination: (2, 2). Thus, the figure that models Arms Race is identical to the payoff matrix for the Prisoners' Dilemma.

[181] *See* DAVIS, *supra* note 1, at 112.
[182] *See id.* at 111 (describing a similar framing).

FIGURE 13
ARMS RACE (A REFRAMING OF THE PRISONERS' DILEMMA)

		Soviet Union: limit missiles, C	Soviet Union: continue building, D
United States: limit missiles, C		3, 3 stop building	1, 4 Soviet ahead
United States: continue building, D		4, 1 U.S. ahead	2, 2 arms race

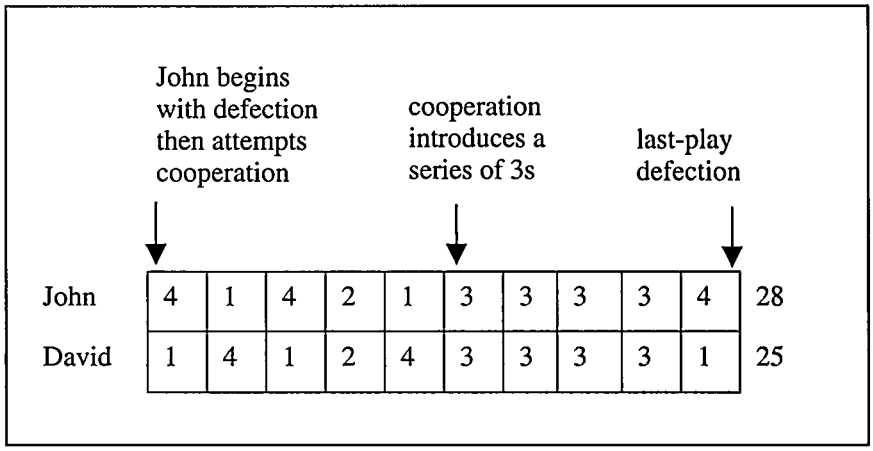
Alternatively, the game can be reframed as "Tragedy of the Commons."[183] Two citizens who must produce all of their consumables inhabit a tiny island together. Cooperation in reducing pollution means a clean environment, for a payoff of (3, 3). But if one of the players defects by polluting at will, while the other cooperates by cleaning up, the defector obtains a reasonably clean environment without expending the effort or bearing the losses entailed in pollution abatement, and this strategy combination produces lopsided welfare effects of (4, 1) or (1, 4). Therefore, both may be motivated to defect, resulting in the worst possible combined payoff: (2, 2). This game demonstrates the value of legislation setting environmental standards.

The implications of the Prisoners' Dilemma for equality are as encouraging as those for Chicken are discouraging. A commitment to cooperative equality has the prospect of producing the highest possible combined welfare for our hypothetical two-person society: (3, 3).[184] This result is dependent, however, upon the removal of ambiguity in the roles of both players. That is to say, both must be willing to rely upon cooperation by the other, or both must be required by a rule or statute to act cooperatively. The trouble is, it also is possible for equality to result from double defection, and this equality produces a distinctly lower level of total welfare. The good news is that if there are repeated encounters governed by the same payoff structure, a strategy of rational mutuality tends to be rewarded (with some important qualifications, as we shall see). In other words, the what-goes-around-comes-around philosophy works, al-

[183] *Cf. id.* at 110–11 (proposing other environmental framings involving use of scarce water and overfarming).
[184] *See id.* at 109.

beit imperfectly, in the prisoners' game.[185] To see why, we next consider the situation in which the same players repeat the prisoners' game ten times, or a hundred times, all in a row. Game theorists refer to this game as the "Iterated Prisoners' Dilemma."[186]

FIGURE 14
AN ACTUAL ITERATED PRISONERS' GAME



The Iterated Prisoners' Dilemma. The Iterated Prisoners' Dilemma changes the calculus of the game, at least psychologically,[187] by making the strategy of cooperation more obviously superior. Figure 14, for example, depicts a ten-iteration sequence played between the author and my then-thirteen-year-old son, John. John approached the game competitively and defected on the first move. I played a strategy known to game theorists as "Tit for Tat,"[188] which we shall consider in greater detail below. After several initial moves, John and I established mutual cooperation, and then we produced a series of maximum combined payoffs: (3,

[185] *Cf. supra* note 174 and accompanying text (discussing the goes-around-comes-around approach).

[186] *See* COLMAN, *supra* note 1, at 136.

[187] One can argue that the effect is psychological, not strategic, by using "backward induction." *Cf. id.* at 104 (explaining backward induction with a game called "Centipede"). Backward induction divides the game into subgames of reciprocal moves, solves the last subgame, and then proceeds backward through each other subgame to the beginning. In the Iterated Prisoners' Dilemma, the last subgame is the last iteration. There is no strategic reason for either player to cooperate in this subgame, and therefore, backward induction begins by concluding that the rational player will defect. Arguably, the rational player also should defect in the next-to-last subgame, knowing that his rational opponent will defect in the last one no matter what the player does. This reasoning then can be applied sequentially to each subgame, back to the first. The paradox is that backward induction produces a logical solution that is in conflict with other rational strategies discussed here.

[188] *See infra* notes 195–196 and accompanying text.

3). Despite strategic defections by both of us, John's payoff was twenty-eight, and mine was twenty-five. This combined payoff of fifty-three significantly exceeded the score of forty that would have resulted from uniform defections, and both players improved on the twenty-point score that each would have obtained.[189] If both parties had played the Tit-for-Tat strategy consistently, however, the combined payoff would have been sixty,[190] the maximum possible. The point is that the structure of the iterated game induces a degree of cooperation not merely for its own sake, but also as a matter of self-interest.

John, whose instructions were to maximize his score, defected on the last iteration, and he was able to increase his score that way at my expense. This strategy, of course, represented rational, game-appropriate behavior by a thirteen-year-old video combat veteran.[191] In most tournaments that are conducted experimentally, however, iterations number in the hundreds, and thus the advantage attributable to this last-move strategy is reduced, though not eliminated.[192] In real life, the number of encounters usually is indefinite and the last-move defection is not a viable strategy because one does not know when an encounter is going to be the last move with that opponent. Theoretically, one can imagine an infinitely iterated prisoners' game in which one never reaches the last move. But the real point is that a strategy of cooperation, based not upon attempting to "win" but upon maximizing payoffs, is what actually produces the greatest individual welfare in the iterated prisoners' game.

The iterated version of this game has led to some fascinating results, none more so than the computer tournaments and evolutionary experiments of game theorist Robert Axelrod.[193] Axelrod invited submissions of programmed strategies to a prisoners' game tournament conducted by computer. Each entry played two hundred iterations against all other programs and against a clone of itself.[194] The winner was a disarmingly simple program entered by game theorist Anatol Rapoport called "TIT FOR TAT," or "TFT."[195]

TFT starts with a cooperative move, and from then on its play simply echoes its opponent's last move, tit for tat.[196] Another program called "RANDOM," because it played cooperate or defect strategies at random, also was entered in Axelrod's tournament. It came in last, as one might guess from its lack of a coherent strategy.[197] Paradoxically, TFT also beat

---

[189] These outcomes result from the (2, 2) payoffs that mutual defection produces.
[190] This outcome results from the (3, 3) payoffs that mutual cooperation produces.
[191] *See also* DAVIS, *supra* note 1, at 147 (explaining the strategy in simpler terms).
[192] *See* COLMAN, *supra* note 1, at 145–49.
[193] *See id.*
[194] *Id.* at 145.
[195] *Id.*
[196] *Id.*
[197] *Id.*

more complex strategies, including those that attempted to obtain advantages by either random or predetermined defections.[198]

Axelrod then conducted a second computer tournament, which attracted four times as many entries, and TFT again won.[199] Notice that TFT can never beat its opponent in a one-on-one contest, because a defecting program stays ahead of it and a TFT clone always produces the same score as the TFT program it is playing against.[200]

After his tournaments, Axelrod took the idea into the next dimension. He simulated natural selection or Darwinian survival of each of the sixty-three programs by adjusting the number of offspring in each successive round, or "generation," based on a strategy's success in the previous round.[201] After one thousand generations of play, weak programs became extinct, and so did some "predatory" programs that had survived by exploiting dwindling programs lower in the food chain.[202] Interestingly, in this game designed to simulate Darwinian natural selection, TFT won again, just as it had in Axelrod's tournaments. Its numbers increased in each round and grew to 14% of the population.[203]

Since that time, other experiments have demonstrated that TFT can be beaten, although that outcome depends on the rest of the population. For example, a program called "PAVLOV," which plays a version of a strategy known as "win-stay, lose-change," outdoes TFT in some populations.[204] But PAVLOV loses in an environment of "all-defect" programs, whereas TFT can survive in such a situation.[205] These experiments are a reminder that "evolution" and "survival of the fittest" are misleading terms. The success of TFT, which literally cannot beat any other program in a head-to-head contest, indicates that the key is survival in the environment, or natural selection, and not "survival of the (abstractly) fittest" or of the fastest, strongest, or most competitive.[206]

To what should one attribute TFT's success? Axelrod himself offered the following analysis:

> What accounts for TIT FOR TAT's robust success is its combination of being nice, retaliatory, forgiving, and clear. Its niceness prevents it from getting into unnecessary trouble. Its retaliation discourages the other side from persisting whenever

---

[198] For example, some programs defected twice upon an opponent's defection, some defected in the last subgame, and some contained sophisticated efforts to base defections on the opponent's entire history. *See* DAVIS, *supra* note 1, at 147.

[199] COLMAN, *supra* note 1, at 145–47.

[200] *See id.*

[201] *Id.* at 148.

[202] *Id.*

[203] *Id.*

[204] *Id.* at 149.

[205] *Id.*

[206] *See id.* at 148.

> defection is tried. Its forgiveness helps restore mutual cooperation. And its clarity makes it intelligible to the other player, thereby eliciting long-term cooperation.[207]

Axelrod thus concluded that a "nice" program, one that never initiates defection, generally beats a "nasty" one.[208] In both tournaments, nice programs clustered in the top rankings, nasty ones in the bottom.[209] For example, a program called "JOSS" resembled TFT, but with ten percent random defections, followed by reversion to cooperative moves.[210] This touch of nastiness did not improve JOSS's score because in playing against TFT, for example, JOSS's first double-cross would trigger a next-move defection by TFT at the same time that JOSS tried to establish cooperation (DC), leading to a series of alternating CD, DC, CD moves. Then, the next time JOSS mixed in a double-cross, it would occur when TFT was itself in defect mode, producing a DD pair. After that, all moves would be DD.[211]

Axelrod concluded that, in real life, double-crossing people act in this way because they have not realized the consequence of a double-cross to their own self-interest.[212] Double-crossers, in other words, arguably may not be dishonest so much as they are foolish. They fail to understand how the world works, and they are ignorant of the proverb that "what goes around comes around," or that their double-crosses will provoke retaliation.[213] A corollary conclusion is that double-crosses are disproportionately costly.

On the other hand, it was not enough for TFT to be nice. TFT also needed to be "retaliatory" in order to succeed as it did. Another program, called "TIT FOR TWO TATS" ("TFTT"), was less successful because nasty opponents gained too much ground by multiple defections.[214] In contrast, whenever the opposing player defects, TFT does so also, on the very next move. If the opponent is rational, it has a powerful incentive to discontinue defection, in which case, TFT will be "forgiving," or forthcoming with cooperation on the next move. TFT's forgiveness enables it to outperform a program called "FRIEDMAN," which assaults defecting opponents with a "mafia" strategy of retaliations repeated to the point of overkill.[215] Finally, TFT is "clear," in the sense that its simplicity makes

---

[207] ROBERT AXELROD, THE EVOLUTION OF COOPERATION 54 (1984).
[208] *Id.* at 33.
[209] *See* COLMAN, *supra* note 1, at 146–47.
[210] *Id.* at 145–46.
[211] *Id.*
[212] *See* AXELROD, *supra* note 207, at 38.
[213] *See supra* notes 174, 185 and accompanying text (explaining why the proverb does not necessarily hold true for Chicken, though it does for the Prisoners' Dilemma).
[214] COLMAN, *supra* note 1, at 147.
[215] *Id.*

its next move easily predictable by the opponent. Its transparency is a virtue, because it enhances cooperation.[216]

What are the implications of the iterated prisoners' game for legislation that affects equality? Even more than the simple prisoners' game, the iterated version leads to some heartening conclusions. A strategy of defection as a means of obtaining a greater-than-equal share of the pie ultimately is self-defeating where there is a Prisoners' Dilemma payoff matrix.[217] Although TFT cannot take advantage of any opponent and indeed is entirely incapable of ever winning any game,[218] its strategy of never attempting to obtain more than its own fair share (but of insisting that it not be repeatedly exploited) wins in the long run, at least in the artificial conditions of the iterated prisoners' game. Axelrod summed it up by advocating the maxim, "don't be envious."[219]

As a more specific application of the game to problems of equality, consider the behavior of the Rambo lawyer. In a book with the intriguing title, *Why Lawyers Lie and Engage in Other Repugnant Behavior*, attorney Mark Perlmutter describes Rambo-style litigation tactics ranging from verbal abuse to theft to perjury.[220] But Perlmutter suggests that most lawyers would prefer to operate under a procedural system free of these (and lesser) abuses.[221] They would prefer, in other words, a cooperative strategy that provides for a level playing field. Perlmutter's own solution to this dilemma is to send each opponent, at the beginning of each lawsuit, a proposed letter of agreement offering cooperative strategies for pretrial development of the case.[222] Some lawyers sign Perlmutter's proposal, some implicitly accept it, and some vehemently reject it.[223] If an agreeing opponent deviates (defects), Perlmutter remonstrates: "Do you think that's consistent with our agreement . . . ?"[224] Such arrangements are probably what Congress and the Supreme Court intended when they promulgated Federal Rule of Civil Procedure 26(f), requiring attorneys at the beginning of litigation to meet, confer, and "attempt[ ] in good faith to agree on [a] proposed discovery plan."[225] The intent of Rule 26(f) is to set up a payoff structure that encourages cooperation, resembling the iterated prisoners' game.

---

[216] Thus, the Prisoners' Dilemma (and some other mixed-motive games) differ from games in which inscrutability is strategic. *See, e.g., id.* at 64 (discussing a game in which concealment of intentions is "vitally important").

[217] *See supra* notes 208–213 and accompanying text.

[218] *See supra* note 200 and accompanying text.

[219] *See* AXELROD, *supra* note 207, at 110–13.

[220] MARK PERLMUTTER, WHY LAWYERS LIE AND ENGAGE IN OTHER REPUGNANT BEHAVIOR pt. 1 (1997).

[221] *Id.* at 31–32.

[222] *Id.* at 132–33.

[223] *Id.* at 133.

[224] *Id.*

[225] FED. R. CIV. P. 26(f).

How well does this real-life "game" of litigation discovery correspond to the prisoners' dilemma? Reasonably closely. A lawyer who defects may gain a temporary advantage, but the behavior usually results in defection by the opponent, either in the form of reciprocal corresponding behavior or in a hearing before the judge. Thus, although Rambo imposes costs upon his opponent, he also imposes them upon himself. Perhaps Perlmutter's description of his strategy is incomplete, because in the iterated prisoners' game, TFT is retaliatory, and without this feature, it would not lead to cooperative equality.[226] The lawyer faced with a Rambo opponent must succeed in drawing a line, somewhere. On the other hand, Perlmutter's description of his proposed agreement demonstrates why niceness works. If both lawyers follow an equal strategy of cooperation, both provide a better service for more clients at a lower cost. Rule 26(f), again, is an effort to produce this result.

4. Leader: The Fourth Mixed-Motive Two-Person Game

The typical framing of Leader is as follows. Two cars are attempting to enter a freeway ramp. Either one can go first. The driver who goes first obtains a payoff of four, and the driver who follows immediately behind him gets nearly the same payoff, three. Thus, the payoffs for these circumstances are (4, 3) and (3, 4). Both the leader and the follower get on the freeway quickly, with the follower taking only slightly less time. But if both drivers wait for each other and neither enters the ramp, both will be delayed and the payoff will be (2, 2). A payoff of (1, 1) results if both enter at the same time, because they will block each other and barely avoid a collision.[227]

The game of Leader might be thought of as justifying the maxim, "*de minimis non curat lex*."[228] Compared to the disadvantages of both equal payoffs, which are (2, 2) (stalemate) or (1, 1) (blockage), the small difference in payoff attributable to the first or second entry onto the freeway is trivial. In the game, the payoff differences between (4, 3) and (3, 4) perhaps are trivial other than to game theorists.[229]

But perhaps there is a deeper meaning to Leader. In situations in which organized, orderly outcomes produce only slight differences in equality, and in which disorganization leads to significant welfare losses, the best legislative strategy may be to adopt deliberately arbitrary rules. If two automobiles reach an intersection at the same time, it matters little whether the legislature assigns the right of way to the car on the left or

---

[226] *See supra* note 189 and accompanying text.

[227] COLMAN, *supra* note 1, at 108–09.

[228] *See* BLACK'S LAW DICTIONARY 443 (7th ed. 1999) ("The law does not concern itself with trifles.").

[229] *See* COLMAN, *supra* note 1, at 109–10 (suggesting "informal" solutions such as the maxim, "first come, first served").

the one on the right, so long as it assigns it definitively. The repetition of this encounter throughout the lives of both drivers will produce a rough approximation of equality, just as, perhaps, a good strategy for iterated games of Leader might produce this cooperative outcome.[230] A similar strategy is reflected in legislation or court decisions adopting the "birthday rule," which assigns health insurance costs between two overlapping insurers to the one whose insured has the earliest birthday in the year.[231] The value of an individual claim is small in the overall scheme of things, the inequality of arbitrarily imposing it on one or the other is likely to even out over the long-term, and the parties rationally should avoid the disaster of litigation with costs that will far exceed that of the claim. A rule calling for a coin flip would be just as effective, and in this scenario, paradoxically, an arbitrary resolution is an equal one. But the real point is that a clear law is preferable to precise equality, even if it is founded on a random factor.

5. Putting It Together: The Legislative Implications of Battle of the Sexes, Chicken, the Prisoners' Dilemma, and Leader

Whether a legislature should intervene to secure equality depends, in part, upon whether the pre-existing payoff matrix fits more closely with that of Battle of the Sexes, Chicken, the Prisoners' Dilemma, or Leader. Of course, the strategy that influences the distribution may not meet any of these conditions. It may, for example, reflect more closely the payoff structure of a zero-sum game.[232] However, these two-person mixed-motive games model many human situations. The modeling, though usually imperfect, can provide useful insights.

The discussion that follows will apply the theories of Battle of the Sexes, Chicken, the Prisoners' Dilemma, and Leader to problems of equality in the context of family law, marriage, and divorce. The usefulness of these theories, however, is not confined to this area of the law. Some of the analysis of these games has been used to model antitrust law applications, for example.[233] But marriage and divorce legislation involves sensitive questions of equality as well as conflicts between equality and other values, and it invites rich development of all four of the mixed-motive games.

---

[230] *Cf. supra* Part IV.B.3.b (describing the Iterated Prisoners' Dilemma).

[231] *See, e.g.*, 215 ILL. COMP. STAT. 5/367(11)(b) (1998); Boon Chapman, Inc. v. Tomball Hosp. Auth., 941 S.W.2d 383, 384 (Tex. App. 1997).

[232] For example, the Battle of the Bismarck Sea had a zero-sum payoff structure. *See supra* Part IV.A.

[233] *See* PORTER, *supra* note 5, at 88–89 (discussing use of Prisoners' Dilemma to model the economic behavior of firms in oligopoly).

First, we apply the Battle of the Sexes, which we also have described as the game of "Marital Cooperation."[234] If the payoff structure between two people fits this game, the legislature should avoid intervening on equitable grounds. Overall welfare is inconsistent with equality. In this case, efforts to ensure rigorous equality will result in overall losses serious enough that governmental action is likely to be disadvantageous. Thus, for example, it would be inadvisable to enforce custody or visitation schedules during marriage, even though divorce decrees must resolve a variety of equalities of this kind, including equal treatment of parents and equilibration of finances, because the game changes upon divorce. During the marriage, welfare maximizing requires sacrifices of equality from moment to moment. Therefore, the costs of enforcing equality while the marriage is still intact, or even of regulating each parent's participation for the best interest of children, would be too high. Occasionally, there are proposals for legislation regulating the relationship of husband and wife during marriage.[235] Short of disasters such as fraud or violence, game theory suggests that these political temptations should be resisted.

The legislature should, however, be on the lookout for payoff structures that resemble the game of Chicken. Such structures do occur in divorce situations. The avenger or the bully who hates his or her former partner to such an extent that psychological victories count more than overall welfare, more than self-preservation, and more than the welfare of the children, can easily present the other spouse with a situation in which the only rational solution is capitulation.[236] Equality in these circumstances depends upon the legislature's ability to recognize this be-

---

[234] *See supra* note 155 and accompanying text.

[235] It occasionally has been suggested, for example, that the law should encourage the formulation during marriage of joint decisions reflecting the consent of both spouses when either disposes of marital property within his or her control. Interview with J. Thomas Oldham, Professor of Law, University of Houston, in Houston, Tex. (Apr. 12, 2000) (explaining views of proponents other than himself). If enacted into law, these suggestions might well decrease overall welfare in marriages in which either spouse makes many transactions, particularly if the parties have good reason to, or prefer to, trust unified management by a single spouse. *See* J. Thomas Oldham, *Management of the Community Estate During an Intact Marriage*, 56 LAW & CONTEMP. PROBS. 99, 106–07 (1993) (discussing different state-law regimes, including laws forcing joint decisions or allowing separate management; demonstrating the "substantial burden" on both commerce and each spouse imposed by laws uniformly requiring joint management); Richard W. Bartke, *Marital Sharing—Why Not Do It by Contract?*, 67 GEO. L.J. 1131, 1172 (1979) (agreeing with states that "provide for exclusive management by the active spouse" in the context of "the pursuit of a business or profession by only one spouse").

At one time, coverture doctrines were designed to protect the "equality" of married women by, in effect, assuming their weakness and need for protection from laws forcing legal decisions to be made through the agency of her husband. *See* BLACK'S LAW DICTIONARY 373 (7th ed. 1999) (defining coverture); Margaret Valentine Turano, *Jane Austen, Charlotte Bronte and the Marital Property Law*, 21 HARV. WOMEN'S L.J. 179 (1998).

[236] *Cf. supra* note 173 and accompanying text (concluding that certain Chicken strategies may be successful, particularly in situations in which there is a limited estate).

havior and to remove the skillful Chicken player's perverse advantage, since courts probably do poorly at this task. It is an inherently difficult assignment, but the gains in equality we would experience from effectively addressing such a problem would be significant. In other words, just as game theory suggests that the law should hesitate to ask who is exploiting whom during the marriage, it indicates that we should be quick to recognize and discourage a bully during divorce. A statutory penalty, in the form of a sanction for unfounded pleadings, seems particularly appropriate in the divorce setting.

The Prisoners' Dilemma is the most studied game in all of the literature of game theory, not only because its strategies pervasively are confounding, but also because it fits a wide variety of human contacts.[237] In the long run, for most people, living through divorce is like the Iterated Prisoners' Dilemma. Cooperative equality is strategic and is encouraged by the payoff matrix itself, and it need not be micromanaged by legislation. For some people whose battles to carve up the pie (and the children) continue forever, the result is long-term low payoffs. But the irrationality of this course of conduct, like continued defection in the Prisoners' Dilemma, actually proves too costly for most people, and they ultimately discover the rationality of cooperation. Therefore, except in extreme, very long-term cases, game theory seems to suggest that continuing intervention to secure equality after divorce should be limited. This is so, that is, unless the parties have converted the game, de facto, into a contest resembling Chicken.

Finally, there are circumstances in which the Leader matrix applies, and in which arbitrary rules, analogous to those that tell us to drive either on the left of the road or the right,[238] become necessary. In the so-called standard divorce decree, for example, a non-custodial parent obtains visitation on Wednesday evenings and first and third weekends. Such a regime may be imposed by legislation, custom, or agreement, but the key point is that it must be clear, just as a law telling us to drive on the left or the right must be clear. The arbitrariness of this visitation arrangement may sacrifice equality to a degree. For example, the noncustodial parent may prefer the second weekend in a given month rather than the third, desiring to take his son to a fly fishing event or her daughter to a soccer tournament. The custodial parent may gain at the noncustodial's (and children's) expense in this particular instance. Definitive resolution of these issues over the life of the custodial arrangement, however, may be more important than instance-by-instance equality, and differences tend to counterbalance each other in the long run.

---

[237] *See supra* notes 178–183 and accompanying text.

[238] *See, e.g.*, N.Y. VEH. & TRAF. LAW § 1120 (McKinney 1996) (requiring vehicles to be driven on the right side of the road).

Again, since real life is not as simple or quantifiable as these games, the insights gained from this analysis need to be taken with a grain of salt. Nevertheless, the insights are real. Statutory proposals for intervening in the affairs of married people as a means of creating equality sometimes do surface,[239] and Battle of the Sexes shows why these proposals should be greeted with skepticism. On the other hand, the law arguably does poorly at recognizing bullies in divorce cases, and the Chicken matrix may show why evenhandedness crucially depends on improving our legislative treatment of this issue. Game theory is not ethics, and it cannot provide infallible prescriptions, but if it is consulted as a means of analysis or of asking these kinds of questions, perhaps it can be useful.

V. STRATEGIES OF MAXIMIZATION AND RAWLS'S THEORY OF DISTRIBUTIVE JUSTICE

In this Part, we shall develop the strategies that game theorists refer to as "maximax," "maximin," and "minimax." These strategies represent different ways of choosing which variable or risk to maximize or minimize.[240] John Rawls's *A Theory of Justice*,[241] one of the landmark works on distributive equality,[242] focuses heavily upon precisely this kind of strategic choice. The linchpin of Rawls's theory is adoption of the maximin strategy for all legislative choices in a society.[243] Here, we shall first sketch Rawls's theory, and then examine the maximin theory and its counterparts.

A. Rawls's Two Principles of Distributive Justice: The Equality Principle and the Difference Principle

Rawls presented his theory as an improvement upon social contract approaches and as an alternative to utilitarianism.[244] His major theme is deontological[245] and sounds Kantian at first blush:[246] "Justice is the first virtue of social institutions .... Each person possesses an inviolability

---

[239] *See supra* note 235.

[240] *See infra* Part V.B.

[241] RAWLS, *supra* note 8.

[242] *Cf.* DAVID W. BARNES & LYNN A. STOUT, THE ECONOMICS OF CONSTITUTIONAL LAW AND PUBLIC CHOICE 1–34 (1992) (including excerpts from *A Theory of Justice* as counterparts to the works of Thomas Hobbes and Jeremy Bentham).

[243] *See infra* note 268 and accompanying text.

[244] RAWLS, *supra* note 8, at 10, 19–24.

[245] Deontology is defined as "ethics dealing esp[ecially] with duty, moral obligation, and right action." RANDOM HOUSE WEBSTER'S COLLEGIATE DICTIONARY 362 (1996). A deontological approach, which emphasizes justice and the right-wrong distinction, is to be contrasted with a "teleological," or purposive, consequentialist, or result-justified philosophy such as utilitarianism.

[246] *See* IMMANUEL KANT, FOUNDATIONS OF THE METAPHYSICS OF MORALS 44 (Lewis White Beck trans., Robert P. Wolff ed., 2d ed. 1995) (1785).

founded on justice that even the welfare of society as a whole cannot override."[247] But Rawls's approach is not that of Kant's universal imitation or categorical imperatives.[248] It uses newer methods and reaches different results.

Furthermore, Rawls's theory is sophisticated, with important defined places for market economics and civil disobedience, among many other non-Kantian arrangements.[249] But the thrust of his conclusion is that justice is "fairness" to all.[250] A major criticism of Rawls's theory is that it centerpieces what arguably may be a stultifying kind of equality—fairness above all else—which some people think suffocates other values such as freedom and autonomy.[251]

1. The Derivation of Rawls's Theory: The Device of the Original Position, with Perfect Knowledge but Behind a Veil of Ignorance

Rawls replaces the social contract, Hobbes's great device,[252] with a similar but different tool, which he calls "the original position."[253] This device is a hypothetical construct, just as the social contract was, to focus the thinking that derives Rawls's theory. Rawls's original position is the condition in which all of the members of society would find themselves if they were transformed out of their current roles and into roles as participants in a kind of legislative conference for a new social order, to which they were to be returned in yet-unknown roles.[254]

Rawls posits that all planners in the original position have "perfect knowledge"[255] of the varieties of status into which they might return. But a key feature of the original position is that all the planning is assumed to be done behind a "veil of ignorance."[256] None of the planners knows what the eventual role of any of them will be in society—whether male, fe-

---

[247] RAWLS, *supra* note 8, at 3.

[248] "There is, therefore, only one categorical imperative. It is: Act only according to that maxim by which you can at the same time will that it should become a universal law." KANT, *supra* note 246, at 38. The application of this summary principle produced for Kant several other universal rules or categorical imperatives, including the anti-objectification principle, which prohibited any person from using another as an object solely to achieve his own ends. *See id.*

[249] RAWLS, *supra* note 8, at 239–42, 319–46.

[250] *See id.* at 10–19.

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

[252] *See* THOMAS HOBBES, LEVIATHAN 186–87 (C.B. Macpherson ed., Penguin 1968) (1651) (theorizing that because, in a state of nature, "the life of man [is] solitary, poor[ ], nasty, brutish, and short," the state exists as a common power to protect subjects by mutual consent, made up by the covenant or contract of every person to accept it as sovereign).

[253] RAWLS, *supra* note 8, at 15–19.

[254] *See id.* at 15–19, 102–60 (expanding on the concept).

[255] Perfect knowledge is the game theory term. Rawls puts it slightly differently: the planners "are presumed to know whatever general facts affect the choice of the principles of justice." *Id.* at 119.

[256] *Id.* at 136–42.

male, butcher, baker, candlestick-maker, disabled, gay, member of a majority or minority race, merchant, or thief. Yet all of the planners know and fully understand these roles.[257] Thus, the purpose of perfect knowledge is to allow informed lawmaking, but the veil of ignorance assures that lawmakers will be concerned for the whole society, not just for themselves. The device of the original position is only that—a device—just as the metaphor of the social contract and Kant's universal imitation were devices from which to derive the real theory.

2. The Equality Principle, the Difference Principle, and Rawls's Choice of the Maximin Strategy

Rawls concludes that lawmakers in the original position, with perfect knowledge of the society but with their own roles hidden behind a veil of ignorance, "would choose two rather different principles: the first requires equality in the assignment of basic rights and duties, while the second holds that social and economic inequalities, for example inequalities of wealth and authority, are just only if they result in compensating benefits for everyone, and in particular for the least advantaged members of society."[258]

These two maxims differ starkly from utilitarianism. As Rawls says, they "rule out" the argument that "hardships of some are offset by a greater good in the aggregate."[259] Here is his basic statement of the two principles:

> [1] First [Rawls's "equality principle"]: each person is to have an equal right to the most extensive basic liberty compatible with a similar liberty for others.
>
> [2] Second [Rawls's "difference principle"]: social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone's advantage, and (b) attached to positions and offices open to all.[260]

The genius of this formulation, arguably, is that although its roots are deontological, it also accommodates versions of institutions more

---

[257] *See id.* at 137.

[258] *Id.* at 14–15.

[259] *Id.* at 15. In contrast, utilitarianism seeks to maximize aggregate happiness. "[T]he foundation of morals [is] 'utility' or the 'greatest happiness principle[,]' [which] holds that actions are right in proportion as they tend to promote happiness; wrong as they tend to produce the reverse of happiness." JOHN STUART MILL, UTILITARIANISM 7 (George Sher ed., Hackett 1979) (1861).

This aggregate approach to happiness, unfortunately, might result in rationalizing slavery if it brought "happiness" to others deemed greater than the unhappiness of the slaves. Kant's anti-objectification principle reaches the opposite result.

[260] RAWLS, *supra* note 8, at 60.

closely associated with utilitarianism, such as political democracy and market economics.

Rawls's ambitious aim is to rank these two principles for all cases, even when they conflict with other principles or with each other. His formulation is that the equality principle is the first priority, and it must in all cases be satisfied. The difference principle is subordinate, and all other principles, in turn, are subordinate to these.[261] This priority avoids "intuitionism," in which the ethical decision-maker is left with no basis to decide among conflicting principles other than intuition.[262]

Rawls's difference principle is the more complex. The equality principle applies only to basic liberties, says Rawls, such as the right to vote and freedom of speech.[263] The difference principle holds that outside of the absolute equality of basic political liberty, there are realms of wealth and social status for which the actual distribution can (and will) be unequal, even though in the original position all may have contemplated an equal chance.[264] There must be leaders of certain institutions, for example, and they may enjoy higher status and rewards than others. Markets mean a different distribution of wealth than precise individual equality. An inventor or developer who through creation or hard work provides benefits to others will wind up with greater-than-average wealth.

For Rawls, these inequalities must always depend on two rigorous conditions that must be requirements of all legislation. First, the opportunity to enjoy these unequal benefits must be open to all.[265] And second, no inequality is tolerable unless it makes everyone better off, not just the benefited individual, or some people, or even the majority.[266] In fact, the true criterion of an acceptable legislated inequality, says Rawls, is that it must benefit "the least advantaged members of society."[267] In other words, there can be no bonuses for high-performing executives unless they also make homeless persons better off. Rawls refers to this as the "maximin" principle: "We are to adopt the alternative the worst outcome of which is superior to the worst outcomes of the others"[268]—worst outcome meaning the impact on the least advantaged person in the entire society.

Maximin is a term of strategy. Game theorists are familiar with it. We shall develop the maximin principle in game theory terms next, together with its counterparts, the maximax and minimax strategies.

---

[261] *Id.* at 61–62.

[262] *Id.* at 34–45. But Rawls does not avoid the problem completely. Two different persons' claimed rights might come into conflict, such as if one person's claimed right of association conflicted with another's claim to freedom of speech, and Rawls's first principle does not resolve this inconsistency.

[263] *Id.* at 61.

[264] *See id.* at 95–100.

[265] *Id.* at 65.

[266] *Id.*

[267] *Id.* at 15.

[268] *Id.* at 152–53.

B. Maximax, Maximin, and Minimax Strategies

There are three abstract forms of strategy: "maximax," which might be regarded as the most optimistic; "maximin," Rawls's choice, which arguably is pessimistic; and "minimax," which might be viewed as a conservative in-between strategy. As it happens, it is the third of these, the minimax strategy, that is rational for both players in the Battle of the Bismarck Sea.[269] Incidentally, the names maximax, maximin, and minimax may be unfortunate because they sound confusingly similar, but this is the terminology adopted by game theorists.

1. The Maximax Strategy (Considering Only the Best Possible Outcome): The Bet-the-Company-on-the-Lottery Approach

If General Kenney had been a romantic in conducting his part of the Battle of the Bismarck Sea, perhaps he would have chosen the "maximax" strategy. This approach consists simply of "maxi"mizing the "max"imum possible outcome by choosing the best possible outcome of all rows.[270] In this event, General Kenney simply would have chosen the column that contained the highest possible payoff. He would have searched south, seeking a payoff of three days of bombing. In doing so, he would ignore the other box in the column, which forecasts the worst possible result, only one day's bombing. Maximax might be called the "bet the company on the lottery" strategy.

In zero-sum games or in any real event with similar parameters, maximax usually is a poor strategy.[271] Had he searched south, Kenney's romantic thoughtlessness would not have given him the maximum return, instead, it would have given him the worst possible result, only one day of bombing (since Imamura rationally could only sail north). The overly optimistic maximax strategy does, however, coincide with the romantic behavior of some people who play the lottery when its payoff is at a maximum, even though the number of other players may reduce mathematical expectation below levels where they might be with lesser jackpots. Maximax, then, sacrifices strategy to wishful thinking, and laws creating state lotteries[272] exploit this logical weakness.

---

[269] *See supra* Part IV.A. In fact, the "minimax theorem" is that if both players are rational, the value of any finite zero sum game is fixed and is determined by each player's minimax strategy. *See* DAVIS, *supra* note 1, at 38–39.

[270] *See* COLMAN, *supra* note 1, at 26.

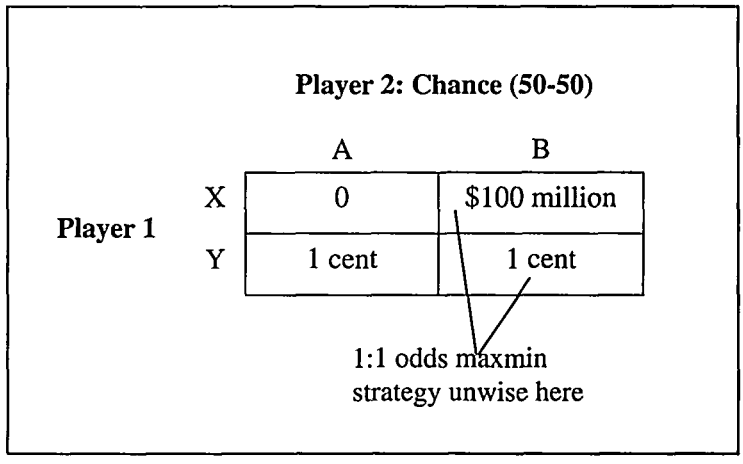
[271] As Colman puts it, although maximax "possesses a certain innocent charm" and is "widely adopted in certain classes of situations," it is "transparently silly." *Id.*

[272] *See, e.g.*, TEX. GOV'T CODE § 466.001–.410 (Vernon 1998 & Supp. 2000).

2. The Maximin Strategy (Choosing the Best of the Worst Possible Outcomes): Avoiding Risk at All Cost

The opposite strategy is to search out the worst possible outcome of each possible move, and then to make the move that would select the best of all the worst possibilities. This is called the "maximin" strategy, because it "maxi"mizes the "min"imum payoffs.[273] A few commentators have suggested this principle,[274] including Rawls.[275] But just as maximax may be too optimistic, maximin may be excessively pessimistic; and in some kinds of games, its disadvantages are a matter of universal proof.[276]

FIGURE 15
THE FALLACY OF A UNIFORM MAXIMIN STRATEGY



Consider the payoff matrix in Figure 15. Assume that "Player 2" actually is chance, with one to one odds of selecting either Column A or Column B. The maximin strategy would tell Player 1, irrationally, always to choose Row Y so as to obtain a sure one-cent payoff, in preference to risking a zero payoff by choosing Row X, even though Row X also contains a one hundred million dollar payoff that has a fifty percent likelihood of being realized![277] In fact, the maximin strategy will always select Row Y, even if the payoff is only a fraction of a cent, no matter how high

---

[273] COLMAN, *supra* note 1, at 26–27.
[274] *See id.*
[275] *See supra* note 267–268 and accompanying text.
[276] *See supra* note 248.
[277] *Cf.* COLMAN, *supra* note 1, at 27 (setting out a different but equivalent game).

the alternative to zero gets in Row X, and without regard to how close the jackpot payoff approaches certainty.

Still, it is impossible to dismiss maximin categorically in all circumstances. Sometimes, when it is important at all costs to avoid the worst-of-the-worst outcome, the maximin strategy may make sense. And let us remember: Rawls specifies the maximin strategy for his ideal society, the society with the best possible distributive justice. Is this choice appropriate? We shall consider that question in a moment, below.

3. The Minimax Strategy (Minimizing Opportunity Loss from the Opponent's Strategy): Taking Account of the Rational Adversary

Game theorists consider the third strategy, the "minimax" strategy, generally to be superior in zero-sum games.[278] The key word here is "generally;" later, we shall see that the minimax strategy has its limits.[279] To see why the minimax strategy is preferred, it is useful to transform the payoff matrix into a "loss matrix" (sometimes called a "regret matrix").[280] Instead of payoffs, each square in the matrix contains the "opportunity loss," or payoff reduction from some maximum, that is represented by that combination of strategies. The player then "mini"mizes the "max"imum loss in the available choices to follow the "minimax" strategy. This strategy is dictated by the strategy of a rational opponent.[281]

FIGURE 16
LOSS MATRIX FOR THE BATTLE OF THE BISMARCK SEA

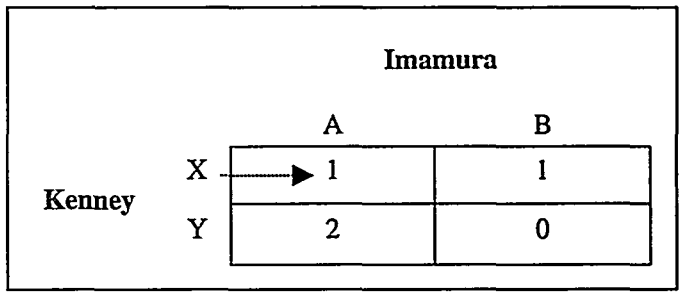


Figure 16 is a loss matrix for Kenney in the game represented by the Battle of the Bismarck Sea. Because the maximum payoff is three, we call the corresponding opportunity loss zero, for "zero opportunity loss." If Kenney gets three days of bombing, this is the best he can possibly do,

[278] In fact, with two equal and rational players, the minimax theorem proves this preference. *See supra* note 269.

[279] *See infra* note 286 and accompanying text.

[280] *See* COLMAN, *supra* note 1, at 28–29.

[281] *See supra* note 269 (explaining the minimax theorem).

and there is zero opportunity loss. We also replace each "two" with a "one," because two days, for Kenney, is one day less than the best possibility of three, and so the opportunity loss is one. And we replace the "one" with a "two" because it is two less than three. The result, then, is the matrix in the figure, a loss matrix.[282]

If he follows the minimax strategy, Kenney makes the choice that will *minimize* his *maximum opportunity loss*. The maximum loss, here, is two, which is contained in Row Y; therefore, the minimax strategy for Kenney is to choose Row X, for which the maximum loss will be only one. This conclusion fits our earlier analysis of the Battle of the Bismarck Sea.[283]

Minimax is preferable for rational two-person zero-sum games with perfect knowledge. The wisdom of the minimax choice for Kenney is confirmed if we consider the independent strategy of Imamura. Column A dominates for him. Imamura therefore must choose A, or north (Column B, south, is inadmissible), and if Kenney had chosen Y, or south, he would suffer his worst possible outcome, an opportunity loss of two. This is true at least for rational zero sum games where each player knows the payoff matrix.[284]

Now, let us introduce another element into the picture to show a fundamental problem with the minimax strategy. Suppose Imamura flips a coin, which produces perfectly random results, and resolves to sail north if the coin comes up heads and south if the coin comes up tails. Thus, it is the coin that really is in control, nature (or chance) is the real second player, and the odds are equal that Imamura will sail either north or south. If Kenney knows about the coin flip, there no longer is any reason for him to consider Row X dominant. His mathematical expectancy from the payoffs in Row X is $(1/2 \times 1) + (1/2 \times 1) = 1$, and from the payoffs in Row Y it is $(1/2 \times 2) + (1/2 \times 0) = 1$. The two expectancies are the same, and therefore the decision-maker is indifferent.[285] The minimax strategy simply does not matter here because we have replaced a strategic opponent with chance.

Furthermore, if we assume a still different game, one in which the coin is "loaded" toward tails, thus favoring a southern trip for Imamura, the minimax strategy becomes decisively inferior. If, say, the odds are three to one that the coin will tell Imamura to sail south, then Row X (north) gives Kenney a mathematically expected loss or regret of $(1/4 \times 1) + (3/4 \times 1) = 1$, while Row Y (south) produces a loss of only $(1/4 \times 2) + (3/4 \times 0) = 1/2$, which is the better choice. The minimax strategy, erroneously, would have sent Kenney north (Row X), decreasing his pay-

---

[282] *Cf.* COLMAN, *supra* note 1, at 29 (setting forth a different matrix that also shows regrets rather than payoffs).

[283] *See supra* Part IV.A.

[284] *See supra* Part IV.A.

[285] This example is original with the author of this Article.

off. Thus, the minimax strategy is not necessarily best in conditions of uncertainty.[286] Also, it does not necessarily fit cooperative games or mixed motive games of the kind that we saw in Part IV.B. It is superior for zero-sum games against a rational opponent where the payoff matrix is known to both players (perfect knowledge),[287] but otherwise it has limits.

C. What Game Theory Shows Us About Strategy: The Choice of Maximax, Maximin, or Minimax Must Be Based on the Circumstances

Maximax may be justified if the other player is chance, with a fifty percent likelihood of indicating either column, and if one contains a very large payoff. This is the case in Figure 15 above, where a more cautious strategy would tell the player to seek a payoff of one cent rather than a fifty percent chance of obtaining one hundred million dollars. But more often, the maximax strategy of betting the company on the lottery is a poor one. This is so even though our society encourages risk-takers with romantic notions; for example, in popular courtroom dramas, the hero attorney always goes for broke and tries the big case, preferring a possibility of a maximum payoff over a sure settlement.[288]

The defects in the maximax strategy can be illustrated by imagining that attorney Joe Smith represents the plaintiff in a negligence suit.[289] The defendant, recognizing the possibility of an adverse verdict, has offered to pay $100,000 to settle the case. Smith figures that he might lose, too, and estimates that only five out of ten juries would return a verdict in his favor (i.e., there is a probability of 0.5 of a plaintiff's verdict). But if he does get a favorable verdict, Smith thinks his chances are one in ten of obtaining $500,000 in damages; the nine other estimated damage verdicts would average to $100,000, exactly the amount offered. Smith knows that the defendant's offer will be withdrawn forever if he does not accept it now. Nevertheless, he does not want to give up his chance to win $500,000.

The appropriate strategy in this circumstance is for Smith to forget the maximax strategy induced by the tempting $500,000 possibility and settle. His mathematical expectation from going to trial is (0.5 x 0) + [0.5 x (9/10)($100,000) + (1/10)($500,000)] = $70,000, which is less than the settlement offer of $100,000. Wishful thinking, competition, and cultural factors do tend to push a player toward the maximax strategy in these kinds of circumstances. As a strategy, however, maximax is dubious here. Legislation encouraging settlement offers, rewarding their accep-

---

[286] *See* COLMAN, *supra* note 1, at 32.

[287] *See supra* note 269.

[288] My own courtroom novels reflect this romantic irrationality. DAVID CRUMP, CONFLICT OF INTEREST (1997); DAVID CRUMP, THE HOLDING COMPANY (2000).

[289] This game framing is original with the author of this Article.

tance, and requiring settlement conferences helps to overcome the psychological tendencies of litigants to adopt maximax foolishly.

We already have seen that uniformly choosing maximin, or avoiding the worst of the worst, also is unwise.[290] But maximin, like maximax, cannot be ruled out in all circumstances. When would a rational player use the pessimistic maximin strategy? The answer is, when a major loss must be avoided at all cost.[291]

The mother's response to the "greedy father's child custody gambit" is an example. Let us say that Mary and John Smith are embroiled in a divorce proceeding. Mary would like her share of the marital assets and a significant amount of child support, but above all else, she needs to win custody of the children. John's desires are to obtain a large share of the assets and to keep support as low as possible. John therefore seeks custody himself but tells Mary that he will drop this claim if she accepts his proposed division of assets and child support. Mary may not think John has much chance of winning custody, perhaps only one chance in ten, but even this risk is unacceptable. It might be rational here for Mary to adopt the maximin strategy, to avoid at all cost the worst possible outcome.[292] This conclusion is tragic, because it makes Mary vulnerable to blackmail and allows John the strategic use of the equality of the genders to extract an unfair financial gain. But unfortunately, this is Mary's reality, unless as yet uninvented legislation can someday change the playing field to discourage John's strategy.

D. The Critique of Rawls's Uniform Maximin Strategy

Rawls's maximin strategy is not exactly mirrored by these games, of course. His use of maximin is to solve all legislated or customary differences within a society. But if game theory shows us that universal choices of strategy are unwise even in the limited context of zero-sum contests, perhaps it should be an indication that we should be suspicious of such a simple solution to address all differences among citizens. In fact, what Rawls's choice of the maximin strategy really means is that every economic and social policy must be evaluated by its effect upon the single person in society who is worst off. Is Rawls correct to choose this pessimistic principle for a just society? Or will his theory stifle growth, knowledge, art, and creativity?

The logic of Rawls's principle leads to the conclusion, for example, that minimum-wage legislation[293] has no place in a just society.[294] Putting

---

[290] *See supra* Part V.B.2.

[291] *See* COLMAN, *supra* note 1, at 27.

[292] This game framing is original with the author of this Article.

[293] *See, e.g.*, 29 U.S.C. § 206(a)(1) (Supp. III 1997) (setting the minimum wage at $5.15 an hour).

[294] Rawls discusses the notion of wage structures without resolving this dilemma.

a floor on the cost of labor benefits some people, and some policymakers will argue that the result is greater justice in the distribution of wealth.[295] There is also, however, a disadvantage to a minimum wage statute, in that a marginal employer, unable to pay the required amount, will decline to offer a job to a person he otherwise would have hired.[296] Rawls apparently would decide the entire question only by reference to people in this latter position. There are substantial efficiency arguments against minimum wages, and the unemployment that they cause is an obvious concern to legislators,[297] but it seems anomalous to decide the issue entirely on the basis of the maximin strategy, as Rawls's difference principle appears to require.

Likewise, the organic legislation creating the Federal Reserve System[298] seems out of bounds under a maximin regime. The Fed's task is to balance growth, employment, and stability.[299] One of its main techniques is to regulate the federal funds rate.[300] Other things being equal, an increase in the rate leads to higher rates of unemployment, while a decrease may create inflationary pressures.[301] Therefore, the Fed, through its Board of Governors, watches indices that predict inflation, and upon signs of undue price instability, it increases the federal funds rate.[302] It thereby puts upward pressure on unemployment.[303] Rawls's logic would require the Fed to abstain from this action if it resulted in layoffs of people who wanted to work.[304] This is the consequence of a uniform maximin strategy. In other words, Rawls's strategy would mean that we could not have legislation setting up a Federal Reserve Board, or if we did, it would be required to stand idly by while inflation spiraled upward, intervening only at that point at which the destruction of the economy became so complete that stopping inflation would benefit even the single worst-off person in the society.

RAWLS, *supra* note 8, at 307–10. He concludes that in a society with equal opportunity, "[t]he relative difference in earnings between the more favored and the lowest income class tends to close." *Id.* at 307. Elsewhere, he justifies welfare payments as a "social minimum" to be met by the "transfer branch" of government. *Id.* at 276–77.

[295] *See* Eric Schmitt, *Minimum Wage Rise of $1 Is Approved*, N.Y. TIMES, Mar. 10, 2000, at A1.

[296] *See id.*

[297] Rawls makes these arguments, in fact. *See* RAWLS, *supra* note 8, at 305–06.

[298] 12 U.S.C. §§ 221–522 (1994 & Supp. IV 1998).

[299] *See* JOSEPH E. STIGLITZ, ECONOMICS 727–28 (2d ed. 1997).

[300] *Id.* at 739.

[301] *Id.* at 739, 797–98.

[302] *Id.* at 739.

[303] *Id.* at 797–98.

[304] In theory, this result would not be required if taxation and transfer payments combined to compensate those who became unemployed. Strategically, however, this compensation could not place recipients in the same position as employed workers, or the latter would have no incentive to work. It would be necessary, therefore, for the Fed to make some people unemployed, and the unemployed thus would be made worse off by government policy benefiting the employed, in apparent violation of Rawls's difference principle.

Rawls seems to suggest that the hallmark of a just society is that its legislature measures every policy by its impact upon the most disadvantaged. Such a society, however, would suffer a ruined economy. Its citizens would have little incentive to save for the future, with a government unable to prevent inflation from destroying savings. Ironically, the next generation would experience a lower standard of living across the board as a consequence of the strategy of concentrating on those who experience the lowest standard today.[305] The society would produce less growth, knowledge, art, and science. Perhaps Rawls's choice of the maximin strategy is a reminder that a legislature in a just society should structure its laws with concern for the least advantaged among its citizens, but game theory should help us to understand why a legislature's relentlessly uniform application of a single, most pessimistic strategy might be undesirable.

VI. Other Games That Illuminate Equality: Arrow's Theorem, Maximizing Difference, and the Dollar Auction

Our coverage here has been arbitrarily selective. There are many other games that can cast light on the meaning of equality. Here, we shall briefly outline Arrow's Theorem, the Maximizing Difference game, and the Dollar Auction. The descriptions of these games will be more summary than those of the other games considered in this Article, but all three have cautionary tales to tell about equality.

A. Arrow's Theorem, Strategic Voting, and Equalities Within Democratic Institutions

Imagine that Alpha, Betty, and Gammy visit the grocery store with the intention of buying a carton of ice cream, which they will share. Alpha prefers vanilla, then chocolate, then strawberry, in that order. Betty's preferences are different: chocolate, then strawberry, then vanilla. And Gammy's are different still: strawberry, then vanilla, then chocolate. We can display these preferences visually by using a chart like Figure 17.

---

[305] Rawls devotes considerable attention to the duty of saving, the time value of money, and the issue of justice between generations, but without solving the problems identified here. *See* Rawls, *supra* note 8, at 284–98.

FIGURE 17

Alpha:	V > C > S
Betty:	C > S > V
Gammy:	S > V > C

This conundrum illustrates Arrow's Theorem, named after Kenneth Arrow, who demonstrated how the votes of multiple decision-makers tend to result in stalemate.[306] A unitary election among Alpha, Betty, and Gammy will result in one vote each for vanilla, chocolate, and strawberry, and the shoppers will be unable to make a choice. Game theorists refer to this phenomenon as "cycling."[307] Arrow showed that for any number of choices and contestants greater than two, cycling of this kind ultimately is unavoidable.[308]

The only ostensibly equal way to resolve such a stalemate short of flipping a coin or buying striped Neapolitan ice cream is to reduce the choices to two. The three shoppers can begin with a choice between vanilla and chocolate and then choose between the winner of that election and strawberry. Alpha and Gammy will vote for vanilla in the first election. Then, however, Betty and Gammy will vote for strawberry. Alpha, having won the first election with her first choice, winds up after the second election as the loser. Strawberry wins, but it is her last choice.

Worse yet, Alpha now realizes that this outcome is the result of an arbitrary institution: the order of voting. If the first vote had pitted chocolate against strawberry, it would have resulted in a second election in which vanilla would have won. Or, if vanilla runs head-to-head with strawberry in the first election, the second produces a win for chocolate. In summary, it is possible to produce a win for any one of the three choices—vanilla, chocolate, or strawberry—just by putting the other two in the first election.[309]

How do real-world voters and legislators avoid the cycling that Arrow's Theorem predicts? They use a variety of techniques relating to what is known as "strategic voting."[310] First, Alpha might try parliamentary maneuvers to make the first election a choice between chocolate and

---

[306] KENNETH J. ARROW, SOCIAL CHOICE AND INDIVIDUAL VALUES 96–100 (2d ed. 1963); *see generally* Frank Easterbrook, *Ways of Criticizing the Court*, 95 HARV. L. REV. 802 (1982) (applying the theorem to demonstrate why inconsistency among decisions of appellate courts is inevitable); BARNES & STOUT, *supra* note 242, at 46–47 (explaining Arrow's Theorem).

[307] *See* BARNES & STOUT, *supra* note 242, at 46; COLMAN, *supra* note 1, at 237.

[308] ARROW, *supra* note 306, at 96–100.

[309] *See* BARNES & STOUT, *supra* note 242, at 46.

[310] *See* DAVIS, *supra* note 1, at 219.

strawberry, so that the end result will be vanilla. Failing that, she can vote strategically for her disfavored choice in the first phase of the decision. If the choice is between vanilla and chocolate, in other words, Alpha should vote not for vanilla, her favorite, but for chocolate, her second choice. In this way, she salvages her second choice in the latter election, because she foresees that otherwise she will get her third.

This strategy avoids the result of Arrow's Theorem. According to Arrow, for cycling to be predictable, votes must be "independent."[311] That is to say, they must not be influenced by future votes or by vote-trading. Moreover, choices must be "transitive."[312] If vanilla is preferred to chocolate and chocolate to strawberry, vanilla must be preferred to strawberry. Alpha has avoided the result of Arrow's Theorem by negating the required conditions.

Arrow's Theorem has led to a vast literature about public choice. Judge Easterbrook, for example, has used it to demonstrate that inconsistency among decisions of an appellate court are virtually inevitable.[313] The Theorem also explains why a run-off election may not equally reflect the democratic preferences of voters,[314] as well as why political parties are necessary, even though they may result in unequal treatment of voters' choices.[315]

In terms of equality, Arrow's Theorem suggests a disappointing conclusion. In a democracy, there may be no way to ensure that every person's vote will be equal to every other person's. If we are to make decisions at all, we must avoid cycling, and therefore we must depend on institutions such as political parties and run-off elections, which narrow the available choices. The order in which sequential choices are made, even if arbitrarily arrived at, may determine the outcome. Log rolling, horse trading, and a go-along-to-get-along philosophy may look like flaws in a legislature, but in reality they are necessary if anything is to be accomplished.

B. The Maximizing Difference Game: Is a Choice of Self-Destructive Inequality Motivated by Ignorance, Competition, or Boredom?

The "Maximizing Difference" Game actually is of little strategic interest to game theorists. In fact, it is trivial in strategic terms.[316] The payoff matrix is in the same form as those for Battle of the Sexes, Chicken, the Prisoners' Dilemma, and Leader, but the maximum payoff for both

---

[311] More precisely, this condition is said to require the "independence of irrelevant alternatives." ARROW, *supra* note 306, at 27.

[312] *Id.* at 13.

[313] *See generally* Easterbrook, *supra* note 306.

[314] *See* DAVIS, *supra* note 1, at 218.

[315] *See id.* at 214.

[316] *See* COLMAN, *supra* note 1, at 136, 138–39.

players is reached in a single square whose row and column are dominant.[317] The Maximizing Difference game is diagramed in Figure 18. The rational strategy for Player 1 is to choose the top row, and it results in the maximum payoff if Player 2 also has acted rationally by choosing the left column. Both players receive uniform maximum payoffs of (4, 4).

FIGURE 18
THE MAXIMIZING DIFFERENCE GAME

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	4, 4	2, 3
	Defect	3, 2	1, 1

In spite of its strategic triviality, Maximizing Difference has been the object of considerable study. This is because although the choices may be indicated clearly by game theory, the behavior of real human beings who play the game remains unpredictable. The results of psychological experiments with real subjects playing the Maximizing Difference game are fascinating. There are surprisingly high frequencies of seemingly irrational strategies in which players tend to choose the "defect" option.[318] Charles McClintock and Steven McNeel, for example, gave pairs of students the opportunity to win money over a hundred iterations of the game.[319] They found that the students chose the competitive strategy over sixty percent of the time even though cooperation always remained the rational choice for both players.[320]

One can only speculate about the mechanism for this result. First, it is conceivable that the subjects misunderstood the payoff structure. In other words, ignorance may have driven their strategies, and they may have believed they would benefit from obtaining a more than equal share. But this explanation seems incomplete, given the simplicity of the game and given the play of experimental subjects in other, more complex games.[321] An alternative explanation is that a competitive inclination led

---

[317] *Id.*

[318] *Id.* at 138–39.

[319] Charles G. McClintock & Steven P. McNeel, *Prior Dyadic Experience and Monetary Reward as Determinants of Cooperative and Competitive Game Behavior*, 5 J. PERSONALITY & SOC. PSYCHOLOGY 282, 284–85 (1967).

[320] *See id.* at 286.

[321] *See* COLMAN, *supra* note 1, at 138.

subjects into irrational strategies. Rather than maximizing their payoffs by cooperative equality, they may have been attempting to "win," even though Maximizing Difference cannot be "won" by defect strategies, in the sense that defection always brings lower payoffs to both players. McClintock accepts this explanation:

> In effect, the high reward subjects were still willing to forego considerable material gain which would have obtained if they had maximized joint gains to ensure that they would have more points than the other, or at least, not fall appreciably behind the other in score.[322]

But there is another possibility, a "disturbing" one, as Colman labels it. "[M]any if not most of the subjects must have been playing the game according to a utility structure different from the payoff matrix."[323] In other words, the players may prefer to shake things up just for the fun of it, even though they lower their own monetary welfare as a result and even though they cause even greater losses to their counterpart players. Equality is not as exciting as combat, and there may be a perverse payoff from disadvantaging one's fellow players, perhaps resulting from simple boredom at the sameness of equal results.

The implications of these experiments for legislation affecting equality are disturbing, too. Just as Arrow's Theorem demonstrates a flaw in democratic institutions, the Maximizing Difference experiments demonstrate a flaw in the individuals that make up a democratic citizenry. They show just how ornery and irrational human beings can be when shared welfare maximization is at issue. The results suggest that perhaps people will engage in invidious discrimination even when they know it is against their interests, because they hope to disadvantage others more than they are disadvantaged. Perhaps these experiments even hint at a motive for hate crimes: ruffians who know they face substantial odds of getting caught and punished may nevertheless derive perverse utility from combinations of their own welfare losses with the satisfaction gained in knowing that their victims' losses are greater.

C. The Dollar Auction: How a Legislature Succumbs to the Concorde Fallacy

The third and final game is the "Dollar Auction." An auctioneer offers a dollar to the highest bidder. The bidding will start at one cent, and

---

[322] C.G. McClintock, *Game Behavior and Social Motivation in Interpersonal Settings*, *in* EXPERIMENTAL SOCIAL PSYCHOLOGY 291 (C.G. McClintock ed., 1972).

[323] COLMAN, *supra* note 1, at 139. Colman finds this disturbing because it undermines the entire field of experimental games. *Id.*

it has no upper limit. But there is a catch: the player who makes the second-highest bid must pay the amount of the bid just as the winner does, but the second-highest bid wins nothing.[324] The Dollar Auction has three magic moments: first, the instant when the second bid occurs, because this means that there must be a loser; second, the instant when the highest and second-highest bids average more than fifty cents, because the auctioneer now is assured of making a profit; and third, the instant in which one of the players bids more than one dollar.[325] This last strategy may be chosen by a competitor who does not want to wind up as the second bidder, having concluded that it is better to pay more than a dollar for the dollar than to pay nearly the same amount without winning the dollar.

In psychological experiments, subjects usually are provided with amounts of money with which to bid (or with points that they understand are exchangeable for money). In the first controlled experiment of this kind, by Richard Tropper, bidders paid more than the value of the prize in sixteen out of thirty cases.[326] Allan Teger's later results were even more striking: the winning bid in forty trials always exceeded a dollar, and sometimes it reached twenty dollars.[327] In subsequent experiments, players often escalated their bids until they exhausted their resources.[328] Players seduced into this bidding spiral responded by exhibiting signs of stress.[329] Later experiments in which subjects' heart rates were monitored showed that their rates increased as the auction approached the moment when the bidding would exceed one dollar.[330]

This phenomenon is similar to the so-called "Concorde fallacy," or entrapment in a losing venture in which the legislature continues to escalate the amount of likely losses because there has been "too much invested to quit."[331] The name comes from the supersonic airliner produced by a consortium funded by the British and French governments. Even after costs had escalated to the point where any well-informed, disinterested observer could tell that the Concorde was uneconomical, the consortium members continued to invest in it. In fact, their determination to see it through, as reflected by their statements, actually increased.[332]

In human subjects, it appears that this behavior, this entrapment in escalation, results from a confusion or shift of objectives. The initial goal is monetary, that of buying something for less than it is worth, but it

---

[324] *See* COLMAN, *supra* note 1, at 192.

[325] *See id.*

[326] *See* Richard Tropper, *The Consequences of Investment in the Process of Conflict*, 16 J. CONFLICT RESOL. 97, 97–98 (1972).

[327] ALLAN I. TEGER, TOO MUCH INVESTED TO QUIT 45–60 (1980).

[328] *See* COLMAN, *supra* note 1, at 197.

[329] *Id.*

[330] *Id.* at 197–98.

[331] *Id.* at 191.

[332] *Id.*

changes in the course of escalating bids into a fear of loss or embarrassment.

> [T]he initial bidding appears to be motivated by economic concerns, but the tendency to bid until you are broke is due to a new motivation that develops during the course of the auction—a motivation toward competition which makes the economic considerations less important.[333]

This psychological entrapment appears to increase when groups are the decision-makers about bids, presumably because of the normative group influence.[334] It can be reduced, if not eliminated, by the imposition of a penalty, or "tax," experimentally imposed on bids that exceed one dollar.

The Dollar Auction, like the Maximizing Difference Game, spells bad news for efforts to preserve equality. The normative group influence, which disposes group bidders even more toward escalation, seems to exacerbate the problem.[335] If two or more groups must share resources distributed through the political system, the Dollar Auction suggests that their initial cooperative motives related to mutual welfare maximization may shift toward mutually destructive competition. In other words, they are pushed psychologically toward an effort to minimize their perceived losses at the expense of competing groups, even when it is clear that this strategy results in lowered welfare for everyone. A legislature will be besieged by lobbyists who not only seek their own welfare but also seek not to lose anything to other groups.

VII. CONCLUSION

A. Game Theory as Strategy: Its Intrinsic Value in the Law

There is a bit of an analogy to the Trojan horse reflected in this Article. If it succeeds in exposing law students and professors to game theory, it will have achieved its objective at least partially. This would be so even if it drew no useful conclusions about legislation or equality. Enhanced awareness of strategy would particularly improve legal education because law school teaches by analyzing the texts of court opinions. Unfortunately, the discussion of legal strategies is not a necessary part of this method, which focuses upon Socratic dialogue about the court opinions themselves.

---

[333] TEGER, *supra* note 327, at 60.

[334] *See* COLMAN, *supra* note 1, at 199 (explaining the effect as the "group polarization phenomenon").

[335] *Id.* at 199–200.

To put the matter another way, law school teaching from casebooks tends to emphasize theoretical differences in result as inferred from reasoning and holding, and upon the policy implications of varying these factors. Although human beings, as strategic entities, will play the system in unpredictable ways that might profoundly change the meaning of the discussion, recognition of this inconvenient fact need not enter into the discussion. For example, one readily can imagine a law school class considering a case presenting an analog of the truel problem without the professor's having noted that the ostensible loser, Clark, holds the tactical advantage. If this Article exposes law students and professors to game theory, it may make them more aware of strategy, which might alter their conclusions when they discuss appellate opinions. In this way, the game theory explicated here has an intrinsic value even apart from its implications for equality.

B. Game Theory, Legislation, and Equality

Game theory also can provide us with useful insights into legislation that affects equality. It convincingly shows us, for example, that there are many equalities. It also tells us that it is impossible to distinguish equal from unequal distributions by logic alone. In spite of this indeterminacy, the logic of strategy indicates that definitive intervention is necessary to redress the inequalities that we decide are unacceptable. In a related way, game theory demonstrates that equality is strategic in that citizens will be influenced by rational self-interest in ways that deviate from societal efforts to enhance equity.

But game theory also can remind us that equality is not the only value of importance, and sometimes its sacrifice to enhance freedom or autonomy may be justified. At the same time, the psychological experiments suggest that people sometimes will react with surprising irrationality in preferring unequal outcomes precisely because they are unequal, even though this behavior decreases everyone's welfare, including their own. Finally, mixed-motive games illustrate the importance of fine differences in payoff structures that may call for intervention in some circumstances but not in others that seem closely similar.

1. There Are Many Equalities, and a Simplistic Focus on a Single Version May Prove Misleading

The first way in which game theory is relevant to legislation affecting equality is that it helps us to recognize that there are many versions of equality. The stockholders' game is particularly forceful in supporting this insight. Simple or naive equality consists of giving each stockholder a numerically equal share. This distribution plausibly can be presented as equal in many situations; after all, it parallels the ingrained concept of

one person, one vote. But a legislated regime mandating individual equality also can be attacked as unequal. It arguably treats people equivalently when they really are not the same, because they have made different choices about indulgence or postponement of satisfactions.

By-the-shares equality is a way of taking account of these differences. But the stockholders' game shows that a by-the-shares distribution is only one of many possible solutions. Efficient use equality, which distributes according to marginal utility, also takes account of differences among the distributees. It simply focuses on differences of a different kind.

Then, there is the Shapley value solution, which distributes equally in accordance with each player's contribution to a majority coalition. This solution seems unequal at first blush, since it gives the largest stockholder an amount exceeding the actual proportion of shares she holds, and it distributes equal amounts to smaller stockholders who actually have different holdings. But there are other ways to view the Shapley solution, ways that make it seem more appealing. It is a relatively stable result of the democratic process. It emphasizes equality of opportunity, whereas other solutions depend only on equality of result. It encourages thoughtful choices. It recognizes and equalizes past and present benefits, contributions, and sacrifices. And if the biggest stockholder has paid a control premium, or even if all stockholders had a chance in the beginning to know the proportions in which they were investing, the Shapley solution may be superior in providing equality, in the sense that it produces equal satisfaction of legitimate expectations.

In legislative battles among real-world people in situations analogous to the stockholders' game, the players probably will argue for the types of distribution that will most benefit them or their constituents individually. This approach is merely rational strategy. It is important to realize, however, that the arguments may not emphasize the relative desserts of each player. Instead, they may center upon determining the rules of the game itself: the shape of the playing field, or the vision of equality that each player prefers. In the stockholders' game, for example, the critical issue is the choice among different concepts of equity. After that choice has been made, applying the chosen distribution to the factual differences among players is straightforward.

The problem becomes more difficult when each player fastens upon a single, idiosyncratically preferred version of equality and lobbies the legislature to accept it, without acknowledging other models of equality. Thus, the small-to-middling stockholder may argue for individual equality because this distribution benefits that shareholder more than any other regime. And this smaller stockholder rationally should attempt to structure the dispute so that individual equality seems to be the only sensible solution. In this situation, a lawmaker must struggle to remain alert to the possibility of alternative equalities. Game theory is useful in reminding

us of this possibility. Unfortunately, its logic cannot provide us with a clear map of all equalities in a given human encounter, nor can it tell us how to choose among them.

In *Dandridge v. Williams*,[336] the Supreme Court faced precisely this problem. The plaintiffs challenged a Maryland regulation that capped welfare benefits in the program for Aid to Families with Dependent Children so that the amount available to larger families was limited.[337] This approach meant that the government paid more on account of an only child than a child in a four-child family.[338] A test based solely on individual equality might have resulted in a holding that the Maryland law was unconstitutional. This view, not surprisingly, was the one championed by the plaintiffs.[339]

Justice Stewart's opinion for the majority, however, recognized Maryland's interest in balancing multiple equalities.[340] For example, the Maryland legislature needed to ensure that the program did not operate to the disadvantage of working poor families.[341] Similarly, although the court did not mention this rationale,[342] it is possible that Maryland had decided that it was important not to disadvantage a parent who thoughtfully acted to limit her number of children. The messy pluralism at the heart of the *Dandridge* opinion may be unsatisfying to readers who would prefer a single, simple theory of legislative equality, but it is reflective of the logic inherent in a complex strategic encounter. Game theory does not provide the human context that enables us to recognize these precise concerns, nor does it steer us to the particular statutory compromise among them that we might think is appropriate. But if it is consulted with proper caution, the logic of strategy can help us to ask the right questions.

2. *Equality Requires Legislation in Circumstances That Cannot Be Determined by Logic*

The recognition that there are many equalities, however, creates a new kind of difficulty. Game theory predicts that players sometimes will form coalitions that freeze out other players simply for the purpose of increasing their own advantage at the others' expense. In the stockholders' game, for example, minimum winning coalition theory indicates that the largest stockholder and one of the smaller ones, if they are rational,

---

[336] 397 U.S. 471 (1970).
[337] *Id.* at 473–75.
[338] *See id.*
[339] *Id.* at 476–77.
[340] *Id.* at 486–87.
[341] *Id.* at 486.
[342] As the Court put it, "We need not explore all the reasons that the State advances in justification of the regulation." *Id.*

should agree to arrogate to themselves benefits that might otherwise be shared with the third stockholder. Minimum winning resource theory is similar, but it predicts that smaller shareholders will coalesce, paradoxically, to freeze out larger ones.

To a game theorist, such behavior is neither bad nor good. It is merely rational and self-interested. Logic, therefore, cannot distinguish exploitive coalitions from just distributions. In fact, there are circumstances in which minimum winning coalition (or resource) distributions can be supported by arguments that make them appear equal, perhaps even persuasively so. Furthermore, it is difficult logically to distinguish exploitive effects of minimum winning coalition or resource theories from distributions such as the Shapley value solution, which have much more to recommend themselves by way of equality.

There is no way out of this dilemma, unfortunately, except to cut the Gordian knot. Imperfectly, and with shifting conceptions of the principles that guide them, legislators must differentiate distributions that rely excessively on power and exploitation from other, logically indistinguishable distributions that they are willing to tolerate as reflecting arguably acceptable equalities. They have nothing, ultimately, except cultural values and social conventions to rely on in making these judgments. This indeterminacy explains why Supreme Court opinions considering the limits of suspect classes or fundamental rights so often seem logically unsatisfying.[343] The distinctions are not logical but empirical, reflecting values that judges have internalized from experience. The process is messy, the results inconsistent. Thus, for example, the *Dandridge*[344] case suggests that wealth differences can reflect an acceptable kind of equality, but *Boddie v. Connecticut*[345] shows that the addition of a single factor, such as access to justice, can make wealth distinctions unjust and unequal. *United States v. Kras*[346] further tells us that relatively minor changes in the nature of the access to justice at issue can reverse the result.

When we conclude that equality requires intervention, our principles should be firm, clear, and hard-edged. The arguable inequality in a minimum winning coalition solution, for example, cannot be reversed except by definitive changes in the rules. By analogy, this conclusion explains why heightened scrutiny is appropriate for legislation that implicates fundamental rights or suspect classes. Game theory does not tell us how to recognize these situations because they are not identifiable by logic, but perhaps it can help us to recognize both the indeterminacy we face in

---

[343] *See supra* notes 113–119 and accompanying text.

[344] Dandridge v. Williams, 397 U.S. 471 (1970).

[345] 401 U.S. 371 (1971) (holding that prohibitively expensive procedures in the form of filing fees unaffordable to poor people violate the Due Process Clause).

[346] 409 U.S. 434 (1973) (holding that a filing fee analogous to that held unconstitutional in the context of divorce proceedings is constitutional in the context of bankruptcy proceedings).

trying to secure equality and the resoluteness we need when we decide that legislative intervention is appropriate.

3. Equality Depends upon Legislative Strategies

The third lesson from game theory is related to the first two. It is that equality is strategic. The truel problem convincingly demonstrates that rational actors will behave in ways that defeat the expectations of policymakers who try to equalize them. It also shows that apparently similar means of legislatively redressing perceived inequalities may have vastly different effects. Government too must act strategically when it chooses among alternatives. It must anticipate the possible strategic responses of those who will play the system.

In other words, game theory reminds us that human beings are autonomous decision-makers whose self-interested rationality, although not perfectly controlling, significantly influences their behavior. It should not be surprising that people often work strenuously to defeat efforts to equalize them. Instead, this behavior is characteristic of human actors who strive, usually commendably, to improve their situations. As a result, some people will try to exploit gaps in the rules so that they can emerge in better positions than their neighbors, rather than in equal ones. Human beings, in one view, are ornery and self-centered, and they act in ways that aggravate imperfections in regimes of equality, rather than tailoring their conduct to advance legislative aims. But then, this is part of the charm of human beings. Less judgmentally, one could conclude that a rule about equality that harnesses the self-interest of the people subject to it is more likely to succeed than one that wishes this factor away.

The picture is complicated, however, because social interactions do not always reflect strategies that seem obvious from the outside. This is apparent from the truel paradox and the phenomenon of strategic voting as a response to Arrow's Theorem. Therefore, before a legislature intervenes to enhance equality, it should take pains to gauge accurately the effects of existing as well as future players' strategies. This assessment may require a careful determination of current advantages, which also may not be obvious. In the truel problem, it takes determinedly counterintuitive analysis to figure out that Clark, the apparently weakest combatant, actually holds the winning hand.

Moreover, the legislature cannot rely on inputs from the players to make this determination. The players' arguments about the strategic situation will themselves be strategic. In the truel game, for example, if the government considers intervening on his behalf, Clark's rational response will be to advertise, lobby, and electioneer in ways that enhance his appearance of helplessness. He will try his best, in other words, to mislead the politicians. And if he succeeds, he then will proceed to con-

tradict his own arguments by exploiting the very government initiative he has precipitated, thus increasing his own unequal advantage.

The failure to account for such strategic responses is one of the major flaws in Rawls's theory of justice. Rawls treats the poorest citizens as purely reactive objects, rather than as autonomous decision-makers. His maximin solution, which makes every legislative decision depend upon the single worst-off person or category of persons, does not allow adjustment for the ways in which different policy choices might dysfunctionally motivate the very people they are designed to serve. In Rawls's strange world, wealth tends to equalize automatically (for reasons that Rawls does not explain),[347] and there is "no need even for the penal law except insofar as the assurance problem [makes] it necessary."[348] Rawls assumes that everyone in a just society accepts the propriety of "motivation duly regulated by a sense of justice."[349] This caricature of human action naturally results in a conception of all poor people as uniformly unresponsive both to incentives to better themselves and to temptations to exploit imperfections in the regime of equality.

Rawls's solution does provide some important insights into the character of a just society, particularly insofar as it focuses attention on the duty to improve the distributive shares of those who have the least. But the relentlessness of his second principle, which insists on always focusing on benefiting disadvantaged people, leaves no room for considering the strategic tendencies of poorer citizens as rational decision-makers. Worse, it treats poor people entirely as an unthinking mass, rather than as human beings who might be influenced somewhat by incentives or strategies.

The majority opinion in *Dandridge v. Williams*[350] recognizes Maryland's legislative interest in discouraging recipients from engaging in what might otherwise be rational but harmful strategies and encouraging desirable strategies for people situated similarly, but not identically, to recipients.[351] Rawls's theory would have required the Maryland legislature to subordinate these considerations to the single criterion of benefiting the most disadvantaged. Game theory would not have identified the precise balance of interests that Maryland should have adopted, but it reminds us to account for rational strategies.

---

[347] *See supra* note 294 and accompanying text.
[348] RAWLS, *supra* note 8, at 315.
[349] *Id.*
[350] 397 U.S. 471 (1970).
[351] *See id.* at 483–84.

4. The Limits of Logic: Just as People May Act Strategically, They Also May Act Non-Strategically or Even Irrationally with Respect to Equality, Particularly If It Conflicts with Other Values

Logical rationality is not the only approach that human beings can adopt in strategic situations. As the psychological experiments demonstrate, they sometimes fail to understand rewards and incentives. Sometimes, they displace mutual or even individual maximization of ostensible payoffs with other motivations, such as the competitive desire to win or the urge to shake things up merely to relieve boredom. Studies of the Maximizing Difference Game vividly illustrate this disturbing phenomenon. And sometimes stress, confusion, and conflict make people act perversely. The entrapment in escalation so characteristic of the Dollar Auction, particularly given the astonishing frequency with which it occurs, reinforces this conclusion.

Sometimes people may reject equality even when everyone, including themselves, would be better off with a solution reflecting cooperative values. Psychological experiments with games may help us to recognize this phenomenon. Thus, an individual may opt for a strategy that damages both himself and those with whom he is interdependent, not in an effort to better his own lot in the absolute sense, but to disadvantage his neighbors more than himself. This phenomenon may mean that insidious forms of discrimination will prove more persistent than we prefer to believe. Worse yet, experiments with group behavior demonstrate that the effect may be exaggerated in groups. Thus, we should not be surprised to find that groups whose mutual welfare would increase as a consequence of cooperative equality nonetheless compete vigorously in the legislative arena in an effort to gain at each other's expense, even though this may be a losing strategy in terms of ostensible payoffs.

5. The Tendency Toward Equality in Mixed-Motive Situations Depends on the Payoff Structure

The four interesting mixed-motive games of Battle of the Sexes, Chicken, the Prisoner's Dilemma, and Leader provide particularly interesting insights into equality. Battle of the Sexes shows that sometimes equality does not maximize mutual or even individual welfare. Sometimes the best solution, not only for the community but for every individual in it, may depend on a deliberate legislative choice to permit inequality. In other circumstances, in which equality may result either in the best or the worst individual welfares, the Prisoner's Dilemma suggests that repeated encounters result in such strong encouragement of cooperation that the prospect of inequality largely can be ignored. Unless the participants misperceive the payoff structure or psychologically covert it into a game with different utilities, self-interest naturally will motivate them

toward strategies that maximize both welfare and equality. Legislation will be counterproductive, except to assure that the level playing field remains intact.

The perverse game of Chicken actually rewards the bully who seeks to increase his share at everyone else's expense. Whenever the payoff structure resembles that of Chicken, or in other words whenever credible recklessness is a rational strategy because it can precipitate deference in rational opponents, equality depends critically upon accurately targeted legislative intervention. Situations in which an estate to be divided is limited, so that the cantankerous bankruptcy lawyer or the avenging spouse can play the role of the skillful Chicken strategist, provide examples of areas where legislation is needed. So do other situations in which the costs of battle far exceed the rewards, such as the types of litigation discovery in which the behavior of a Rambo-like lawyer may prove successful. In these circumstances, unlike those of the Prisoners' Dilemma, equality does not result from a hands-off philosophy, or from reliance on the maxim that what goes around, comes around. Instead, iterated games aggravate the inequality by strengthening bullies and weakening the rational actors who defer to them. Courts seem to perform poorly both at adjudicatively differentiating the Chicken payoff structure from that of the Prisoners' Dilemma and at intervening effectively when the game resembles Chicken. Instead, legislative intervention creating principles that discourage or penalize the reckless bankruptcy or divorce player, and that convert the payoff structure from Chicken to Prisoners' Dilemma, may be more successful.

Finally, the game of Leader recognizes the need for rules that sacrifice precise equality in favor of welfare-maximizing coordination. Whether we drive on the right side of the road or the left does not matter much, but it matters a great deal that the legislature provides a uniform convention governing this issue. Laws such as the birthday rule reflect minor sacrifices of immediate equality, but they tend to produce even results over many encounters, and they create welfare gains for all parties that outweigh their incidental inequalities. In this situation, paradoxically, arbitrarily determined outcomes may lead both to utility increases for everyone and, for iterated encounters, to rough equality.

6. Game Theory Cannot Provide Precise Prescriptions for Securing Equality

I conclude with the disclaimer that began this Article. Game theory reduces strategic encounters to simplistic structures so that logic can resolve them. Even a relatively complex game such as chess, whose solution eludes our current capabilities, features defined moves for each piece, a finite (if astronomical) range of strategies, and clear payoffs in the form of wins, losses, and draws. In contrast, human encounters reflect

infinite variations. Game theory cannot tell us how people will react to strategic situations. It cannot include all factors that might influence human behaviors, and it cannot supply the customs and values that make a given solution seem acceptable or unacceptable, equal or unequal.

Nevertheless, game theory sometimes can alert us to possibilities that we otherwise might overlook. It can help us to generate different views of acceptable equalities, to which we then can apply accepted values. It can help us to remember the strategies that individuals may use (and that they characteristically will use) when they are motivated to avoid equalization by bettering themselves. In related ways, it demonstrates that equality is not the only value of interest and that sometimes it conflicts with other values so profoundly that inequality is tolerable or even preferable to a relentless rule forcing each player to the same level.

At the same time, game theory can help us to avoid the logic that might equate exploitive situations with equitable ones. It reminds us that fine alterations in payoff structures can motivate people toward either cooperative equality or mutual destruction, and it suggests ways that legislation can change incentives to avoid the latter. It tells us that two apparently equivalent legislative interventions may produce entirely different results in terms of equality. In summary, if the logic of strategy is consulted as a heuristic for generating questions and exploring options, rather than as a prescriptive algorithm, it can provide limited but valuable guidance for legislative efforts to enhance equality.

ARTICLE

WHAT BUY-OUT RIGHTS, FIDUCIARY DUTIES, AND DISSOLUTION REMEDIES SHOULD APPLY IN THE CASE OF THE MINORITY OWNER OF A LIMITED LIABILITY COMPANY?

SANDRA K. MILLER*

With the advent of the new IRS "check-the-box" regulations that permit a limited liability company ("LLC") to choose whether to be taxed as a corporate or partnership entity, LLCs no longer need to structure themselves to avoid the corporate characteristic of continuity of life. As LLCs are beginning to provide for an indefinite life, minority shareholders are facing the elimination of their default withdrawal rights. Given the inherently unequal bargaining power of such relationships, minority shareholders are put at a distinct disadvantage, especially given the lack of developed law for LLCs. This Article argues that minority shareholders must be protected by maintaining statutory buy-out rights at least until LLC statutes protect the rights of minority shareholders explicitly. Professor Miller discusses fiduciary duties in the LLC context and recommends two types of statutory protections, both of which are contained in the Uniform Limited Liability Company Act, namely a prohibition on unreasonable reductions in fiduciary duties and a mechanism for judicial dissolution.

The elimination of rigid tax requirements for obtaining partnership tax treatment for unincorporated business entities has already begun to have a dramatic impact on the development of business entities. As a result of the Internal Revenue Service's adoption of so-called "Check-the-Box" regulations, partnership and limited liability company ("LLC") statutes need not defeat the "corporate" characteristic of continuity of life.[1] It is no longer necessary for partnership and LLC statutes to contain

---

* Associate Professor, Widener University School of Business Administration, B.S., S.U.N.Y. at Stony Brook, 1976; J.D. Hofstra University School of Law, 1978; M.S., American University Graduate School of Business, 1981; LL.M., Villanova University School of Law, 1990. Professor Miller was an observer of the drafting committee on the Uniform Limited Liability Company Act of the National Conference of Commissioners on Uniform State Laws. The author wishes to acknowledge the helpful comments and support of Kenneth C. Miller, the research assistance of Gina Newman and Nicole Gallipoli, and the complimentary services provided by LEXIS/NEXIS.

[1] *See generally* Susan Pace Hamill, *The Limited Liability Company: A Catalyst Exposing the Corporate Integration Question*, 95 MICH. L. REV. 393 (1996) (discussing the revision in the income tax regulations governing LLCs and its tax policy implications); Thomas M. Hayes, Note, *Checkmate, the Treasury Finally Surrenders: The Check-the-Box Treasury Regulations and Their Effect on Entity Classification*, 54 WASH. & LEE L. REV. 1147 (1997) (exploring the impact of the change in Treasury regulations concerning tax classification of taxpayers); David A. Weisbach, *Line Drawing, Doctrine, and Efficiency in the Tax Law*, 84 CORNELL L. REV. 1627 (1999) (analyzing tax trends under new develop-

events that potentially dissolve the firm in order to gain favorable flow-through tax treatment as a partnership.[2]

Before the legislative ink on the early rounds of statutes dried,[3] a growing number of LLC statutes eliminated the LLC member's right to withdraw from the company prior to its dissolution and winding up, or otherwise eliminated the right to be paid the fair market value of the LLC

---

ments in the tax law).

[2] *See* I.R.C. § 7701 (2000); Treas. Reg. § 301.7701-1(a) to -7(f) (2000) (setting forth tax classification rules that permit taxpayers to elect to be taxed as partnerships or corporations irrespective of whether the taxpayer possesses limited liability, continuity of life, or other characteristics traditionally associated with corporate status); BORIS I. BITKER & JAMES S. EUSTICE, FEDERAL INCOME TAXATION OF CORPORATIONS AND SHAREHOLDERS § 2.01 (6th ed. 2000) (discussing the definition of a corporation, the history of the tax classification regulations, and the current elective regulations for business entities); *see also* 18 STAND. FED. TAX REP. (CCH) §§ 43,084–43,087.30 (2000) (explaining the current tax classification rules and the relaxation of the prior rigid regulations); Karen C. Burke, *The Uncertain Future of Limited Liability Companies*, 12 AM. J. TAX POL'Y 13 (1995) (analyzing the evolution of the tax classification rules and questioning the continued viability of taxing corporate income at the entity level, as well as corporate dividend distributions on the individual shareholder level under subchapter C of the I.R.C.); Hayes, *supra* note 1, at 1159 (discussing the revision of the tax classification regulations under pressure from the legal and business communities).

[3] *See* Carter G. Bishop, *Treatment of Members Upon Their Death and Withdrawal From a Limited Liability Company: The Case for a Uniform Paradigm*, 25 STETSON L. REV. 255, 259 (1995) (discussing the "first generation bulletproof" LLC statutes that mandated compliance with formerly rigid tax classification regulations, and the later LLC statutes that permitted increased flexibility in drafting LLC agreements). This first round of statutes included: ALA. CODE § 10-12-1 to -61 (1998); ALASKA STAT. § 10.50.010–.995 (Michie 1998); ARIZ. REV. STAT. § 29-601 to -857 (1999); ARK. CODE ANN. § 4-32-102 to -1316 (Michie 1997); CAL. CORP. CODE §§ 17000–17705 (West 1999); COLO. REV. STAT. ANN. § 7-80-101 to -913 (West 1999); CONN. GEN. STAT. ANN. § 34-100 to -243 (West 1997); DEL. CODE ANN. tit. 6, § 18-101 to -1107 (1999); D.C. CODE ANN. §§ 29-1301-1.1 – 32-552 (1998); FLA. STAT. ANN. § 608.401 – .703 (West 1999); GA. CODE ANN. § 14-11-100 to -1109 (Harrison 2000); IDAHO CODE § 53-601 to -672 (Michie 1998); ILL. COMP. STAT. 180/1-1 to 205/7-1 (1998); IND. CODE ANN. § 23-16-10.1-1 to 23-18-12-11 (Michie 1998); IOWA CODE ANN. § 490A.100-1601 (West 1997); KAN. STAT. ANN. § 17-7601 to -7651 (1997); KY. REV. STAT. ANN. §§ 275,001–275,455 (Michie 1998); LA. REV. STAT. ANN. § 12:1301 to 1369 (West 1999); ME. REV. STAT. ANN. tit. 31, §§ 601–762 (West 1997); MD. CODE ANN., CORPS. & ASS'NS § 4A-101 to -1103 (1999); MASS. ANN. LAWS ch. 156C, §§ 1–68 (Law. Co-op. 1999); MICH. COMP. LAWS ANN. § 450.4211 to -5200 (West 1999); MINN. STAT. ANN. § 322B.01 to .960 (West 1999); MISS. CODE ANN. § 79-29-101 TO -112 (1999); MO. ANN. STAT. § 347.010 to .740 (West 1997); MONT. CODE ANN. § 35-8-101 to -1307 (1998); NEB. REV. STAT. ANN. § 21-2601 to -2653 (Michie 1999); NEV. REV. STAT. ANN. §§ 80.010–86.571 (Michie 1999); N.H. REV. STAT. ANN. § 304-C:1 to -C:81 (1999); N.J. STAT. ANN. § 42:2B-1 to 2B-69 (West 1999); N.M. STAT. ANN. § 53-19-1 to -73 (Michie 1998); N.Y. LTD. LIAB. CO. LAW §§ 101–1402 (McKinney 2000); N.C. GEN. STAT. § 57C-1-01 to -10-07 (1999); N.D. CENT. CODE § 10-32-01 to -155 (1999); OHIO REV. CODE ANN. § 1705.01 to .58 (Anderson 1998); OKLA. STAT. tit. 18, §§ 2000–2056 (1998); OR. REV. STAT. § 63.001–.955 (1997); PA. CONS. STAT. §§ 8901–8996 (1998); R.I. GEN. LAWS § 7-16-1 to -75 (1998); S.C. CODE ANN. § 33-43-101 to -1409 (Law. Co-op. 1998); S.D. CODIFIED LAWS § 47-34-1 to -59 (Michie 1999); TENN. CODE ANN. § 48-201-101 to 248-606 (1999); TEX. BUS. CORP. ACT ANN. art. 1528n, §§ 1.01–11.07 (Vernon 1999); UTAH CODE ANN. § 48-2B-101 to -157 (2000); VT. STAT. ANN. tit. 11, §§ 3001–3162 (1999); VA. CODE ANN. § 13.1-1000 to -1073 (Michie 1999); WASH. REV. CODE § 25.15.010–.285 (1999); W. VA. CODE § 31-1A-1 to -69 (1999); WIS. STAT. ANN. § 183.0102–.1305 (West 1997); WYO. STAT. ANN. § 17-15-101 to -144 (Michie 2000).

interest upon withdrawal, unless otherwise provided in the agreement.[4] Originally, most LLC statutes provided that the LLC member could withdraw and obtain the fair market value of his or her LLC interest, less damages caused by the withdrawal, unless the agreement provided to the contrary.[5] The new restrictions on the LLC member's withdrawal and distribution rights are designed to enhance the limited liability company as an estate and gift tax-planning vehicle. The driving force behind the reforms is to facilitate estate and gift tax valuation discounts for minority interests in family-owned limited liability companies.[6] Under an excruciatingly complex web of estate and gift tax rules applicable to family limited partnerships and other family-owned entities, the Internal Revenue Service recognizes discounts for restrictions on the ability to transfer or liquidate a minority interest in the entity only if the restrictions are rooted in state law, rather than in a partnership or operating agreement that has been drafted for the family.[7]

---

[4] ALA. CODE § 10-12-30 (1998); ARIZ. REV. STAT. § 29-707 (1999); ALASKA STAT. § 10.50.185 (Michie 1998); CAL. CORP. CODE § 17252 (West 1999); COLO. REV. STAT. ANN. § 7-80-603 (West 1999); DEL. CODE ANN. tit. 6, § 18-603 (1998); FLA. STAT. ch. 608.427 (1999); GA. CODE ANN. § 14-11-601 (Harrison 1998); OHIO REV. CODE ANN. § 1705.16 (Anderson 1998); R.I. GEN. LAWS § 7-16-29 (1998); VA. CODE ANN. § 13.1-1032 (Michie 1999); *see also* Laurel Wheeling Farrar & Susan Pace Hamill, *Dissociation from Alabama Limited Liability Companies in the Post Check-The-Box Era*, 49 ALA. L. REV. 909 (1998) (providing an excellent overview of the incipient trend to eliminate default buy-out rights in Alabama and elsewhere, the potential adverse impact upon minority owners of small businesses, and a policy recommendation that Congress reconsider the estate and gift tax valuation rules that are the driving force behind this trend).

[5] Bishop, *supra* note 3, at 261 (indicating that in most respects the acts adopt the partnership rather than the corporate paradigm regarding the effect of a member dissociation on that member's right to be bought out by the company and to cause a liquidation); *see also* Sandra K. Miller, *What Remedies Should Be Made Available to the Dissatisfied Participant in a Limited Liability Company?*, 44 AM. L. REV. 465, 505 (1994) (indicating that several states permitted members to withdraw at will).

[6] *See* I.R.C. §§ 2701–2704 (2000) (providing special valuation rules in the case of transfers among family members); Thomas I. Hausman, *Family Limited Partnerships*, 98 TAX NOTES 2, 2–72 (1998) (discussing the evolution of family limited partnerships developed to reduce estate and gift tax burdens and the implications of using LLCs to hold family businesses or property); Joseph M. Mona, *Advantages of Using a Limited Liability Company in an Estate Plan*, 25 EST. PLANNING 167, 167–71 (May 1998) (analyzing the estate and gift tax issues presented in using an LLC to own a family business).

[7] A common estate planning technique has been to have the owner of a family business or family investments place the business and/or assets into a limited partnership and transfer all but a minority interest in the limited partnership to his or her children. Russel Standaland, Note, *Valuation Discounts After* Estate of Novell v. Commissioner*: A Clear Formula for Reducing Estate Taxes*, 30 GOLDEN GATE U. L. REV. 679, 684–85 (2000). Upon death, the decedent's estate would claim discounts in the valuation of the retained partnership interest because of the minority nature of the interest, which could not be sold or liquidated. To enhance the discounts for minority ownership and lack of marketability, tax planners have developed creative partnership agreements that place severe restrictions on the taxpayer's ability to sell, redeem, or otherwise liquidate his or her minority partnership interest. *See id.* at 687. To combat the exploitation of minority discounts and discounts for lack of marketability by owners of family businesses and investments, Congress adopted I.R.C. § 2704(b) in 1990. The Omnibus Budget Reconciliation Act of 1990, I.R.C. 2701–2704 (1999). Section 2704(b) provides that if there is a transfer of an interest among

The purpose of this Article is to provide a critique of the present movement to eliminate buy-out rights of limited liability company members in order to achieve estate tax-related objectives. It refers to the reforms in buy-out rights as modifications of "default exit rights" since under the reforms, buy-out rights are eliminated or the withdrawal of a member is prohibited unless the parties have an agreement to the contrary. The Article suggests that changes in fundamental rights such as default exit rights *should not* be driven by purely tax considerations.

Parts I and II of the Article provide an overview of business entity exit rights and argue against their elimination. Part III recommends the adoption of explicit minority protections including a prohibition on unreasonable reductions in fiduciary duties or standards of care, and a mechanism for seeking judicial dissolution in the event of deadlock or specifically enumerated types of majority misconduct. Part IV discusses alternative strategies for handling exit rights in business entities.

The Article emphasizes the continued importance of giving the LLC member a right to withdraw from the LLC and receive the fair market value of his or her interest in the absence of an agreement to the contrary in light of: (1) the illiquidity of an investment in a private firm; (2) the intended use of the LLC as a vehicle for the informal conduct of a wide variety of business ventures; (3) the potential lack of other built-in statutory protections against foul play and abusive behavior; and (4) the considerable uncertainties surrounding both the duty of loyalty and standard of care for members and managers, and the mechanisms for asserting a breach of such duties.[8]

Many LLC statutes were drafted to avoid formalities in the conduct of business,[9] and as a result may lack many of the statutory rules and re-

---

family members, and the transferor and his family control the organization, the restrictions on the transferor's liquidation rights will be entirely ignored for valuation purposes to the extent that they are *more* restrictive than the limitations that would apply under state law in the absence of an express private agreement. I.R.C. § 2701(b). Thus, there is now a tax incentive to create restrictive state law provisions regarding the ability to sell, redeem, or otherwise liquidate an ownership interest in a business entity. *See* Robert R. Keatinge, *Universal Business Organization Legislation: Will it Happen? Why and When*, 23 DEL. J. CORP. L. 29, 52–62 (1998) (discussing the estate and gift tax provisions and their impact on the dissociation rights of members of partnerships, corporations and LLCs).

[8] *See* Sandra K. Miller, *Minority Shareholder Oppression in the Private Company in the European Community: A Comparative Analysis of the German, United Kingdom, and French "Close Corporation Problem,"* 30 CORNELL INT'L L.J. 381, 382 (1997) (discussing the universal nature of the vulnerability of minority investors in private businesses); Sandra K. Miller, *What Standards of Conduct Should Apply to Members and Managers of Limited Liability Companies?*, 68 ST. JOHN'S L. REV. 21 (1994) (discussing the development of a paradigm for applying fiduciary duties and a standard of care in the context of the LLC); Sandra K. Miller, *What Remedies Should Be Made Available to the Dissatisfied Participant in a Limited Liability Company?*, 44 AM. U. L. REV. 465 (1994) (recommending the development of equitable remedies for LLC members who may be squeezed out of the company by majority owners).

[9] *See* CAL. CORP. CODE § 17059 (West 1999) (indicating that members have the power to adopt an operating agreement, but avoiding a mandatory requirement that an operating

quirements that protect minority owners in corporations.[10] Not all LLC statutes provide for dissenters' rights in the case of certain mergers or acquisitions,[11] and relatively few provide an equitable remedy for a dissolution or buy-out in the event of certain illegal or fraudulent acts or other misconduct.[12] LLC statutes do not typically contain corporate-style notice provisions and other protections commonly contained in corporate statutes.[13] Without either the liquidity found in a general partnership or the protective remedies found in the corporation, a minority LLC owner without a strong bargaining position and a favorably negotiated operating agreement may be locked into a hybrid entity offering the worst, rather than the best, of the partnership and corporate worlds.

This Article suggests that prior to eliminating default exit rights to achieve estate tax goals for family-owned businesses and investments, legislators should consider the policy goals of the state's LLC statute, the

---

agreement be executed); D.C. CODE ANN. § 29-1302 (1998) (noting that the LLC is formed by filing articles of organization); *id.* § 29-1318 (recognizing that entering into an operating agreement is optional); FLA. STAT. ANN. § 608.423 (West 1999) (indicating that members may enter into an operating agreement, which need not be in writing). Some statutes, however, require that certain records be kept. *See id.* § 608.4101 (setting minimum standards for keeping records, including records of the amount of cash and property contributed).

[10] *See* Larry E. Ribstein, *The Deregulation of Limited Liability and the Death of Partnership*, 70 WASH. U. L.Q. 417, 417–38 (1992) (providing insight into the history of the LLC and its goal of combining the limited liability offered by corporations with the flexibility available to partnerships).

[11] For a sample of statutes that do provide for dissenters' rights, see CAL. CORP. CODE §§ 17600–17613 (West 1999) (providing for dissenters' rights with regard to certain reorganizations or mergers of LLCs); FLA. STAT. ANN. § 608.4381(4)(d) (West 1999) (referring to offers required in connection with dissenters' rights); N.Y. BUS. CORP. LAW § 1005 (McKinney 1994) (providing for payments to dissenting members in the case of certain mergers or consolidations); OHIO REV. CODE ANN. § 1705.40 (Anderson 1998) (outlining members' entitlement to relief as dissenting members). *But see* DEL. CODE ANN. tit. 6, § 18-210 (1998) (providing that an LLC agreement or merger agreement may provide appraisal rights, but the statute fails to provide appraisal rights in the absence of such contracts).

[12] For a sample of statutes that do provide equitable relief, see ARIZ. REV. STAT. ANN. § 29-785 (West 2000) (providing for involuntary dissolution in certain cases involving deadlock, illegal or fraudulent conduct, or the wasting, misapplication, or diversion of substantial assets); COLO. REV. STAT. ANN. § 7-80-808 (West 1999) (providing for involuntary dissolution if the LLC exceeds or abuses its authority); CAL. CORP. CODE § 17351 (West 1999) (permitting dissolution when it is necessary for the protection of the rights and interests of the complaining members, in the event of deadlock or internal dissension, or where those in control have been guilty of or have knowingly countenanced persistent and pervasive fraud, mismanagement, or abuse of authority); FLA. STAT. ANN. § 608.441 (West 1999) (authorizing a circuit court to dissolve an LLC if the managers or members are deadlocked, they are unable to break the deadlock, and irreparable injury will result or the LLC's assets are being misappropriated or wasted).

[13] *See* DEL. CODE ANN. tit. 6, § 18-302 (1998) (indicating that unless otherwise provided in the operating agreement, on any matter that is to be voted upon the members may take action without a meeting, without prior notice, and without a vote if consent in writing setting forth the action taken is signed by members having not less than the minimum number of votes that would be necessary to authorize the action).

availability of other business entities to achieve estate and gift tax savings, and the state's practical experience with LLCs.

Most importantly, legislators should analyze whether the LLC statute contains mechanisms *other than* default buy-out rights that can protect minority LLC owners from fraudulent and other opportunistic misconduct by majority owners. Does the particular LLC statute articulate fiduciary duties or standards of care to guide members' conduct or establish non-waivable minimum standards for the duty of loyalty? Does the LLC statute provide remedies in the event of deadlock or dispute? What are the provisions concerning the potential ability to set salaries, amend the operating agreement, approve sales of assets, and vote upon mergers or the expulsion of members? Does the LLC statute provide a clearly defined process for pursuing grievances with the LLC or with other LLC members such as a suit for an accounting or other equitable remedy? This Article maintains that the elimination of default exit rights poses a threat to minority owners in states whose LLC statutes lack other features that are designed to protect minority owners against majority squeeze-outs and other majority misconduct.

The Article recommends that states re-visit their business entity statutes as a whole, consider how all of their business entity statutes interface, and in the process, carefully consider the Uniform Limited Liability Company Act ("ULLCA").[14] The ULLCA provides for at-will withdrawals for LLCs without a term, as well as restricted withdrawals for LLCs with a specific operating term.[15] While the dichotomy between at-will and term LLCs may not be acceptable to those states determined to eliminate default exit rights, ULLCA contains several other protections that could be helpful in all states, especially where default exit rights have been eliminated. ULLCA's provisions restricting the ability to contractually reduce the duty of loyalty or the standard of care,[16] the process for authorizing extraordinary events such as changes in the operating agreement and mergers,[17] and the remedies in the case of internal dissension[18] serve as useful models.

In light of the intended use of the LLC as a vehicle for a wide variety of business ventures, the infancy of the case law interpreting the duty

---

[14] UNIF. LTD. LIAB. CO. ACT (1996).

[15] *Id.* § 701.

[16] *Id.* §§ 103, 409 (prohibiting the operating agreement from unreasonably reducing the duty of loyalty, standard of care, or access to information and records and providing for express fiduciary duties and a standard of care limited to refraining from grossly negligent or reckless conduct, intentional misconduct, or a knowing violation of law).

[17] *Id.* § 404(c) (requiring unanimous agreement for certain important events, including changes in the operating agreement, authorizations or ratifications of acts that would violate or modify the duty of loyalty, changes in the articles of organization, mergers, and the sale or other disposal of goodwill).

[18] *Id.* § 801 (providing for dissolution in the event of certain misconduct by managers or members in control).

of loyalty and standard of care for members and managers, and the mechanisms for asserting a breach of these duties, this Article strongly advocates the enactment of at least two statutory protections for minority LLC owners: (1) a prohibition on contractual provisions that unreasonably restrict or reduce fiduciary duties and the standard of care; and (2) the right to seek a dissolution, or a buy-out in lieu of a dissolution, in the case of deadlock or where the managers or members in control of the company have engaged in specific types of misconduct delineated in the statute, i.e., illegal, fraudulent, oppressive or unfairly prejudicial conduct. Such protections have been included in the ULLCA.[19]

I. EXIT RIGHTS IN LLCS AND OTHER BUSINESS ENTITIES: THE NEW WAVE OF REVISIONS

The right of an investor to withdraw from a business and recover his or her capital investment, as well as the consequences of such withdrawal to the firm, have traditionally varied depending upon the business entity. In order to comprehensively evaluate the consequences of eliminating default LLC exit rights in the broader context of each state's business entities, it is necessary to gain an understanding of an investor's exit rights in the other business entities with which the LLC co-exists.

Overall, the corporation has stood at one end of the business entity spectrum with regard to exit rights and their consequences for the firm, and the partnership has stood at the other. In general, although stock of a public company could usually be *sold*, corporate shareholders of both public and private corporations have traditionally lacked a general statutory right to *withdraw* from the corporation and receive the fair market value of their stock.[20] In contrast, until the latest round of reforms and draft proposals, partners in general partnerships, in some cases general partners in limited partnerships, and members of limited liability companies have had a default right to withdraw from the firm and get paid the

---

[19] *See id.* § 103 (providing that the operating agreement may not unreasonably restrict the duty of loyalty or the duty of care); *id.* § 801 (enumerating the occurrence of events upon which the LLC is dissolved, including where "the managers or members in control of the company have acted, are acting, or will act in a manner that is illegal, oppressive, fraudulent, or unfairly prejudicial to the petitioner").

[20] *See* 2 F. HODGE O'NEAL & ROBERT THOMPSON, O'NEAL'S CLOSE CORPORATIONS § 9.02 (3d ed. 2000) (discussing the fact that shareholders of a private corporation possess an illiquid investment and frequently must litigate to obtain relief from misconduct by the majority); Harry J. Haynsworth, *The Effectiveness of Involuntary Dissolution Suits as a Remedy for Close Corporation Dissension*, 35 CLEV. ST. L. REV. 25 (1987) (discussing generally the dissension in private corporations that has led to litigation as a result of the illiquidity of the shareholders' investments and containing an empirical study of the litigation); J.A.C. Hetherington & Michael P. Dooley, *Illiquidity and Exploitation: A Proposed Statutory Solution to the Remaining Close Corporation Problem*, 63 VA. L. REV. 1, 2–3 (1977) (emphasizing the differences in the liquidity of a partnership interest and a shareholder's interest in a private corporation).

fair market value of their ownership interests, unless an ex ante agreement provided to the contrary, subject of course to a variety of special rules and restrictions as well as potential offsets for wrongful conduct that violated the applicable agreement among the parties.[21] General partners who wrongfully withdraw would not be entitled to receive immediate payment.[22]

The default buy-out right was first recognized in the Uniform Partnership Act ("UPA"), which gave a partner the right to dissolve the partnership at any time and get paid the fair market value of his or her interest, subject to special rules for withdrawals in violation of an agreement.[23] Although the Revised Uniform Partnership Act ("RUPA") overhauled the UPA's dissolution rules by introducing the concept of partnership "dissociation," the revised statute retains the partner's default buy-out rights in the case of the "at-will" partnership—a partnership formed without a specified term or time period.[24] Default buy-out rights similar to those contained in RUPA are mirrored in the ULLCA.[25] Finally, the existing Revised Uniform Limited Partnership Act contains a default rule permitting a general partner to withdraw from the limited partnership.[26] However, the proposed revisions to the RULPA, the Re-RULPA, would

---

[21] Bishop, *supra* note 3, at 260–61 (1995).

[22] *See id.* at 280.

[23] *See* UNIF. P'SHIP ACT § 31(1)(b) (1914) (discussing the causes of dissolution, including dissolution by express will of any partner when no definite term or particular undertaking is expressed).

[24] *See* REVISED UNIF. P'SHIP ACT §§ 601, 801 (1997). The UPA was initially amended in 1994 (to become RUPA) and significant modifications included: changing the dissolution rules, utilizing the entity rather than aggregate model, recognizing the primacy of the partnership agreement by providing a series of default rules that become operative only in the absence of an agreement to the contrary and adopting express obligations of loyalty, due care, and good faith. RUPA was amended in 1997 to incorporate provisions for limited liability in the case of a limited liability partnership. *See* NAT'L CONFERENCE OF COMM'RS ON UNIF. STATE LAWS, A FEW FACTS ABOUT THE UNIFORM PARTNERSHIP ACT (1994) (1997) [hereinafter A FEW FACTS ABOUT THE UPA], *at* http//www.nccusl.org/uniformact_factsheets/uniformacts-fs-upa9497.htm. (last visited May 22, 2001). The fact sheet indicates that Connecti-cut, West Virginia, and Wyoming have adopted RUPA with the 1994 amendments. Jurisdictions that have adopted RUPA with the 1997 amendments (the current form of RUPA) include Alabama, Alaska, Arizona, Arkansas, California, Colorado, Delaware, District of Columbia, Florida, Hawaii, Idaho, Iowa, Kansas, Maryland, Minnesota, Montana, Nebraska, New Jersey, New Mexico, North Dakota, Oklahoma, Oregon, Puerto Rico, Texas, U.S. Virgin Islands, Vermont, Virginia, and Washington. *Id.*; *see also* 6 U.L.A. 153 (Supp. 2000) (listing states that enacted the 1914 Act and states that have recently adopted the revisions).

[25] *See* UNIF. LTD. LIAB. CO. ACT § 701(a), (f) (1996) (providing rules for the purchase of a distributional interest of an at-will company for its fair market value and for the purchase of a term company as of the date of the expiration of the term, subject to offsets for damages for wrongful dissociation). Several jurisdictions have already adopted the ULLCA, including Alabama, Hawaii, Illinois, Montana, South Carolina, South Dakota, U.S. Virgin Islands, Vermont, and West Virginia. *See* NAT'L CONFERENCE OF COMM'RS ON UNIF. STATE LAWS, A FEW FACTS ABOUT THE UNIFORM LIMITED LIABILITY COMPANY ACT, *at* http//www.nccusl.org/uniformact_factsheets/uniformacts-fs-ullca.htm (last visited Feb. 18, 2001).

[26] *See* UNIF. LTD. P'SHIP ACT OF (1976) WITH 1985 AMENDMENTS § 604 (1985).

eliminate the right of a person to receive any distribution on account of dissociation, largely to achieve estate tax-related benefits.[27]

In light of the growing trend to eliminate default exit rights under several state LLC acts and in the Re-RULPA, the investor's buy-out rights and dissolution rules in each business entity should be carefully analyzed.

A. Exit Rights in the Corporation: The Concept of Continuity of Life

Historically, the corporation has been recognized as a separate legal entity distinct from its owners.[28] The life of a corporation remains largely unaffected by changes in the identity of its shareholders.[29] A corporate redemption or a sale of stock by a shareholder has never triggered a corporate dissolution. The corporate capacity to remain unchanged in spite of changes in ownership has long been referred to as the corporate characteristic of continuity of life.[30] Aside from the relatively recent advent of dissenters' rights,[31] and the narrow dissolution and/or buy-out rights recognized in the event of shareholder deadlock or illegal or oppressive conduct,[32] the corporate shareholder of a private company is very much locked into his or her investment absent a buy-sell agreement or other agreement that provides for a shareholder buy-out under specified triggering events.

B. Partner Withdrawal Rights and Dissolution Events in the General Partnership

Under the UPA, a general partnership dissolves through the expiration of a term or particular undertaking, the express will of the partners, the expulsion of a member, events that make it unlawful to conduct business, court decree, death, incapacity, or bankruptcy of a partner.[33] Thus, a

---

[27] *See* PROPOSED REVISIONS OF UNIF. LTD. P'SHIP ACT (1976) WITH 1985 AMENDMENTS § 505 (completion targeted for 2001) (on file with author).

[28] *See* MODEL BUS. CORP. ACT § 14.02 (1999) (providing for dissolution upon a majority vote of a quorum of the voting groups); *see also* 1 JAMES D. COX ET AL., CORPORATIONS § 2.3–2.7 (2000) (discussing the evolution of the American corporation); O'NEAL & THOMPSON, *supra* note 20, § 8.18 (2000) (discussing the close corporation as a separate entity, subject to the veil-piercing exception); 1 WILLIAM H. PAINTER, PAINTER ON CLOSE CORPORATIONS § 1.3 (3d ed. 1999) (discussing the legal treatment of a corporation as a separate entity).

[29] Continuity of life has been regarded as the hallmark of a corporation. *See* Morrissey v. Comm'r, 296 U.S. 344, 358 (1935); U.S. v. Kintner, 216 F.2d 418, 422 (9th Cir. 1954).

[30] *See* Larson v. Comm'r, 66 T.C. 159, 173 (1976) (comparing partnerships and corporations).

[31] *See* MODEL BUS. CORP. ACT § 13.02 (1999) (granting to shareholders appraisal rights and rights to obtain payment of the fair value of their stock when certain corporate actions occur, including merger, share exchange, and disposition of corporate assets).

[32] *See id.* § 14.30(2).

[33] UNIF. P'SHIP ACT §§ 31–32 (1914).

general partner can technically withdraw at-will, although if the withdrawal contravenes the partnership agreement,[34] the wrongfully withdrawing partner could be subject to several significant punitive measures.[35] Under the UPA, dissolution is defined as "the change in the relation of the partners caused by any partner ceasing to be associated in the carrying on as distinguished from the winding up of the business."[36] Simply put, under the UPA, whenever there is a change in the constituent members of the partnership, the partnership dissolves.[37]

If an event triggering dissolution occurs, the partnership business can either be sold at a judicial sale or the business can be continued by a new, successor partnership.[38] If the withdrawal was wrongful, the goodwill of the partnership is not used in determining the buy-out price.[39] Also, the wrongfully withdrawing partner has no right to bid for the business and his or her buy-out price is reduced by the damages resulting from the wrongful withdrawal.[40]

The unilateral right to dissolve a partnership under the UPA has been subject to considerable criticism.[41] In fact, the American Bar Association submitted no less than sixty-six changes to the UPA's dissolution provisions.[42] Commentators complained that the full liquidation right injected

---

[34] *Id.* at § 20 (defining dissolution as the change in the relation of the partners caused by any partner ceasing to be associated in the carrying on of the business); *id.* § 31(1)(b) (indicating that dissolution is caused without violation of the agreement by express will of any partner when no definite term or particular undertaking is specified); *id.* § 31(2) (indicating that dissolution is caused in contravention of the agreement where the circumstances do not permit a dissolution).

[35] *See id.* § 31(1)(b), (1)(c), (2) (providing that dissolution is caused without violation of the agreement by express will of a partner—when no particular term or undertaking is specified—and is caused in contravention of the agreement where the circumstances do not permit dissolution); *see* Farrar & Hamill, *supra* note 4, at 916, 917; Robert W. Hillman, *The Dissatisfied Participant in the Solvent Business Venture: A Consideration of the Relative Permanence of Partnerships and Close Corporations*, 67 MINN. L. REV. 1, 11–14 (1982) (discussing the consequences of wrongful withdrawals from the general partnership).

[36] UNIF. P'SHIP ACT § 29 (1914).

[37] These dissolution rules are premised on both the aggregate theory of the partnership and the principle of *deluctus personae*, or choice of the person. The aggregate theory of the partnership regards the partnership as a conduit for the partners. The concept of *deluctus personae* recognizes that an investor should have control over the identity of his or her partners. Both concepts have been regarded as leading to the rule that a partner has the power to dissolve the partnership even in violation of the partnership agreement. *See* Donald J. Weidner, *A Perspective To Reconsider Partnership Law*, 16 FLA. ST. U. L. REV. 1, 14 (1988) (explaining that whenever the cast of characters changes the partnership dissolves under the aggregate concept of the partnership on which the UPA is based); Donald J. Weidner, *Three Policy Decisions Animate Revision of Uniform Partnership Act*, 46 BUS. LAW. 427, 428 (1991) (discussing the theories on which the UPA was based).

[38] *See* Farrar & Hamill, *supra* note 4, at 915, 916.

[39] UNIF. P'SHIP ACT § 38(2)(c)(II) (1914).

[40] *See id.*

[41] *See generally* Weidner, *A Perspective to Reconsider Partnership Law*, *supra* note 37 (indicating that dissolution and its consequences remain troublesome areas).

[42] *See id.* at 14.

instability into the partnership relationship and was so frequently abused that courts struggled to find oral or implicit agreements among partners to continue the partnership.[43] The threat of a dissolution can be a powerful weapon in the hands of dissatisfied participants in a partnership and creates the potential for wealthy partners to squeeze-out their less well-to-do partners by entering unfairly low bids for the business.[44]

In response to widespread criticism, the RUPA limits the circumstances that can cause partnership dissolutions.[45] Although the UPA continues to be relevant in some states, a growing number of states have revised their partnership statutes in accordance with RUPA and its new dissolution rules.[46]

RUPA creates two types of partnerships: the "at-will partnership" which has no stated term, and a "term partnership" which does have a stated term.[47] The consequences of a partner's separation from the partnership differ depending on whether it is an "at-will" or "term" partnership.

If a partner withdraws from an "at-will" partnership, the partner is considered to be "dissociated" from the partnership, and the partnership will be dissolved unless the remaining partners, including the rightfully withdrawing partner, agree to waive the dissolution.[48] In an at-will partnership, the dissociation is "wrongful" if it violates an express provision of the partnership agreement.[49] In effect, a withdrawing partner of an at-will partnership without express contractual restrictions on withdrawals may demand a judicial sale of the partnership's assets by withdrawing.[50]

---

[43] *See id.* at 20.

[44] *See* Farrar & Hamill, *supra* note 4, at 916.

[45] *See* REVISED UNIF. P'SHIP ACT (1997).

[46] For a list of jurisdictions that have adopted RUPA, see *supra* note 24.

[47] *See* REVISED UNIF. P'SHIP ACT § 101(8) (1997) (indicating that a "partnership at-will means a partnership in which the partners have not agreed to remain partners until the expiration of a definite term or the completion of a particular undertaking").

[48] *See id.* § 801(1) (indicating that in an at-will partnership, the partnership is dissolved and its business must be wound up upon notice of the partner's will to withdraw in situations *other than* those identified in section 601(2) to (10), which include an event agreed to in the partnership, the expulsion of the partner in certain cases, a judicial determination, bankruptcy, death, or termination); *id.* § 802 (providing that after the dissolution and before the winding up of the business, all of the partners, including the dissociating partner unless he or she is a wrongfully dissociating partner, may waive the right to have the partnership's business wound up and the partnership terminated). Whether the dissociation is rightful or wrongful presumably will be at the heart of litigation in partnership disputes under RUPA.

[49] *See id.* § 602(b)(1). Under section 602(b)(2) of RUPA, in the case of a term partnership, a dissociation that occurs before the expiration of a term or particular undertaking is wrongful if a partner withdraws before the expiration of the term or undertaking, is expelled by certain specified judicial determinations of wrongful conduct, becomes bankrupt, or is wrongfully terminated.

[50] *See id.* § 801(1) & cmt. 3 (indicating that section 801 retains two basic rules from the UPA: the rule that any member of an at-will partnership has the right to force a liquidation, and the rule that partners who wish to continue the business of a term partnership cannot be forced to liquidate the business).

If the partner's separation from the at-will partnership occurs because the partner dies, becomes bankrupt, is expelled, or becomes incapacitated, the partner is also treated as being "dissociated" from the partnership.[51] However, the partnership itself does not technically dissolve in such instances of dissociation.[52] Where a dissociation event occurs and the partnership does not dissolve, regardless of whether the dissociation is considered to be rightful or wrongful, the dissociated partner is entitled to be paid a price equal to the greater of the liquidation value of the partnership or the value of the partnership as a going concern.[53] Damages, however, may be assessed and the purchase price will be reduced if the dissociation is considered to be wrongful.[54]

In a "term" partnership, death, bankruptcy, withdrawal, or expulsion are also treated as "dissociation" events.[55] Although a voluntary withdrawal causes a dissolution in the case of an at-will partnership, it does not necessarily lead to the dissolution of a term partnership. A dissociation event of a term partnership will only lead to a dissolution of the partnership if certain specified events, such as death or bankruptcy, have occurred and at least half of the remaining partners agree to wind up the partnership.[56] Another difference between the consequences of at-will and term partnership dissociations concerns the dissociating partner's buy-out rights in the event that the partners vote to continue the firm after a dissolution event. If there is a dissociation and the partnership continues, the dissociated partner of a term partnership does not typically have an immediate right to be paid the fair market value of his or her partnership interest. Dissociation from a term partnership before the term has expired will usually be defined as wrongful and the withdrawing member must wait for his or her buy-out until the term of the partnership has expired.[57] As in the case of the at-will partnership, the buy-out price of a

---

[51] *See id.* § 601(1)–(10) (enumerating the events causing a partner's dissociation).

[52] *Id.* §§ 601(1)–(10), 801(1).

[53] *Id.* § 701(a)–(b).

[54] *See id.* §§ 602, 701(c). A partner's dissociation is wrongful if it breaches an express provision of the partnership agreement.

[55] *See id.* §§ 601–602.

[56] *See id.* § 801 (indicating that for a term partnership a dissolution and winding up occur only if at least half of the remaining partners express the will to wind up and if the dissociation event is death or the other dissociation events listed in section 601(6) to (10), including bankruptcy, an assignment for the benefit of creditors, the appointment of a trustee, guardian or conservator, distribution of all of an estate's or trust's interests, or certain terminations of a partner).

[57] *See id.* § 602(b) (defining wrongful dissociation to include a dissociation in violation of an express provision of the partnership agreement or, in the case of a term partnership, a dissociation before the expiration of the term where the partner withdraws by express will, unless the withdrawal follows another partner's dissociation by death or otherwise in the circumstances enumerated in section 601(6) to (10)); *id.* § 701(h) (indicating that a partner who wrongfully withdraws before the expiration of a term or undertaking is not entitled to payment until the expiration of the term or undertaking, unless the partner establishes to the satisfaction of the court that earlier payment will not cause undue hardship to the business).

partner who has wrongfully dissociated from the partnership is reduced by damages caused by the wrongful dissociation.[58] Unlike UPA, however, under RUPA a wrongfully dissociating partner's buy-out price will include any goodwill of the partnership unless the partner's withdrawal has damaged the goodwill of the business.[59]

C. Withdrawal Rights and Dissolutions of Limited Partnerships

Historically, limited partnerships have also differed from corporations with regard to dissolution events. Under the RULPA, the limited partnership dissolves at the time specified in the certificate of limited partnership, the happening of events specified in the partnership agreement, the written consent of all partners, the entry of a judicial dissolution decree, or the withdrawal of a general partner, unless there is one other general partner and the partnership agreement permits the continuation of the partnership.[60] In any case, however, the dissolution can be avoided if, within ninety days of a withdrawal, all partners agree in writing to continue the business.[61]

Under the RULPA, withdrawal rights differ depending on the partner's status either as a general partner or a limited partner, and depending on the terms of the partnership agreement.[62] A general partner can withdraw at any time by giving written notice to the other partners.[63] The general partner is entitled to receive the fair market value of the partnership interest within a reasonable amount of time following the withdrawal of the partner, subject to an offset for damages.[64] Of course, a partnership agreement can modify these payment provisions and make it wrongful to withdraw prior to the expiration of the term.[65]

In contrast, a limited partner may be locked into a limited partnership and is likely to lack the broad withdrawal power of the general partner.[66] A limited partner may withdraw upon the happening of events specified in the written partnership agreement.[67] If the partnership agreement does not specify the time or events of withdrawal or does not contain a definite term for dissolution and winding up, the limited partner

---

[58] *See id.* § 701(c).
[59] *See id.* § 602 & cmt. 3.
[60] UNIF. LTD. P'SHIP ACT OF (1976) WITH 1985 AMENDMENTS § 801 (1985).
[61] *Id.*
[62] *Id.* §§ 602, 603.
[63] *Id.* § 602.
[64] *Id.* §§ 602, 604.
[65] *See* Bishop, *supra* note 3, at 289 (discussing exit rights in partnerships).
[66] *See id.* at 290 (discussing the differences between the withdrawal powers of general and limited partners).
[67] *Id.* § 603.

may withdraw upon not less than six months' prior written notice to each general partner.[68]

Since limited partnerships are formed by filing a certificate of limited partnership with the secretary of state, and the certificate must include the latest date for dissolution, most limited partnership agreements repeat the term stated in the certificate.[69] Thus, most limited partners will be subject to an agreement with a definite term and will not have the power to withdraw early.[70] In a minority of cases, it is possible that the partnership agreement will contradict the term stated in the certificate, or will lack a stated term. In such an event, the limited partner would be able to receive a distribution of the fair value of the partnership interest within a reasonable time after the withdrawal.[71] Again, if the withdrawal is wrongful, damages may offset a distribution.[72]

As more fully discussed below, the Re-RULPA proposes to change the dissociation rules of the limited partnership by eliminating distributions upon departures.

D. The Evolution of Exit Rights and Dissolution Rules in the Limited Liability Company

Initially, most LLC statutes contained withdrawal rights and dissolution rules based on a partnership paradigm. The majority of LLC statutes provided that unless otherwise provided in the operating agreement, an LLC member had the power to withdraw from the company.[73] Most states provided that the withdrawing LLC member was entitled to receive the fair market value of his/her interest, reduced by any damages caused by a wrongful withdrawal.[74] Partnership-like contingent dissolution provisions were typically employed, requiring that the LLC dissolve upon the occurrence of stated events such as bankruptcy, death, or incompetency, unless all of the remaining members agreed to continue the LLC or a majority in interest agreed to continue.[75]

The LLCs' contingent dissolution provisions were designed to maximize the chances of partnership characterization under the IRS's previous mandatory business entity classification scheme.[76] Under these now obsolete rules, in order to be classified as a partnership for tax pur-

---

[68] *See* UNIF. LTD. P'SHIP ACT OF (1976) WITH 1985 AMENDMENTS § 603 (1985).
[69] *See* Bishop, *supra* note 3, at 293.
[70] *Id.*
[71] UNIF. LTD. P'SHIP ACT OF (1976) WITH 1985 AMENDMENTS § 604 (1985).
[72] *See id.* § 604; Bishop, *supra* note 3, at 293.
[73] *See* Bishop, *supra* note 3, at 297.
[74] *See id.*
[75] *See* Sandra K. Miller, *Increased Flexibility of Limited Liability Company Operating Agreements Raises Questions About Tax Treatment*, TAXES, Oct. 1994, at 622, 624 (discussing the pre-1997 tax classification regulations in the context of LLCs).
[76] Prior to 1997, the IRS employed a mandatory tax classification scheme.

poses, it was helpful and sometimes critical for the LLC to lack the corporate characteristic of continuity of life. Under prior IRS Regulation 301.7701-2(a)(3), an unincorporated entity would not be taxable as a corporation unless the organization possessed more *corporate* than *noncorporate* characteristics. The four operative factors were: continuity of life, centralization of management, limited liability, and free transferability of interests.[77] Under Treasury Regulations, the corporate characteristic of continuity of life was present if the death, insanity, bankruptcy, retirement, resignation, or expulsion of any member *would not* cause a dissolution of the organization.[78] Further, under regulation 301.7701-2(b)(1) continuity of life was absent in limited partnerships even if the dissolution could be avoided by the unanimous agreement, or by a majority in interest, of the remaining partners to continue the business.[79]

Prior to the enactment of the Check-the-Box Regulations, a majority of states gave LLC members an optional default power to withdraw from the LLC unless the operating agreement provided otherwise.[80] Under most statutes, dissociated members would receive the fair market value of their interest reduced by any damages for wrongful conduct.[81]

Initially, the IRS was slow to recognize the LLC, and legislators proceeded with caution.[82] The *raison d'être* of the LLC was to combine

---

[77] The IRS regulations codified the decision in *United States v. Kintner*, 216 F.2d 418 (9th Cir. 1954). In that case, a medical group was treated as a corporation rather than a partnership. The regulations were initially written to prevent taxpayers from incorporating to gain favorable employee benefits. Subsequently, when tax shelters became prevalent before the Tax Reform Act of 1986, the regulations facilitated the formation of partnerships used as tax shelters. *See* Miller, *Increased Flexibility of Limited Liability Company Operating Agreements Raises Questions About Tax Treatment*, *supra* note 75, at 622–23.

[78] *See* Treas. Reg. § 301.7701-2(b) (1) (1993).

[79] Treasury Regulation 301.7701-2(b) (1) was amended in 1993. Prior to 1993 it provided that, for limited partnerships, if the retirement, death, or insanity of a general partner causes a dissolution of the partnership, unless all remaining members agree to continue the partnership, continuity of life does not exist. *See* Miller, *Increased Flexibility of Limited Liability Company Operating Agreements Raises Questions About Tax Treatment*, *supra* note 75, at 624.

[80] *See* Bishop, *supra* note 3, at 297.

[81] *See id.*

[82] Joseph A. Rodriguez, *Wyoming Limited Liability Companies: Limited Liability and Taxation Concerns in Other Jurisdictions*, 27 LAND & WATER L. REV. 539, 544–45 (1992). In 1988, the IRS issued Revenue Ruling 88-76, 1988-2 C.B. 361, which stated that a Wyoming limited liability company, none of whose members or designated managers were personally liable for any debts of the company, was to be classified as a partnership for federal income tax purposes. In an effort to distinguish between a corporation and a partnership, the IRS focused on the following factors: (1) continuity of life; (2) centralization of management; (3) liability for corporate debts limited to corporate property; and (4) free transferability of interests. Applying the principles enunciated in *Larson v. Commissioner*, 66 T.C. 159 (1976), and Treasury Regulation 301.7701-2, if the entity possesses more corporate characteristics than non-corporate characteristics, then the entity will be taxed as a corporation rather than as a partnership. In Revenue Ruling 88-76, the IRS found that a Wyoming limited liability company lacked the corporate characteristic of continuity of life since the limited liability company is dissolved upon the death, retirement, resignation, expulsion, bankruptcy, or any other termination of a membership, unless the business is

in a single business entity the benefit of limited liability with the advantage of flow-through taxation as a partnership. Legislators sought to ensure that the LLC would lack the corporate characteristic of continuity of life.[83] The first wave of LLC statutes provided relatively little flexibility with regard to provisions concerning dissolution, transferability, and management. Eventually, some LLC statutes were liberalized to allow flexibility with regard to dissolution events, but contained a default rule that, in the absence of an agreement to the contrary, an event such as a member's death, bankruptcy, or incapacity would cause the firm to dissolve unless a *majority in interest* of the remaining members agreed to continue the entity.[84] Most LLC statutes tracked some or all of the language of IRS Regulation 301.7701-2(b)(2) in an effort to ensure the absence of the corporate characteristic of continuity of life.[85]

As practitioners gained more experience with LLCs, they began to express concern that unanimous approval to continue an LLC after a dissolution event of a member would prove impractical, disruptive, and unstable with regard to the underlying business. They became more aggressive in seeking partnership status where dissolution would occur in increasingly narrow circumstances. Partnership status was sought where the LLC dissolved only if a *single* dissolution event occurred with respect to a *single* LLC member. Practitioners also tried to obtain partnership status where LLC members entered a pre-formation agreement to continue the business following a dissolution event. The IRS became in-

---

continued by the consent of all the remaining members. Further, the Wyoming limited liability company lacked the corporate characteristic of free transferability of interest. A member's interest was transferable only with the consent of all remaining members. Thus, the failure to possess more corporate than non-corporate features resulted in classification for tax purposes as a partnership.

[83] *See* Susan P. Hamill, *The Limited Liability Company: A Possible Choice For Doing Business?*, 41 U. FLA. L. REV. 721 (1989) (discussing the major features of the limited liability company and its advantages); Miller, *What Standards of Conduct Should Apply to Members and Managers of Limited Liability Companies?*, *supra* note 8, at 23 (discussing the central features of the limited liability company); Larry E. Ribstein, *The Deregulation of Limited Liability and the Death of Partnership*, 70 WASH. U. L.Q. 417, 417–38 (1992).

[84] *See* Glensder Textile Co. v. Comm'r, 46 B.T.A. 176 (1942) (concluding that a limited partnership lacked continuity of life because upon the death, retirement, or incapacity of a general partner, the remaining general partners would have to agree to continue the partnership, and there was no assurance that they would do so). The court in Glensder observed that the continuity of a partnership was not analogous to the chartered life of a corporation, which continues regardless of the death or resignation of its directors or stockholders. *See* Miller, *Increased Flexibility of Limited Liability Company Operating Agreements Raises Questions About Tax Treatment*, *supra* note 75, at 624–25 (discussing the typical withdrawal rules that were in effect in LLC statutes prior to the enactment of the Check-the-Box Regulations).

[85] *See* I.R.S. Notice 95-14, 1995-1 CB 197 (announcing on April 3, 1995 the intention to amend the tax classification regulations through the development of an elective check-the-box system); Treas. Reg. §§ 301.7701-1, 301.7701-2, and 301.7701-3 (adopted on December 18, 1996 and effective January 1, 1997) (enacting the check-the-box elective tax classification regulations).

undated with private letter ruling requests, and eventually realized that its mandatory tax classification rules created an administrative nightmare.[86]

Finally, nearly twenty years after the recognition of the first LLC, the IRS jettisoned its mandatory tax classification scheme and enacted the so-called "Check-the-Box" regulations that automatically treat unincorporated business entities, including LLCs, as partnerships.[87]

E. Eliminating Exit Rights in LLCs: A New Tax-Driven Trend

Subsequent to the enactment of the Check-the-Box regulations, a number of LLC statutes were modified to eliminate the so-called contingent dissolution provisions. For example, the LLC statutes of Alabama,[88] Alaska,[89] Arizona,[90] California,[91] Colorado,[92] Delaware,[93] District of Columbia,[94]and Florida[95] no longer link dissolution to events such as death,

---

[86] Larry E. Ribstein, *The New Choice of Entity for Entrepreneurs*, 26 CAP. U. L. REV. 325, 329 (1997) (discussing the collapse of prior tax classification rules).

[87] *See* BITTKER & EUSTICE, *supra* note 2, § 2.202[3] (discussing the new regulations that presume that unincorporated domestic taxpayers are classified as partnerships for tax purposes).

[88] ALA. CODE § 10-12-37 (1998) (providing that events of dissolution occur upon the first of the following to occur: (1) events specified in the operating agreement; (2) written consent of all members to dissolve; (3) when there is no remaining member subject to some special exceptions; (4) when the LLC is not the successor in a merger or consolidation; or (5) when there is a judicial dissolution because it is not reasonably practical to carry on business).

[89] ALASKA STAT. § 10.50.400 (Michie 1998) (providing for dissolution upon the first of the following to occur: (1) the time or events specified in the operating agreement; (2) the consent of all the members to the dissolution; or (3) a decree by the superior court that it is impossible to carry on the purposes of the company).

[90] ARIZ. REV. STAT. ANN. § 29-71 (West 2000) (providing for dissolution upon the first of the following to occur: (1) events specified in the articles of organization or operating agreement; (2) written consent of more than one-half of the members and by one or more members who would be entitled to more than one-half of the assets; or (3) entry of an administrative dissolution for failing to comply with enumerated rules).

[91] CAL. CORP. CODE § 17350 (West 1999) (providing that an LLC will be dissolved and its affairs wound up upon the first of the following to occur: (1) the time specified in the articles of organization or the happening of events in the operating agreement or articles of organization; or (2) by vote of a majority in interest of the members, or a greater percentage specified in the articles of organization or the operating agreement).

[92] COLO. REV. STAT. § 7-80-801 (1999) (indicating that an LLC will be dissolved upon the unanimous written agreement of all members or at the time of events specified in the operating agreement).

[93] DEL. CODE ANN. tit. 6, § 18-801(a) (1999) (providing for dissolution upon the first of the following to occur: (1) the time specified in the agreement; (2) events specified in the agreement; (3) consent of members to dissolve (unless otherwise provided); (4) an absence of members or (5) issuance of a judicial decree of dissolution).

[94] D.C. CODE ANN. § 29-1347 (Lexis 2000 Supp.) (providing that the LLC is dissolved upon the first of the following, as indicated in the operating agreement, to occur: (1) upon unanimous consent; (2) the entry of a judicial dissolution, or (3) when the LLC has no more members for 90 consecutive days).

[95] FLA. STAT. ch. 608.441 (1999) (providing for dissolution upon the first of the following to occur: (1) the time specified in the articles of organization or operating agreement; (2) events listed in the articles of organization; (3) written consent of all members,

bankruptcy, or dissolution of a member. While there is considerable variation among the states' dissolution provisions, the common approach appears to provide for dissolution upon the earliest of: (1) the time specified in the written operating agreement; (2) the events stated in the articles of organization or the operating agreement; (3) the time of approval by all members (or in some cases less than all); or (4) a court decree that it is no longer practicable to carry on the business. The Delaware LLC statute presumes a perpetual existence unless the operating agreement provides to the contrary.[96] Of course, there remains considerable variation in the LLC dissolution provisions of different states.[97]

In addition to changing dissolution provisions, a growing number of states have also enacted restrictions on an LLC member's right to have the LLC or other LLC members purchase his or her interest upon withdrawal from the LLC. Initially, the most common approach was to give the LLC member the power to withdraw, unless the operating agreement provided to the contrary, and to give the dissociated member the fair value of his or her interest, less damages caused by the withdrawal.[98] To enhance the LLC's usefulness as an estate and gift-tax planning vehicle, a growing number of LLC statutes now eliminate the member's right to be paid a distribution of the fair market value of the interest upon withdrawal in the absence of an agreement expressly bestowing such buy-out rights.[99] Several other states do not expressly eliminate "distribution rights" as such, but contain other provisions that restrict an LLC member's right to withdraw and liquidate his or her LLC interest before dissolution of the company. For example, Delaware's LLC statute provides that a member may resign only at the time or upon the happening of an event in the LLC agreement and that, unless the LLC agreement provides otherwise, the member may not resign from the LLC prior to its dissolution and winding up.[100] The Florida, New York, and Oklahoma LLC stat-

---

unless otherwise provided in the articles of organization or the operating agreement; (4) if there are no members; or (5) the entry of a court decree that it is not reasonably practical to carry on business).

[96] *See* DEL. CODE ANN. tit. 6, § 18-801(a) (1999).

[97] GA. CODE ANN. § 14-11-602 (2000) (providing for dissolution upon the first of the following to occur: (1) the time specified in the written operating agreement; (2) events in the articles of organization or the operating agreement; or (3) the time of approval by all members; unless otherwise provided in the articles of organization or the operating agreement after 90 days of a dissociation event of any member; unless the LLC is continued by written consent of all members; a judicial decree has been issued because it is not reasonably practicable to carry on business).

[98] *See* Bishop, *supra* note 3, at 297. *See, e.g.*, ARK. CODE ANN. § 4-32-602 (Michie 1997); CONN. GEN. STAT. § 34-180(c) (1997); D.C. CODE ANN. § 29-1327 (1998); WYO. STAT. ANN. § 17-15-120 (Michie 2000).

[99] ALA. CODE § 10-12-30 (1998); ARIZ. REV. STAT. § 29-707 (2000); CAL. CORP. CODE § 17252 (West 1999); COLO. REV. STAT. ANN. § 7-80-603 (West 1999); GA. CODE ANN. § 14-11-601 (1998); OHIO REV. CODE ANN. § 1705.16 (Anderson 1998); MICH. COMP. LAWS ANN. § 450.4305 (West 2000); R.I. GEN. LAWS § 7-16-29 (1998).

[100] DEL. CODE ANN. tit. 6, § 18-603 (1998) (prohibiting an LLC member from resign-

utes provide similar restrictions on withdrawal rights.[101] Washington's LLC statute provides that a member may withdraw from the LLC at or upon the happening of events in the operating agreement and that if the operating agreement does not provide otherwise, the member may not withdraw prior to the time for the dissolution and winding up of the company without written consent of all other members.[102] These restrictions on withdrawal and/or distribution rights effectively "lock in" the LLC owner whose agreement lacks a provision expressly bestowing buy-out rights or a provision addressing deadlocks or disputes. In these circumstances, the minority LLC member is potentially more vulnerable to a squeeze-out than a minority shareholder in a closely held corporation, because many corporate statutes follow the Revised Business Corporation Act's statutory mechanism for providing dissolution or buy-out in lieu of dissolution where there is deadlock, where those in control have engaged in illegal, fraudulent, or oppressive conduct, or where there is a wasting of corporate assets.[103]

A primary advantage of eliminating statutory default buy-out rights in the LLC (and the limited partnership as well) is that it arguably enables the minority owner of a family limited partnership or a family-owned LLC to qualify for discounts in the valuation of the family-owned

---

ing prior to the dissolution of the company unless the LLC agreement provides otherwise); DEL. CODE ANN. tit. 6, § 18-604 (1998).

[101] *See* FLA. STAT. ch. 608.427 (1999); N.Y. LTD. LIAB. CO. LAW § 606 (McKinney 2000); OKLA. STAT. tit. 18, § 2036 (1998); *see also* Steven T. Ledgerwood, *A Focus on Unincorporated Businesses: Oklahoma LLCs v. Limited Partnerships: Choice of Entity for Valuation Discounts After 1997*, 22 OKLA. CITY U. L. REV. 611 (1997) (discussing valuation discounts for estate and gift tax purposes and discussing the effects of locking in the LLC owner).

[102] WASH. REV. CODE § 25.15.130 (1999) (restricting the right to withdraw prior to dissolution).

[103] *See* 3 MODEL BUS. CORP. ACT § 14.3 (3d ed. 1999) (providing for a dissolution where:

> (i) the directors are deadlocked in the management of the corporate affairs, the shareholders are unable to break the deadlock, and irreparable injury to the corporation is threatened or being suffered, or the business and affairs of the corporation can no longer be conducted to the advantage of the shareholders generally, because of the deadlock;
>
> (ii) the directors or those in control of the corporation have acted, are acting, or will act in a manner that is illegal, oppressive, or fraudulent;
>
> (iii) the shareholders are deadlocked in voting power and have failed, for a period that includes at least two consecutive annual meeting dates, to elect successors to directors whose terms have expired; or
>
> (iv) the corporate assets are being misapplied or wasted.
>
> If a dissolution action has been instituted, Section 14.30 provides that the corporation or one or more shareholders may elect to purchase all of the shares owned by the petitioning shareholder.).

entities.[104] If the applicable state limited partnership statute or state LLC statute contains a default provision that permits owners to be bought out of the partnership or LLC at will, then, estate planners argue, a minority owner will not qualify for a discount for marketability or for lack of control even though he or she may have entered a private agreement with other family members to relinquish all control of the entity and rights to a buy-out.[105]

Generally, if a taxpayer owns a minority interest in a family corporation, the value of his or her interest is discounted for lack of marketability and lack of control over fundamental business policies.[106] Taxpayers have also attempted to obtain valuation discounts for investments in publicly traded stocks or other assets by placing the assets in a limited partnership or an LLC owned by family members and gifting all but a minority interest in the entity to family members. By holding only a minority interest in the entity, the donor or decedent typically asserts that the value of the interest should be discounted for lack of marketability, illiquidity, and lack of voting control.[107] The overriding sticking point, however, is section 2704(b) of the Internal Revenue Code, which provides that certain "applicable restrictions" will be disregarded for purposes of estate and gift tax valuation of the transferred interest.[108] Under the Treasury regulations, restrictions on voting rights or buy-out rights rooted in partnership agreements or LLC operating agreements will be ignored where the owner's family members have a right to remove the restriction.[109] The Treasury regulations define "applicable restrictions" as including a limitation on the ability to liquidate the entity that is more restrictive than the limitations that would apply under state law generally.[110]

Estate and gift tax planners now suggest that states modify their limited partnership and limited liability company laws to contain restrictions on the ability of an investor to liquidate his or her interest.[111]

---

[104] *See* Mona, *supra* note 6, at 167 (indicating that the drive to obtain valuation discounts has prompted widespread efforts across the country to amend default liquidation rights).

[105] *See id.*

[106] *See* Rev. Rul. 59-60, 1959 C.B. 237 (outlining fundamental factors in valuing a business); 4 ESTATE PLANNING TAX COORDINATOR §§ 82,264–82,366 (2000) (discussing the valuation of a business and discounts for lack of marketability and for minority interests); RICHARD L. LAVOIE, *TAX MANAGEMENT ESTATES, GIFTS, & TRUSTS: VALUATION OF CORPORATE STOCK* 831-2d TAX MGMT. (BNA) A20 to A43 (1998) (discussing the special valuation rules and their implications for estate and gift tax planning).

[107] *See* Dale A. Osterle & Wayne M. Gazur, *What's in a Name? An Argument for a Small Business "Limited Liability Entity Statute (With Three Subsets of Default Rules),"* 32 WAKE FOREST L. REV. 101, 139 (1997) (discussing the estate tax planning issues presented by business entities).

[108] I.R.C. § 2704(b) (1999).

[109] Treas. Reg. § 25.2704-2(a) (1999).

[110] Treas. Reg. § 25.2704-2(b) (1992).

[111] *See* Mona, *supra* note 6, at 167–68.

Moreover, in furtherance of the drive to enhance the usefulness of the limited partnership as an estate and tax planning vehicle, the latest draft version of the Revised Uniform Limited Partnership Act eliminates the right to receive a distribution upon withdrawal from the partnership.[112]

The elimination of default buy-out rights arguably benefits those minority investors who will create family limited partnerships and/or LLCs. However, the consequences of eliminating default buy-out rights may not be beneficial outside the limited context of family estate and gift tax planning. The elimination of default exit rights could create a trap for the unwary minority LLC owner without adequate protection under an LLC agreement.[113] Legislators thus should not react without thorough analysis. The modification of something as significant as buy-out rights should not be driven by a tax technicality, but rather should be based on a thorough consideration of sound business and legal policy.

II. THE ADVANTAGES AND DISADVANTAGES OF DEFAULT EXIT RIGHTS: WHY DEFAULT EXIT RIGHTS SHOULD BE RETAINED

In light of the recent trend to eliminate an LLC owner's rights to withdraw, it is important to analyze the policy interests in support of as well as against the retention of voluntary rights to withdraw from the LLC and obtain the fair market value for the LLC interest.

Several important factors support the view that, unless the LLC agreement provides otherwise, an LLC participant should *not* be entitled to withdraw at-will and obtain the fair market value of his or her LLC interest. A default rule that "locks in" the LLC participant may be desirable in light of the elimination of personal liability in the LLC. In addition, the elimination of a default exit right furthers the business and community interests in the stability of the business enterprise. The elimination of default exit rights also prevents the opportunity for the abusive exercise of the right to demand a buy-out.

On the other hand, a compelling argument can be made in support of the continuation of a default buy-out right for withdrawing LLC members. Factors supporting the default buy-out right include the extensive mutual agency powers possessed by LLC members, the illiquidity of the LLC interest, the intended use of the LLC as a vehicle for small, informal business ventures without extensive operating agreements, the uncertainties surrounding the duty of loyalty and the standard of care owed by LLC members and managers, and the lack of guidance regarding judi-

---

[112] PROPOSED REVISIONS OF UNIF. LTD. P'SHIP ACT (1976) WITH 1985 AMENDMENTS § 505 (completion targeted for 2001) (on file with author) (providing that a person has no right to receive any distribution on account of dissociation).

[113] *See* Farrar & Hamill, *supra* note 4, at 909.

cial mechanisms for a breach of the duty of loyalty and/or care in the context of an LLC.

A. The Case for Eliminating Default Exit Rights in the LLC

It has been argued that the personal liability exposure of partners in a general at-will partnership justifies giving each general partner a right to dissolve the partnership, and thus the right to compel a dissolution serves no meaningful goal in an LLC because the LLC eliminates the investor's personal liability.[114] The specter of personal liability arguably created the need to allow a general partner to expeditiously terminate the agency powers of other partners by affecting a dissolution of the partnership under the UPA.[115] Similarly, it has been said that the prospect of personal liability under the RUPA justified the decision to retain a partner's right to dissociate at will within the fabric of RUPA.[116] The freedom from personal liability in the LLC arguably eliminates the need to provide LLC members with a mechanism for terminating the agency powers among partners.

The policy interest in providing stable business relationships also has supported the adoption of a default rule that eliminates at-will exit rights in the LLC. Commentators have long argued that at-will exit rights characteristic of general partnerships reflect a flawed and unstable business model for a business entity.[117] In fact, the problems associated with giving a single partner the power to dissolve the partnership led to the introduction of the term partnership embraced by RUPA.[118]

Concern also has been expressed regarding the potential for opportunistic conduct created by a power to dissociate from or dissolve a business entity at will.[119] The elimination of statutory default rights to exit from the business arguably frees participants from fear of getting bought out by an economically powerful partner at an unfairly low price.[120]

It is difficult to dispute the notion that the interests of the participants and the community are best served by rules that foster the continuity and stability of businesses. However, the elimination of a default buyout right does not necessarily contribute to the stability of the business. On the contrary, in some circumstances the business as a whole may

---

[114] *See* Farrer & Hamill, *supra* note 4, at 923.

[115] *See* Robert W. Hillman, *Indissoluble Partnerships*, 37 U. FLA. L. REV. 691, 701 (1985) (indicating that the risk of unlimited partner liability created the need for an expeditious method of terminating the agency powers of other partners, and at-will dissolution provided the needed mechanism).

[116] *See* Farrer & Hamill, *supra* note 4, at 920.

[117] *See* Hillman, *supra* note 115, at 719 (discussing the flawed nature of the partnership).

[118] *See id.* at 735.

[119] *See* Farrer & Hamill, *supra* note 4, at 912.

[120] *Id.* at 922–23.

benefit if dissatisfied participants can readily disengage from the business.

B. *The Importance of Retaining Default Exit Rights at the Current Stage of Development of the LLC*

Although substantial estate and gift tax savings may be achieved by some investors if default exit rights are eliminated, there are many non-tax related reasons to retain default buy-out rights in the LLC, including: (1) the extensive mutual agency powers possessed by LLC members; (2) the illiquidity of private investment and the difficulty facing minority partners in negotiating for protection; (3) the intended use of the LLC as a vehicle for the informal conduct of a wide variety of business ventures; (4) the uncertainty regarding the duty of loyalty and the duty of care in the LLC; and (5) the uncertainty regarding judicial actions for breach of the duty of loyalty and/or the duty of care by LLC members and managers. Other, albeit less convincing, arguments that support default buy-out rights include the principle of *delectus personae*, or the right to control one's own business associates, and the impracticality of fashioning remedies between antagonistic parties.[121]

C. *Extensive Agency Powers Create a Need for Expeditious Withdrawals from the LLC*

The legal community has eagerly awaited judicial interpretations of the corporate-like shield from personal liability. Early indications are that courts will embrace broad concepts under agency principles to protect third parties in their dealings with LLCs and that LLC members may be vulnerable to personal liability under theories of agency law. For example, in *Water, Waste & Land, Inc. v. Lanham*,[122] the Supreme Court of Colorado held that Lanham, a manager/member of an LLC, was liable personally on a contract where a fellow LLC member, Clark, apparently led a third party to believe that the contract was being executed with Lanham in his individual capacity. Clark had given the plaintiff the LLC's business card, which included the initials of the LLC (P.I.I.) and Lanham's personal address, which also served as the address of the LLC. However, the business card failed to indicate what the acronym P.I.I. meant. The court applied the principles of agency law and held Lanham personally liable on the contract because the third party presumably thought that Clark was acting as Lanham's agent rather than as an agent of the LLC.

---

[121] These factors were discussed in connection with default buy-out rights of partnerships. *See supra* note 115 and accompanying text.

[122] 955 P.2d 997 (Colo. 1998).

The decision in *Water, Waste & Land, Inc. v. Lanham* underscores the LLC participants' vulnerability to personal liability under agency principles. The case highlights that participants need to have an efficient mechanism for terminating relations with the LLC, particularly where the LLC is a member-managed company. While the potential vulnerability to personal liability through agency principles and/or veil-piercing judicial doctrines is arguably present for owners of all types of business entities, it appears to be more acute in the general partnership and the member-managed LLC, which are structured to have all of the participants manage the business. Arguably, entities with a greater number of active business participants have a greater likelihood that any single member's conduct will be imputed to either the business entity or to the other owners personally through the use of agency theories than for business entities with fewer participants actively managing the business, such as in the case of a limited partnership, a manager-managed LLC, or a close corporation with passive stockholders.

D. The Illiquidity of the Private Investment and the Difficulty of the Minority in Negotiating for Protection

The fact that shareholders are locked into their investments in private corporations has created a breeding ground for majority shareholder oppression of minority owners.[123] Tactics employed by majority shareholders have typically included: withholding dividends from the minority, restricting or precluding employment, paying excessive compensation to majority owners, withholding information from minority members, violating procedural restrictions on corporate governance, and depriving the minority a voice in corporate decision-making.[124] LLC owners have been regarded as less vulnerable than their corporate counterparts precisely because the dissatisfied LLC participant has typically had the statutory default right to receive the fair market value of his or her interest upon withdrawal.[125]

The minority LLC owner without default exit rights would be locked into his or her investment in the same way a corporate shareholder is locked into the corporation unless he or she had the foresight and negotiating power to contractually address the consequences of a deadlock or dispute and has negotiated in advance for buy-out rights.

---

[123] *See* Sandra K. Miller, *A Note on the Definition of Oppressive Conduct by Majority Shareholders: How Can the Reasonable Expectation Standard Be Reasonably Applied in Pennsylvania?*, 12 J.L. & COM. 51, 53 (1992).

[124] Orchard v. Covelli, 590 F. Supp. 1548, 1557 (W.D. Pa. 1984) (outlining the typical tactics of minority shareholder oppression).

[125] *See* Ribstein, *supra* note 86, at 340 (discussing the minority shareholder problem in the closely held corporation).

The vulnerability of a minority investor in an LLC is perhaps most acute where the minority actively participates in the business and is both an investor and an employee. The minority's vulnerability stems from at least three major sources, including: (1) the majority's ultimate power to control LLC business decisions; (2) the significant financial stake the minority typically has in an enterprise that serves both as a chief employer and chief financial investment; and (3) the inability to sell the LLC interest to anyone other than the majority.[126] As noted by Justice D.J. Weber in a dispute between minority and majority owners, the minority owner may well find himself "on the horns of a dilemma, he can neither profitably leave nor safely stay with the corporation."[127]

The difficulty a minority LLC investor may have in effectively negotiating buy-out rights and obtaining other contractual protections from majority overreaching cannot be over-emphasized. The observations of F. Hodge O'Neal were made in connection with the corporation but are equally applicable to the limited liability company:

> Statutory protection is needed for minority shareholders who fail to bargain for and obtain protective contractual arrangements. Although most state corporation statutes validate special charter and bylaw provisions and shareholder's agreements designed to protect minority shareholders, no statute—not even any of the separate, integrated close corporation statutes—furnishes adequate self-executing protection for minority shareholders who have failed to bargain for special charter provisions or for protective clauses in shareholders' agreements . . . . He (the minority shareholder) may be unaware of the risks involved, or his bargaining position may be so weak that he is unable to negotiate for protection.[128]

Although some limited liability companies may be subject to detailed operating agreements that have been negotiated at arms length, others may not have extensive agreements that establish the parties' rights and responsibilities. Some limited liability companies may not have an agreement at all. Others, including some small limited liability companies, may be subject to boilerplate agreements that do not contain detailed provisions dealing with deadlocks or disputes among owners

---

[126] *See* Meiselman v. Meiselman, 307 S.E.2d 551, 558 (N.C. 1983) (indicating that when the personal relationship among shareholders breaks down in the close corporation, "the majority shareholder, because of his greater voting power, is in a position to terminate the minority's employment and to exclude him from participation in management decisions").

[127] *Orchard*, 590 F. Supp. at 1557.

[128] F. Hodge O'Neal, *Close Corporations: Existing Legislation and Recommended Reform*, 33 BUS. LAW. 873, 881–83 (1978).

and/or managers. In fact, many small businesses are staffed by friends or family members who have such a close personal relationship that they may not consider that disputes could arise at a later time.[129] To the extent that an LLC agreement is executed, it may not address the grievances that eventually develop between members.[130] Despite the existence of a boilerplate LLC agreement, the understanding between the parties to small informal business ventures may be primarily oral, or at least not fully articulated in the operating agreement.[131]

An assumption of a level contractual playing field may be unrealistic in the context of many family limited liability companies. In second-generation businesses, for example, children may join parents or relatives in a business enterprise. Minority interests may be inherited and the business may be operated by a variety of relatives or acquaintances over time. The original LLC operating agreement, however, may have been personal to those parties who initially formed the limited liability company and negotiated and drafted the agreement. Members who inherit an LLC interest or who join an LLC years after its formation may lack either the foresight or the leverage to renegotiate the fundamental workings of an existing LLC agreement. Those minority owners who inherit an LLC interest as assignees and are inactive may lack the sophistication or

---

[129] *See Mesielman*, 307 S.E.2d at 558 (discussing close corporations).

[130] *Cf.* McCauley v. McCauley, 724 P.2d 232 (N.M. Ct. App. 1986) (involving a lawsuit by a divorced wife against her former husband, parental in-laws and two sons, alleging that she had been ousted from the board and denied a voice in the company's affairs, that the defendants improperly failed to pay dividends, and that they converted corporate assets to personal use. Nothing in the company by-laws or contractual agreements between the parties addressed the grievances at issue.); Fox v. 7L Bar Ranch Co., 645 P.2d 929 (Mont. 1982) (involving an action for dissolution of a family corporation. Neither the articles of incorporation nor any other contractual documents were cited. The complaint focused largely on a failure to pay dividends.). In a number of cases alleging "oppressive majority conduct," neither the plaintiff nor the defendant referred to the company's articles of incorporation or to other contractual documents. *See* Baker v. Commercial Body Builders, Inc., 507 P.2d 387 (Or. 1973) (involving exclusion of the plaintiff as an employee of the corporation); Masinter v. Webco Co., 262 S.E.2d 433 (W. Va. 1980) (involving a squeeze-out of the minority shareholder from the Board of Directors).

[131] *See* O'Neal, *supra* note 128, at 884. Professor O'Neal, the leading authority on close corporations, explains:

> Yet, many participants in closely-held corporations are "little people," unsophisticated in business and financial matters. Not uncommonly a participant in a closely-held enterprise invests all his assets in the business with an expectation, often reasonable under the circumstances even in the absence of an express contract, that he will be a key employee in the company and will have a voice in business decisions.

Id. Many close corporations are conducted without observing corporate formalities. *See* Balvik v. Sylvester, 411 N.W.2d 383 (N.D. 1987) (observing that the articles of incorporation had provided that a separate buy-sell agreement would be entered into but the parties never executed it).

the bargaining power to modify the fundamental terms of the inherited LLC operating agreement.[132]

Years of experience with the traditional corporate model illustrate that minority owners of closely held businesses require special statutory protections from fraudulent or opportunistic majority misconduct.[133] The history of dispute resolution among business participants illustrates that there will be failures in contractual agreements as well as failures in human relationships in the context of LLCs, as in other private business entities.[134] Difficulties between majority and minority LLC members, therefore, should be anticipated.[135]

---

[132] This issue is discussed at length in connection with whether reasonable expectations can be imputed to minority shareholders who obtained their rights in a close corporation without bargaining for them. *See* Steven C. Bahls, *Resolving Shareholder Dissension: Selection of the Appropriate Equitable Remedy*, 15 J. CORP. L. 285, 326 (1990) (indicating that minority owners frequently obtain their shares without bargaining with majority shareholders and courts have thus refused to attribute the reasonable expectations of the transferor to the transferee).

> A person taking a minority position in a close corporation often leaves himself vulnerable to squeeze-out or oppression .... He may be unaware of the risks involved, or his bargaining position may be so weak that he is unable to negotiate for protection. Further, he may have been given or may have inherited his minority interest.

O'Neal, *supra* note 128, at 883.

[133] *See* O'Neal, *supra* note 128, at 883 (analyzing the status of legislation for closely held corporations and indicating the inadequacy of traditional corporate standards when applied to close corporations); Frank H. Easterbrook & Daniel R. Fischel, *Close Corporations and Agency Costs*, 38 STAN. L. REV. 271 (1986) (reviewing the cost of managing close corporations); Victor B. Brudney & Marvin A. Chirelstein, *A Restatement of Corporate Freezeouts*, 87 YALE L.J. 1354 (1978) (reviewing and classifying corporate freezeouts from the perspective of public corporations, including a discussion of going private transactions); Bahls, *supra* note 132, at 288, 312, 320 (reviewing the illiquidity and plight of minority shareholders, reviewing legislative efforts to address dissension, and proposing standards for determining appropriate remedies for shareholder dissension); Robert W. Hillman, *The Dissatisfied Participant in the Solvent Business Venture: A Consideration of the Relative Permanence of Partnerships and Close Corporations*, 67 MINN. L. REV. 1, 75 (1982) (reviewing the power to dissolve partnerships and close corporations and focusing on the importance of investors' expectations in close corporations); Richard A. Mann, *A Critical Analysis of the Statutory Close Corporation Supplement to the Model Business Corporation Act*, 22 AM. BUS. L.J. 289 (1984) (containing an analysis of the Statutory Close Corporation Supplement to the Model Business Corporation Act, indicating that while closely held corporations comprise the large majority of incorporated business entities in the United States, their special needs differ from those of the public issue corporation and, until recently, have not received statutory protection); John E. Davidian, *Corporate Dissolution in New York: Liberalizing the Rights of Minority Shareholders*, 56 ST. JOHN'S L. REV. 25 (1981) (providing an overview of the development of New York legislation enhancing the rights of minority shareholders).

[134] *See* Miller, *Minority Shareholder Oppression*, *supra* note 8, at 580–81 (discussing the scope of conflicts among business participants).

[135] The magnitude of dissension in private business entities should not be underestimated.

> Close corporations account for most of American business. Family owned businesses alone represent ninety-five percent of the jobs in the United States. How-

At a minimum, a default buy-out right should be made available to LLC owners who find themselves in a dispute with the majority without any operating agreement or with ineffectual contractual protection. Even if most LLC owners tend to have a contractual buy-out clause, default statutory protection should exist for those few who do not.[136]

Further, a variety of statutory protections in addition to default buy-out rights are important to deter fraudulent and opportunistic majority conduct.[137] Now that Check-the-Box regulations permit an LLC to have perpetual life, a growing number of LLC members will find themselves locked into the LLC for a specified term or indefinitely. Like corporate shareholders, members who are locked into an LLC should have the remedy of dissolution in cases where those in control have engaged in illegal, fraudulent, or unfairly prejudicial conduct. Most states provide shareholders the right to seek a corporate dissolution in the event of certain majority misconduct, and it is suggested that a similar remedy be provided to locked-in minority LLC members.[138]

> ever, in a study by Professor John L. Ward of Loyala Universtiy of Chicago, eighty percent of the Chicago area family-owned corporations that were in existence in 1924 and had at least twenty employees were no longer going concerns in 1984. Some of the major reasons for the business failure of closely held businesses are "typical family problems [such] as sibling rivalry [and] competition between the generations," that result in shareholder dissension and corporate succession problems. Costs associated with dissension include ineffective use of management time, resource-draining litigation, loss of a business's ability to obtain necessary financing, as well as the obvious costs associated with business failure.
>
> Dissension within a company produces non-economic losses as well. If allowed to escalate, it can destroy sound family relationships and lead to vindictiveness. In one sense, dissension in close corporations is like dissension in a marriage. In both circumstances, complex emotional and financial relationships exist that courts cannot easily dissolve without losses. Just as courts have developed standards in dissolution of the marital relationship to minimize hardship, courts also must develop standards to resolve dissension in close corporations.

Bahls, *supra* note 132, at 287.

[136] *See* Farrar & Hamill, *supra* note 4, at 929–34 (noting that the elimination of dissociation rights exposes unsophisticated business owners to oppression and squeeze-out techniques).

[137] Franklin A. Gevurtz, *Limited Liability Companies: Squeeze-outs and Freeze-outs in Limited Liability Companies*, 73 WASH. U. L.Q. 497, 507–16 (1995) (discussing LLC operating rules in various jurisdictions and the potential for majority misconduct).

[138] *See* 3 MODEL BUS. CORP. ACT ANN. § 14.30 (3d ed. 1999). According to the statutory comparison, virtually all states provide for an involuntary dissolution in defined circumstances. *See* ALA. CODE § 10-2B-14.30 (1994) (providing for involuntary dissolution where the directors or those in control have acted, are acting, or will act in a manner that is illegal, oppressive, or fraudulent); ARK. CODE ANN. § 4-27-1430 (Michie 1996) (same); ARIZ. REV. STAT. ANN. § 10-1430 (West 1996) (same); COLO. REV. STAT. ANN. § 7-114-301 (West 1998) (same); CONN. GEN. STAT. ANN. § 33-896 (West 1997) (providing for judicial dissolution, in a proceeding by shareholders, to occur if it is established that the directors or those in control have acted, are acting, or will act in a manner that is illegal, oppressive, or fraudulent or where the assets are being misapplied or wasted); IDAHO CODE § 30-1-1430 (Michie 1998) (providing for involuntary dissolution in the event of illegal,

E. The Intended Use of LLCs as a Vehicle for the Informal Conduct of a Broad Spectrum of Businesses

Although the primary motivation for eliminating the default buy-out rights of LLC members is to enable taxpayers to qualify for estate and gift tax valuation discounts, a substantial number of LLCs are not, in fact, formed for use as a family held investment. Indeed, the LLC was enacted to serve as the business entity of choice for a broad array of privately owned businesses.[139]

Nationwide data suggest that use of the LLC form is increasing rapidly and that LLCs are housing a wide variety of business enterprises.[140] It thus does not appear that LLCs are being used for the relatively narrow purpose of holding family investments.

---

oppressive, or fraudulent conduct and where irreparable injury to the corporation is thereby threatened or suffered); ILL COMP. STAT. ANN. § 5/12.50, 5/12.55 (West 1993) (recognizing the involuntary dissolution action and alternative remedies for public and private companies where illegal, oppressive, or fraudulent conduct has occurred); MISS. CODE ANN. § 79-4-14.30 (1999) (authorizing involuntary dissolution for illegal, oppressive, or fraudulent conduct); MONT. CODE ANN. § 35-1-938 (1999) (providing for dissolution on the basis of illegal, oppressive, or fraudulent conduct); N.J. STAT. ANN. § 14A:12-7 (Supp. 1998) (authorizing dissolution where there are 25 or fewer shareholders and where the directors or those in control have acted fraudulently or illegally, have mismanaged the corporation or abused their authority as officers, or have acted oppressively or unfairly toward the minority shareholders in their capacities as shareholders, directors, officers or employees); N.Y. Bus. CORP. § 1104-a (McKinney 1986) (permitting holders of 20% or more of outstanding shares to petition for dissolution where directors or those in control have been guilty of illegal, fraudulent, or oppressive actions toward the complaining shareholders); 15 PA. CONS. STAT. ANN. § 1981 (West 1998) (permitting involuntary dissolution in the case of illegal, fraudulent, or oppressive conduct); VA. CODE ANN. § 13.1-747 (Michie 1999) (providing for involuntary dissolution on the grounds of illegal, oppressive, or fraudulent conduct). *But see* FLA. STAT. ANN. § 607.1430 (West 1998) (providing for dissolution where the directors or those in control have acted, are acting, or are reasonably expected to act in a manner that is illegal or fraudulent but excluding language for oppressive conduct); GA. CODE ANN. § 14-2-1430 (1994) (allowing for involuntary dissolution for illegal or fraudulent conduct but not for oppressive conduct); HAW. REV. STAT. ANN. § 415B-97 (Michie 1997) (allowing for involuntary dissolution for illegal or fraudulent conduct but excluding any language concerning oppressive conduct); IND. CODE ANN. §§ 23-1 to 47-1 (West 1976) (omitting oppression as a ground for involuntary dissolution but authorizing dissolution in part, in the event of fraud, abuse of authority, or deadlock); TEX. BUS. CORP. ACT ANN. § 7.01 (Vernon 1980) (omitting dissolution on the basis of oppressive conduct but authorizing involuntary dissolution on the grounds of fraud, misrepresentation, operating outside its incorporated scope or where the corporation is convicted of a felony or a high managerial agent has engaged in a persistent course of felonious conduct).

[139] *See* Brian L. Schorr, *Limited Liability Companies: Features and Uses*, C.P.A. J., Dec. 1992, at 32–34 (indicating that LLCs are desirable for use in the following: corporate joint ventures, entrepreneurial businesses, family businesses, start-up businesses, high technology and research businesses, oil and gas investments, investments in theatrical productions, real estate investments, transactions involving international investors, management leveraged buy-outs, structured finance arrangements, and commodity pools).

[140] Conrad S. Ciccotell & C. Terry Grant, *LLCs and LLPs: Organizing to Deliver Professional Services*, BUS. HORIZONS, Mar.-Apr. 1999, at 85 (analyzing the growth of LLCs through an analysis of LLC and LLP registrations in 49 states).

According to a recent analysis of records of the individual secretaries of state compiled by the International Association of Corporation Administrators, new registrations of corporations and limited partnerships grew by 13 and 15% respectively between 1992 and 1996, whereas LLCs grew by over 2300% in the same time period.[141] By 1996, nearly one in six new business registrations were LLCs.[142] A sample of registration records show that a wide variety of industries are utilizing the LLC form.[143] LLCs span a significant range of business classifications, including engineering and management support services, real estate, construction and general contracting, investments, retail, health services, amusement and recreation, agriculture, oil and gas, restaurants, and leasing services.[144] Of the approximately 1200 LLCs analyzed, 26% were in engineering and management support, 19% were in real estate, 12% were in construction, and 9% were investment companies.[145] The remaining one-third of the sample was spread over a number of different industries.[146]

While the data do not indicate the LLCs used within the context of family estate planning, the diversity of LLC registrations suggests that the LLC is serving the needs of a broad base of the business community. In light of this apparent diverse use of LLCs, it may be more sensible to retain default buy-out rights in each state's LLC statute while removing default buy-out rights in a different statute such as the state's limited partnership statute. If, for example, it appears that most estate planners commonly use the family limited partnership as a means of obtaining minority discounts for estate tax purposes, it may make the most sense to eliminate default exit rights in the state's family limited partnership statute, while retaining default exit rights in the state's LLC and general partnership statutes.

The Re-RULPA eliminates distributions when a partner withdraws from the limited partnership.[147] However, the National Conference of Commissioners on Uniform State Laws, which drafted the Re-RULPA, includes the default buy-out right in the ULLCA in the case of at-will limited liability companies without a specific term.[148]

---

[141] *Id.* at 87.

[142] *Id.*

[143] *Id.*

[144] *Id.* at 89; *see also* Ribstein, *supra* note 10, at 427–33 (suggesting that non-professional as well as professional firms may elect to use the LLC).

[145] Ciccotello & Gant, *supra* note 140, at 89.

[146] *Id.*

[147] PROPOSED REVISIONS OF UNIF. LTD. P'SHIP ACT (1976) WITH 1985 AMENDMENTS § 504 (completion targeted for 2001) (on file with author) (providing that a partner has no right to a distribution before the dissolution and winding up of the limited partnership unless the limited partnership decides to make an interim distribution).

[148] UNIF. LTD. LIAB. CO. ACT § 701(a)(1) (1996) (containing a default buy-out right for at-will LLCs, a construct that was originally developed before the Check-the-Box Regulations were enacted but which, for the time being, remains in effect under the ULLCA).

It may be preferable to eliminate default buy-out rights in a limited partnership statute rather than in a widely used LLC statute. Alternatively, a different entity without buy-out rights could be used by those seeking minority discounts for estate and gift tax purposes. A business entity created largely for the purpose of achieving estate and gift tax goals is likely to be the subject of extensive tax and business planning. Participants in such a business entity would be on notice that their statutory rights and remedies may be abridged in deference to estate and gift tax planning goals. For example, Colorado has created the Colorado Limited Partnership Association.[149] Under its default rules, an interest in a Limited Partnership Association may only be transferred as specified in the by-laws, and the default rule is non-transferability.[150] The creation of a special entity without default buy-out rights serves the dual purpose of providing a vehicle for achieving tax goals, while placing investors on notice that special rules may apply. In this manner, the default buy-out rights of a diverse range of unsuspecting LLC investors will not be compromised.

F. The Uncertainties Surrounding the Standard of Care and the Duty of Loyalty Increase the Importance of Default Buy-Out Rights and Other Statutory Protections

Since the LLC is still in its infancy, there is not a well-developed body of law pertaining to the standard of care and the duty of loyalty applicable to LLC members and managers. Given the lack of clear-cut judicial guidance regarding standards of LLC conduct, and uncertainties as to the processes for enforcing these standards, LLC participants should be strongly urged to address standards of conduct, remedies for breach, and processes for resolving deadlock in their own operating agreements. Legislators also should be extremely cautious about eliminating protective statutory provisions such as default buy-out rights. In fact, now that LLCs may have a perpetual existence without jeopardizing their tax status as a partnership, it is an excellent time to review the presence, or absence, of statutory protections for minority LLC owners.

G. How Will Courts Interpret the Standard of Care for LLC Members and Managers and How Expansively Will They Apply the Business Judgment Rule?

The Prototype Limited Liability Company Act ("PLLCA") drafted by the Business Law Section of the American Bar Association ("ABA") provides that a member or manager of an LLC will not be liable or ac-

---

[149] COLO. REV. STAT. ANN. § 7-63-101 to -117 (West 1999).
[150] Oesterle & Gazur, *supra* note 107, at 139.

countable in damages or otherwise unless the act or omission constitutes gross negligence or willful misconduct.[151] The Uniform Limited Liability Act similarly embraces a gross negligence standard, providing that the partner's duty of care is to refrain from engaging in grossly negligent or reckless conduct, intentional misconduct, or a knowing violation of law.[152] According to the comments contained in the PLLCA, the gross negligence standard of care is similar to the standard commonly applied to corporate directors, managing partners, or general partners of limited partnerships.[153]

Although the ABA states that gross negligence is the standard commonly applied to corporate directors, managing partners, or general partners of limited partnerships, there is some uncertainty as to how courts will develop the precise boundaries of the standard of care in the context of LLCs and how courts will apply the business judgment rule in various settings.[154] Although a leading treatise on limited liability companies indicates that the standards of good faith and prudence applicable to LLC managers and members mirror those found under corporate law, there may be considerable variation in how those standards are applied in different contexts.[155] Will the extent of judicial intervention depend in part on whether the aggrieved participant is an active or passive LLC member?[156] Will the conduct of an LLC manager be subject to increased judicial scrutiny where the other LLC members are relying heavily on the special expertise of the LLC manager? Will differences in the economic power of the majority and minority LLC owner influence the judicial posture? How expansively will courts apply the business judgment rule to insulate the majority from liability where, for example, the majority has

---

[151] *See* PROTOTYPE LTD. LIAB. CO. ACT § 402(A) (1992) (establishing the duty to refrain from gross negligent conduct).

[152] *See* UNIF. LTD. LIAB. CO. ACT § 409 (1996) (setting forth the standards of conduct for members and managers).

[153] *See* PROTOTYPE LTD. LIAB. CO. ACT § 402(A) cmt.

> (Subsection (A) sets forth the gross negligence standard of care for those participating in management. This is similar to the standard commonly applied to corporate directors, managing partners, or general partners of limited partnerships. In general, as long as managers avoid self-interested and grossly negligent conduct, their actions are protected by the business judgment rule.).

[154] *See id.*

[155] *See* THOMAS A. HUMPHREYS, LIMITED LIABILITY COMPANIES § 4–29 (1998) (discussing the duty of care for LLCs).

[156] Some courts have scrutinized the conduct of majority shareholders particularly where the minority shareholder was a passive owner of the private company. *See* Kelley v. Axelsson, 687 A.2d 268, 274 (N.J. Super. Ct. App. Div. 1997) (involving suit for compulsory buy-out by passive shareholder where the majority failed to keep proper books and records and maintain a basic system of internal control); Bonavita v. Corbo & Corbo Jewelers, Inc., 692 A.2d 119, 121, 130 (N.J. Super. Ct. Ch. Div. 1996) (ordering shareholder buy-out where the widow of a former shareholder did not actively participate in management).

decided to modify its services or products and in the process has fired a minority LLC owner who has actively managed the business for many years?[157]

The precise boundaries of the duty of care required of members and managers of LLCs have yet to be fully developed, thus making the prospects of litigation alleging LLC member or manager misconduct somewhat uncertain. The judicial uncertainties regarding the appropriate standard of care make contractual and statutory protection even more important for the minority LLC member.

A review of LLC statutes reveals no uniform language pertaining to the standard of care. A number of states, including Delaware, fail to adopt any express standard of conduct and fail to elaborate on the member's or manager's responsibilities.[158] Other states contain standards of conduct for managers but not for members.[159] At least one LLC statute contains language indicating that a member is not liable solely by virtue of being a member where the LLC is managed by a manager.[160]

A number of states employ statutory language similar to that applied to directors under the Model Business Corporation Act.[161] For example, Alaska, Colorado, Georgia, Louisiana, Maine, Michigan, Minnesota, Mississippi, New York, North Carolina, North Dakota, Oklahoma, Pennsylvania, and Virginia basically adopt a corporate style standard of care for management, which considers the good faith of the managers in discharging their duties, the best interests of the LLC, and what an ordinary, prudent person in a like position would have done.[162] Some, although not

---

[157] *See* Muellenberg v. Bikon Corp., 669 A.2d 1382 (N.J. 1996) (ordering the majority to buy out minority shareholder where the minority had actively managed the business for the majority owner).

[158] *See* DEL. CODE ANN. tit. 6, § 18-1101 (1999) (containing no express standards of conduct for members and managers but indicating that the operating agreement may expand or restrict any duties at law or equity); TEX. REV. CIV. STAT. ANN. art. 1528n, § 2.12–2.20 (Vernon 1999) (containing rules pertaining to managers but failing to specify a standard of care). Other states that fail to provide a statutory duty of care include Arizona, Kansas, Nebraska, Nevada, South Dakota, Utah, and Wyoming. *See* HUMPHREYS, *supra* note 155, § 4–32 (observing the different approaches to standards of care and listing the states that do not appear to adopt an express standard of care).

[159] *See, e.g.*, COLO. REV. STAT. ANN. § 7-80-406 (West 1999); FLA. STAT. ANN. § 608.4225 (West 2000).

[160] ARK. CODE ANN. § 4-32-402 (1996).

[161] *See* MODEL BUS. CORP. ACT ANN. § 8.30 (1999). This Act provides in part that a director shall discharge his duties as a member of a committee:

> 1) in good faith;
> 2) with the care an ordinarily prudent person in a like position would exercise under similar circumstances; and
> 3) in a manner he reasonably believes to be in the best interest of the corporation.

Id.

[162] ALASKA STAT. § 10.50.135 (Michie 1998); COLO. REV. STAT. ANN. § 7-80-406(1) (West 1999); GA. CODE ANN. § 14-11-305 (1994); LA. REV. STAT. ANN. § 12:1314 (West Supp. 2000); ME. REV. STAT. ANN. tit. 31, § 652 (West 1996); MICH. COMP. LAWS

all, of these statutes apply the corporate-style standard of care to managers without expressly extending it to members in a member-managed LLC.[163] At least one state requires that a violation of managerial duty be proved by clear and convincing evidence.[164]

Other LLC statutes follow the standard of care articulated in the RUPA, requiring that members and/or managers refrain from engaging in grossly negligent or reckless conduct, intentional misconduct, or a knowing violation of law.[165] For example, the LLC statutes in Alabama, Arkansas, Florida, and Hawaii impose a duty to refrain from grossly negligent conduct.[166]

Although the above formulations of the standard of care may ultimately boil down to a standard premised upon gross negligence, even the commentary to the ABA Prototype Limited Liability Company Act recognizes that there are important differences among firms and there may be variations in applying a general standard.[167] According to Alan R. Bromberg and Larry E. Ribstein, partners are not subject to the ordinary care standard applicable to a paid agent, and the consensus appears to be that LLC members similarly are not subject to an ordinary care standard.[168] However, recognizing the vulnerability of minority owners who do not actively participate in the business enterprise, courts may, and arguably should, take a more active posture in policing managerial misconduct in the case of LLCs owned by passive investors than in the case of those actively managed by all owners. Increased judicial supervision may well be justified in the case of LLCs with passive members in light of the illiquidity of the passive private investment, the reliance that passive in-

---

§ 450.4404 (West 2000); MINN. STAT. ANN. § 322B.663 (West Supp. 1998); MISS. CODE ANN. § 79-29-402 (2000); N.Y. LTD. LIAB. CO. LAW 32A § 409 (McKinney 2000); N.C. GEN. STAT. § 57C-3-22 (2000); N.D. CENT. CODE § 10-32-96 (1999); OKLA. STAT. tit. 18, § 2016 (1999); 15 PA. CONS. STAT. § 8943 (1998); VA. CODE ANN. § 13.1-1024.1 (Michie 2000).

[163] For example, the standards articulated in Alaska, Minnesota, Michigan, Mississippi, New York, North Carolina, Oklahoma, and Pennsylvania apply expressly to managers. The standard established in Georgia and Louisiana applies to both members and managers. The North Dakota standard appears to be potentially applicable to both managers and members who have been delegated managerial duties. *See supra* note 162.

[164] *See* OHIO REV. CODE ANN. § 1705.29 (Anderson 2000).

[165] *See* REVISED UNIF. P'SHIP ACT § 404(c) (1997); *see also* HUMPHREYS, *supra* note 155, § 4-29 to 36 (discussing the standards of care applicable in the context of LLCs).

[166] *See* ALA. CODE § 10-12-21(g), -21(k)(2) (2000); ARK. CODE ANN. § 4-32-402(a) (Michie 2000); FLA. STAT. ANN. § 608.4225 (West 2000); HAW. REV. STAT. § 428-409 (2000).

[167] *See* PROTOTYPE LTD. LIAB. CO. ACT § 402 commentary (1992).

[168] *See* ALAN R. BROMBERG & LARRY E. RIBSTEIN, 2 BROMBERG & RIBSTEIN ON PARTNERSHIP § 6.07(f) (2000) (indicating that partners are not subject to the ordinary care standard applicable to agents, who must act with the skill and care standard of the locality for the kind of work the agent is employed to perform and also must exercise any special skill the agent has) (citing RESTATEMENT (SECOND)OF AGENCY §§ 379, 400–402; Wyler v. Feuer, 149 Cal. Rptr. 626 (Cal. Ct. App. 1979)). Both ULLCA and the ABA Prototype embrace a gross negligence standard. *See* UNIF. LTD. LIAB. CO. ACT § 404 (1996); PROTOTYPE LTD. LIAB. CO. ACT § 402A.

vestors must necessarily place in management, and the lack of regulatory and accounting controls normally applicable to private business enterprises.[169] The facts and circumstances of particular cases—i.e., where an LLC member is admitted specifically because of special expertise, where others have delegated special responsibilities to a member, or where a particular member has been granted exclusive power over certain functions—may also justify heightened judicial scrutiny.

The degree to which the business judgment rule should apply in the context of LLCs also has yet to be fully developed in the case law. The business judgment rule has been part of the common law for at least 150 years.[170] Under the business judgment rule, courts will not second-guess an informed business decision, even if the decision subsequently proves to be ill-conceived.[171] The rationale of the business judgment rule is to encourage risk-taking and innovation and to limit litigation and judicial intrusiveness into private-sector business decision-making.[172] In both general and limited partnerships, some courts have applied the business judgment rule liberally to shield managing partners.[173]

Indeed, there are sound reasons why management should be protected from liability for their business judgments. As one court noted:

> [T]the controlling group in a close corporation "must have some room to maneuver in establishing the business policy of the corporation. It must have a large measure of discretion, for example, in declaring or withholding dividends, deciding whether to merge or consolidate, establishing the salaries of corporate officers, dismissing directors with or without cause, and hiring and firing corporate employees.[174]

The business judgment rule has yet to be fully developed in litigation in the context of LLCs. While LLC managers must have discretion

---

[169] *See* Hetherington & Dooley, *supra* note 20, at 35 (observing that the inability to set workable limits on managerial prerogatives gives the majority opportunities to enhance its own interests, which it cannot be expected to resist entirely).

[170] *See* Gries Sports Enters., Inc. v. Cleveland Browns Football Co., 496 N.E.2d 959 (Ohio 1986).

[171] *See id.*

[172] *See, e.g.*, A. Gilchrist Sparks III & Lawrence A. Hamermesh, *Common Law Duties of Non-Director Corporate Officers*, 48 BUS. LAW. 215, 230 (1992).

[173] *See* Bane v. Ferguson, 890 F.2d 11, 14 (7th Cir. 1989) (sheltering the managing partner of a law firm from claims of negligent management); ARTRA Group, Inc. v. Salomon Bros. Holding Co., Inc. 680 N.E.2d 769, 773–74 (Ill. App. Ct. 1997) (holding that projections that were overly optimistic reflected business judgment and did not give rise to an action); Wyler v. Feuer, 149 Cal. Rptr. 626, 633–634 (Cal. Ct. App. 1979) (shielding general partners from liability for mistakes made in good faith in the exercise of business judgment in connection with a movie production); BROMBERG & RIBSTEIN, *supra* note 168, § 6.07(f) (discussing the application of the business judgment rule in general and limited partnerships).

[174] Wilkes v. Springside Nursing Home, Inc., 353 N.E.2d 657, 663 (Mass. 1976).

in the conduct of business, and should not be liable for poor business decisions otherwise made in good faith,[175] an overly expansive application of the business judgment rule could trivialize the standard of care, and could operate as a dangerous obstacle to the removal of poor and incapable management.[176] Courts should be wary of applying the business judgment rule so expansively that it undermines the standard of care.

In light of the infancy of the case law on LLC standards of conduct, and the traditional judicial deference shown to business judgment, legislators should be reluctant to relinquish statutorily based protections for minority owners, such as default buy-out rights, and should consider statutorily based strategies that will strengthen minority protections because many LLCs now have a perpetual existence or an extended term.

H. How Will Courts Interpret the Duty of Loyalty of LLC Members and Managers?

The judicial approach to the duty of loyalty, as distinguished from the duty of care, has traditionally varied between corporations and partnerships. Overall, it may be said that judicial oversight of fiduciary duties has been more exacting in the partnership context than in the context of public corporations.[177] A stricter interpretation of the duty of loyalty and greater judicial supervision has traditionally been observed in partnerships, where Justice Cardozo eloquently stated that "joint adventurers, like copartners, owe to one another . . . the duty of finest loyalty . . . . Not honesty alone, but the punctilio of an honor the most sensitive . . . ."[178] This heightened duty of loyalty has been extended to close corporations.[179]

Because the LLC is a hybrid entity, containing both partnership and corporate features, the level of judicial oversight of the duty of loyalty as applied to LLC members and managers cannot be certain, and may well vary among jurisdictions, depending on the philosophy toward judicial monitoring to protect minority owners of private enterprises. Some jurisdictions might apply a heightened duty of loyalty to an LLC. Others may

---

[175] *See In Re* Gary Smith, 546 N.Y.S.2d 382, 384 (N.Y. Sup. Ct. 1989) (reversing an order granting dissolution where there was no evidence to support a finding that salaries were paid to the majority in lieu of dividends or that the work they performed was duplicative); Exadaktilos v. Cinnaminson Realty Co. Inc., 414 A.2d 994 (N.J. Super. Ct. App. Div. 1980), *aff'd*, 400 A.2d 554, 561–62 (N.J. Super. Ct. Law Div. 1979) (refusing to grant a dissolution where there was a business purpose in firing the plaintiff).

[176] *See* Bahls, *supra* note 132, at 288–94 (1990) (analyzing the plight of the minority shareholder and proposing standards for selecting the appropriate equitable remedy).

[177] *See* John C. Coffee, Jr., *No Exit? Opting Out, The Contractual Theory of the Corporation, and the Special Case of Remedies*, 53 BROOK. L. REV. 919, 940 (1988) (discussing the elimination of due care liability in corporations).

[178] Meinhard v. Salmon, 164 N.E. 545, 546 (1928).

[179] Donahue v. Rodd Electrotype Co. of New England, 328 N.E.2d 505, 516 (Mass. 1975).

follow a growing trend to avoid the imposition of broad fiduciary duties in a close corporation setting and extend it to the limited liability company. For example, in *Riblet Products Corp. v. Nagy*, the Delaware Supreme Court refused to recognize the alteration of the duties of majority shareholders because of the status of the company as a close corporation.[180] In *Nixon v. Blackwell*, the Delaware Supreme Court also declined the opportunity to develop any judicially created rules to protect minority shareholders of a corporation.[181] Further, in *Olsen v. Seifert*, a Massachusetts court, applying Delaware law, refused to provide relief to a minority shareholder when the defendant had sought to repurchase the minority shareholder's shares at the original purchase price, rather than at their higher fair market value, on the eve of a favorable merger.[182] The Massachusetts court indicated that Delaware would review a particular transaction for overall fairness where it is alleged that a controlling shareholder has benefited excessively, but would not impose broad fiduciary duties on close corporations.[183]

The trend to curtail broad judicial statements of fiduciary duties is also reflected in the statutory language of a number of LLC statutes. Although some LLC statutes are silent on the subject,[184] others contain express statements that the fiduciary duty of members and/or managers is limited to certain duties specifically designated in the statute. For example, some statutes include duties to account for certain benefits, to refrain from dealing with the LLC on behalf of an adverse party, or to refrain from competing with the LLC.[185] This approach has been taken in the

---

[180] 683 A.2d 37, 39 (Del. 1996) (involving a lawsuit by a 15% minority shareholder/employee of the corporation who claimed that the majority breached its duty by dismissing him as an employee and indicating that the company's status as a closely held company did not alter the duties of stockholders inter se).

[181] 626 A.2d 1366, 1377 (Del. 1993) (refusing to hold that the board of directors breached their fiduciary duties to non-employee minority shareholders by failing to offer the minority the same liquidity offered to employee-shareholders, who were covered by an ESOP plan and key man life insurance); *see also* Theresa L. Kelly, Nixon v. Blackwell: *Fairness But Not Equality For Minority Shareholders*, 19 DEL. J. CORP. L. 533 (1994) (analyzing the implications of *Nixon v. Blackwell*).

[182] No. 97-6456, 1998 Mass. Super. LEXIS 592, at *15 (Mass. Super. Ct. Aug. 28, 1998).

[183] *Id. But see* VGS Inc. v. Castiel, No. 17995, 2000 Del. Ch. LEXIS 122, at *15 (Del. Ch. Aug. 31, 2000) (holding that two of three LLC board members breached their duty of loyalty to the LLC, its investors, and a third board member because the two board members failed to notify the third board member of the proposed merger, which had the effect of divesting the third member of his majority control of the business).

[184] *See* ALASKA STAT. § 10.50.130 (Michie 1998); ARK. CODE ANN. § 4-32-402 (Michie 1996); DEL. CODE ANN. tit. 6, § 18-1101 (1999); D.C. CODE ANN. § 29-1321 (1996).

[185] *See* ALA. CODE § 10-12-21(f) (1999) (indicating that the duty of loyalty is limited to accounting and holding as trustee all property, profit, or benefit, including the appropriation of the LLC's opportunity, refraining from dealing with the LLC as or on behalf of a party having an adverse interest, and refraining from competing with the LLC); CAL. CORP. CODE § 17153 (West Supp. 2001) (providing that the fiduciary duties a manager owes to the LLC and to the members are those of a partner to a partnership); HAW. REV. STAT. § 428-409 (1998) (indicating that the only fiduciary duties a member owes are those

ULLCA[186] and is mirrored in the RUPA.[187] A number of states adopt a duty to account to other members for profits made from LLC transactions.[188]

LLC statutes also vary in the extent to which they permit contractual modifications to the statutory standard for fiduciary duties. Some LLC statutes contain express restrictions on the right to modify the members' or managers' standards of conduct contractually in the articles of organization or operating agreement.[189] The extent to which courts will respect contractual limitations on fiduciary duties remains to be developed on a case-by-case basis.[190]

The growing movement away from the judicial implication of broad fiduciary duties makes it increasingly important for the minority LLC member to obtain express contractual protections and for the LLC statute itself to provide a variety of minimum statutory protections, including a default buy-out right and possibly a dissolution remedy. Practitioners' sentiments against judicial monitoring of private enterprises are dramatically changing the business law landscape and should not be taken lightly. As noted by Donald Weidner:

> [V]ague broad statements of a powerful duty of loyalty cause too much uncertainty .... [E]ven if there are no bad holdings, overly broad judicial language has left practitioners uncertain about whether their negotiated agreements will be voided .... [A]ttorneys and their clients want to be able to negotiate transactions, reduce their agreements to writing, and have some com-

---

specified in the statute and that the duty of loyalty is limited to accounting for property, profits, or benefits, refraining from dealing with the LLC as an adverse party, and refraining from competing with the LLC); 805 ILL. COMP. STAT. 180/15-3 (1998) (specifying the fiduciary duties in subparagraphs, including the duties to account for certain benefits, to act fairly when acting on behalf of a party with an adverse interest to the company, and to refrain from competing with the company); MONT. CODE ANN. § 35-8-310 (1999) (specifying the member's duties to account for certain benefits, to refrain from dealing with the company on behalf of a person having an interest adverse to the company, to refrain from competing with the company, and to refrain from engaging in grossly negligent conduct).

[186] UNIF. LTD. LIAB. CO. ACT § 409 (1996).

[187] REVISED UNIF. P'SHIP ACT § 404 (1997).

[188] *See* HUMPHREYS, *supra* note 155, § 4-33 (1999).

[189] *See* ALA CODE § 10-12-21 (1999) (providing that, effective January 1, 2001, an operating agreement may modify the member's or manager's duties but may not unreasonably restrict rights to information or records or eliminate the duty of loyalty); COLO. REV. STAT. ANN. § 7-80-108 (West 1999) (providing that an operating agreement may not unreasonably restrict access to books and records, unreasonably reduce the duty of care, or eliminate good faith requirements); D.C. CODE ANN. § 29-1320 (1996) (indicating that liability may be limited or eliminated in the articles of organization, except if the manager or member engaged in willful misconduct); FLA. STAT. ch. 608.423 (2000) (providing that the agreement may not unreasonably restrict the right to information or records, the duty of loyalty, or the duty of care).

[190] *See* Elf Atochem v. Jaffari, 727 A. 2d 286, 291–92 (Del. 1999) (permitting LLC members, like limited partners, the broadest possible discretion in drafting their contractual agreements).

> fort that those agreements will not be undone by "fuzzy" notions of fiduciary duties.[191]

Default buy-out rights take on increased importance where the prevailing judicial philosophy opposes proactive judicial intervention in the resolution of disputes among owners of private business enterprises and/or where the LLC contains express statutory language purporting to limit the imposition of fiduciary duties.

It is incumbent upon legislators to develop statutory parameters of conduct and statutory protections and remedies in light of the growing pressures on the judiciary to play a less active role in monitoring private business ventures.

I. What Enforcement Mechanisms Will Be Developed Where an LLC Member or Manager Has Breached Fiduciary Duties and the Standard of Care?

Considerable variation may be found among states as to how an LLC owner should proceed if a manager or member violates his or her fiduciary duty and/or the applicable standard of care. Should a suit be instituted for a direct action alleging damages to the individual plaintiff as an LLC owner, or should a derivative lawsuit be initiated alleging injury to the LLC as an entity? Alternatively, should something akin to an accounting be sought, as is the case in partnership disputes where a partner alleges that another partner has breached a fiduciary duty? The minority LLC owner may be entering new territory with regard to both substantive *and* procedural issues concerning fiduciary duties and the standard of care.

In the corporate world, minority shareholders may be entitled to institute a shareholder's derivative suit, an action that permits the individual shareholder to sue on behalf of the corporate entity to remedy or prevent harm to the corporation.[192] The shareholders' derivative lawsuit is instituted in a broad array of situations, including, but not limited to, cases asserting a breach of the directors' duty of care and/or duty of loyalty, and claims of negligence, mismanagement, self-dealing, excessive compensation, or usurpation of corporate opportunities.[193]

The distinction between a direct action and a derivative action turns on whether the injury alleged is one to the corporation or to the share-

---

[191] Weidner, *Three Policy Decisions Animate Revision of Uniform Partnership Act*, *supra* note 37, at 462 (discussing standards of conduct in connection with the revision of the UPA).

[192] *See* O'NEAL & THOMPSON, *supra* note 20, § 9.22 (discussing the basic features of a shareholders' derivative action).

[193] *Id.*

holder individually.[194] However, the American Law Institute ("ALI") has observed that the concept of a corporate injury that is distinct from an injury to the shareholders is essentially a fiction in the case of the closely held corporation.[195] Additionally, a shareholders' derivative action is limited somewhat in the context of a private company because the recovery obtained in the shareholders' derivative suit goes to the corporation, and the payment may therefore become subject to the control of the majority shareholder who has committed the misconduct.[196] It therefore recommends that in the case of a private company, the court should have discretion to treat an action raising derivative claims as a direct action if to do so would not: (1) unfairly expose the corporation or defendants to a multiplicity of actions; (2) materially prejudice the interests of creditors of the corporation; or (3) interfere with a fair distribution of the recovery among interested parties.[197]

Several states authorize the right to bring an equitable action for an accounting, although at least one state permits the operating agreement to provide to the contrary.[198] In addition, the majority of LLC statutes expressly provide a right to bring a derivative action or an action in the name of the LLC, although there is considerable variation in the statutes regarding the circumstances that permit the derivative suit.[199] The Ameri-

---

[194] Aurora Credit Servs., Inc. v. Liberty W. Dev., Inc., 970 P.2d 1273, 1280 (Utah 1998) (indicating that derivative suits are those which enforce any right belonging to a corporation, but in a direct action, the plaintiff can prevail by showing an injury to him/herself that is distinct from the injury suffered by the corporation).

[195] 2 AM. LAW INST., PRINCIPLES OF CORPORATE GOVERNANCE: ANALYSIS AND RECOMMENDATIONS § 7.01(d) cmt. e (1994) (indicating that in light of the difficulty in discerning a difference between an injury to the corporation and an injury to the shareholders individually in a private company, the court should have discretion to allow direct actions by a minority shareholder).

[196] *See id.*

[197] *Id.* § 7.01(d).

[198] *See* LA. REV. STAT. ANN. § 12:1319(B)(3) (West 1994) (authorizing a member to demand a formal accounting unless the operating agreement or articles of organization provide otherwise). Other states that authorize an accounting include Colorado, Hawaii, Michigan, Oklahoma, and Pennsylvania. *See* COLO. REV. STAT. ANN. § 7-80-712 (West 1999); HAW. REV. STAT. § 428-410 (Supp. 1998); 805 ILL. COMP. STAT. 180/10-15-20 (1998); MICH. COMP. LAWS ANN. § 450.4503(4) (West Supp. 2000); OKLA. STAT. tit. 18, § 2021(B)(3) (2000); 15 PA. CONS. STAT. ANN. § 8911 (West 1995 & Supp. 2000). Though Delaware does not contain express authorization for an accounting, it does permit an action to be brought in the Court of Chancery to enforce a right to certain information and records. DEL. CODE ANN. tit. 6, § 18-305 (1999). For further discussion of accounting actions and other remedies for LLCs, see JAMES R. BURKHARD, PARTNERSHIP & LLC LITIGATION MANUAL: ACTIONS FOR ACCOUNTING & OTHER REMEDIES §§ 5.02–8.07 (1995).

[199] *See* ALA. CODE § 10-12-25 (1999); ALASKA STAT. § 10.50.735 (Michie 1998); ARIZ. REV. STAT. ANN. § 29-831 (West 1998); ARK. CODE ANN. § 4-32-1102 (Michie 1996); CAL. CORP. CODE § 17500 (West Supp. 2001); DEL. CODE ANN. tit. 6, §§ 18-1001, 18-1002 (1999); D.C. CODE ANN. § 29-1343 (1996); FLA. STAT. ch. 608.601 (2000); GA. CODE ANN. § 14-11-801 (1994); IDAHO CODE § 53-659 (Michie 1994); MASS. GEN. LAWS ch. 156C, § 56 (1998); N.J. STAT. ANN. § 42:2B-60 (West Supp. 2000); 15 PA. CONS. STAT. ANN. §§ 8991–8992 (West 1995); *see also* BURKHARD, *supra* note 198, § 5.07 (discussing derivative actions in LLCs).

can Bar Association's Prototype Limited Liability Company Act contains a provision permitting one or more members of an LLC to institute a suit by or against an LLC in its own name.[200]

Because of the closely held nature of the LLC, there may be little practical difference between a direct suit and a derivative suit. Therefore, the ALI's analysis of derivative and direct suits with respect to close corporations may well apply to privately owned LLCs. Some courts have been willing to recognize direct actions in close corporations for a breach of fiduciary duty and have not imposed derivative pleading requirements on close corporations.[201] Courts are increasingly recognizing the unique nature of close corporations and the appropriateness of direct actions for damages resulting from a breach of fiduciary duty.[202] This view has been endorsed by the ALI.[203] However, the recognition of direct actions for a breach of fiduciary duty has not been universal in the case of close corporations.[204]

A direct action should not be barred in an LLC, particularly where it would be the most efficient manner of resolving the dispute.[205] Statutory silence on remedies for a violation of the standard of care or the duty of loyalty should not be interpreted as barring direct actions by LLC members. In the absence of an express statutory prohibition, LLC participants should be entitled to institute whatever forms of action that will lead to a resolution of the dispute with minimum time and expense.[206]

J. The Importance of the Buy-Out Right Amidst Substantive and Procedural Uncertainties in Judicial Remedies

Default buy-out rights, and perhaps other statutory remedies that would become operative in the event of disputes between owners, are

---

[200] PROTOTYPE LTD. LIAB. CO. ACT § 1101 (1993).

[201] A summary of the trend to avoid a derivative pleading requirement in the context of closely held corporations is contained in *Wessin v. Archives Corp.*, 581 N.W.2d 380 (Minn. Ct. App. 1998), which indicates that Minnesota law recognizes a distinction between large publicly held corporations and small, closely held companies and supports the rule annunciated by the ALI permitting minority shareholders a direct action based on the view that there is little possibility of a disinterested board of directors in a close company and little possibility of multiple lawsuits.

[202] *See* O'NEIL & THOMPSON, *supra* note 20, § 9.22 (observing that courts are permitting direct suits for the misuse of corporate authority, misappropriation of corporate assets, and selling corporate assets too cheaply).

[203] 2 AM. LAW INST., PRINCIPLES OF CORPORATE GOVERNANCE: ANALYSIS AND RECOMMENDATIONS § 7.01 & Reporter's Note, at 17–33 (1994). It is noteworthy that some states do not recognize a direct action by minority shareholders of private companies. *See* Bagdon v. Bridgestone/Firestone, Inc., 916 F.2d 379, 384 (7th Cir. 1990) (indicating that Delaware does not follow the ALI approach to derivative actions).

[204] *See Bagdon*, 916 F.2d at 383–84 (disallowing a direct suit in the case of a close corporation).

[205] *See* BURKHARD, *supra* note 198, § 5.04 (discussing possible forms of action by aggrieved LLC members).

[206] *See id.*

particularly important for LLC members precisely because LLC members are swimming in uncharted procedural, as well as substantive, waters. The uncertainty regarding standards of conduct for LLC members is compounded by the lack of authority on the process for asserting breaches of fiduciary duties and violations of the standard of care. The elimination of a default buy-out right in LLC statutes makes it impossible for an LLC member without an operating agreement to extricate himself from the LLC without resort to litigation, in which the process and outcome may be unclear. At least until precedents are established to create the parameters of acceptable member and manager conduct, an LLC participant should have the protection of a default exit right, and perhaps other statutory protections as well. As Steven C. Bahls points out:

> [Majority owners] control the books and records of the company and have been known to alter them. Likewise, majority shareholders, with the benefit of hindsight, and the Business Judgment Rule, often re-characterize a questionable transaction or find new and acceptable justifications for the transaction. Ascertaining whether the shareholder-manager has been benefited is like putting Humpty Dumpty back together again . . . .[207]

Considering the relative imbalance of power between majority and minority LLC owners, and the uncertain prospects for litigation concerning a breach of fiduciary duty and duty of care, LLC statutes should retain a default buy-out rule to protect minority LLC owners who may have lacked the bargaining power or the foresight to obtain reasonable buy-out protection in an operating agreement.

III. Statutory Provisions to Facilitate or Combat Minority Oppression: Assessing the Elimination of Buy-Out Rights in the Context of the LLC Statute as a Whole

Prior to modifying any fundamental rights of the LLC owner, legislators should carefully analyze the usage of LLCs within the state. It is also important to analyze the provisions of the LLC statute as a whole to assess whether it contains rules that are likely to establish conditions under which majority oppression could flourish, or alternatively contains provisions that would likely deter majority over-reaching.[208] The elimination of a default buy-out right is arguably more harmful to minority members where the LLC statute lacks minority protections other than

---

[207] *See* Bahls, *supra* note 132, at 293.

[208] *See* Gevurtz, *supra* note 137, at 508 (discussing the structural provisions in LLC statutes that facilitate or frustrate squeeze-outs and freeze-outs).

buy-out rights than where the LLC statute contains a full range of built-in minority protections.

Several statutory provisions have been identified as important in enhancing a majority owner's capacity to freeze-out a minority participant.[209] Such provisions include: (1) the capacity of majority LLC owners to change salaries of other members; (2) the ability to amend the operating agreement or articles affecting members' fundamental rights; (3) the ability to sell all of the business's assets notwithstanding the objection of a minority; (4) the power to approve mergers with less than unanimous vote; and (5) the right to expel members from the LLC.[210] The elimination of a default statutory buy-out right would be particularly problematic in states with such LLC provisions that tend to facilitate majority squeeze-outs.

One of the most important mechanisms for achieving squeeze-outs from a close corporation is the reduction or elimination of the minority's salary or the payment of excessive compensation to the majority.[211] Some LLC statutes provide fixed guidelines for making interim distributions to LLC members based on profit sharing percentages or capital accounts, unless otherwise provided in the LLC operating agreement.[212] At least one state requires unanimous approval for distributions if the operating agreement fails to specify the times or events for distributions prior to dissolution.[213] Some jurisdictions, including the District of Columbia and Delaware, fail to provide a default rule for interim distributions in the absence of an operating agreement, thus creating the opportunity for a majority to make distributions in an unfair manner to the detriment of the minority.[214]

---

[209] *See id.*

[210] *See id.*

[211] *Cf.* Bernard v. Buttner, No. CV95 032 30 62, 1995 Conn. Super. LEXIS 3048, at *8 (Conn. Super. Ct. Oct. 30, 1995) (alleging the defendant breached her fiduciary duty in part by paying herself excessive compensation); Barth v. Barth, Jr., 659 N.E.2d 559, 560–61 (Ind. 1995) (alleging the majority shareholder breached his fiduciary duty by paying excessive salaries to himself and his immediate family members, providing corporate services to family members without compensation, lowering dividend levels, and appropriating corporate funds for personal investments); Wessin v. Archives Corp., 581 N.W.2d 380, 381–82 (Minn. Ct. App. 1998) (alleging the majority shareholder paid his wife a salary even though she performed no services for the company); Kelly v. Axelsson, 687 A.2d 268, 270–71 (N.J. Sup. Ct. 1997) (seeking dissolution where the majority shareholder paid excessive salaries to himself and to his children); Crosby v. Beam, 548 N.E.2d 217, 224 (Ohio 1989) (upholding a minority shareholder's right to bring an action for a breach of fiduciary duty against controlling shareholders in part for paying themselves unreasonable compensation); Jorgensen v. Water Works, Inc., 582 N.W.2d 98, 105–06 (Wis. Ct. App. 1998) (recognizing a direct action for breach of fiduciary duty where defendants allegedly paid themselves excessive compensation instead of dividends to the detriment of the plaintiff).

[212] *See, e.g.*, ALA. STAT § 10-12-29 (Michie 1998); ALASKA CODE § 10.50.300 (1998); ARIZ. REV. STAT. ANN. § 29-703 (West 2000).

[213] *See* GA. CODE. ANN. § 14-11-404 (1998).

[214] DEL. CODE ANN. tit. 6, § 18-601 (1999); D.C. CODE ANN. § 29-1326 (1998).

It has been suggested that LLC statutes should prohibit interim distributions unless provided for in the operating agreement as a strategy for deterring manipulation of distributions by majority owners.[215] Default rules that either prohibit interim distributions or provide reasonable guidelines for determining such distributions would be particularly helpful if default buy-out rights were to be eliminated.

The capacity to amend the operating agreement, sell assets, or approve mergers by majority vote also has been regarded as a potential weapon in the hands of a majority owner.[216] While some states provide a default rule requiring unanimity for amending the operating agreement,[217] some business decisions, including the approval of a merger, may only require a majority vote unless otherwise specified in the operating agreement.[218]

Historically, governance by majority rule, particularly when coupled with the business judgment rule, has created a fertile ground for minority oppression. The LLC member without effective contractual protection could be as vulnerable as the minority shareholder of a close corporation if a default buy-out right is not provided.

The elimination of default buy-out rights or other restrictions on withdrawals is most troubling in states with LLC provisions that lack other statutory mechanisms for protecting minority owners. For example, the position of the minority LLC owner is seriously compromised in Delaware, which prohibits the LLC member from resigning prior to the dissolution and winding up of the LLC unless the LLC agreement provides to the contrary.[219] The Delaware LLC statute lacks even the most basic of provisions that normally protect minority owners. It fails to contain any minimum standards of good faith or fiduciary duty, but rather leaves the parties free to expand or contract the member's or manager's duties and liabilities.[220] It is silent regarding the manner in which an operating agreement may be amended,[221] possibly leaving the door open to adverse majority-driven changes in the fundamental rights and duties of the minority.[222] Further, it is conceivable that under the Delaware LLC

---

[215] *See* Gevurtz, *supra* note 208, at 509–10.

[216] *See id.* at 511–12.

[217] *See* ALA. CODE ANN. § 10-12-24 (Michie 1998); ALASKA STAT. § 10.50.095 (Michie 1998); ARIZ. REV. STAT. ANN. § 29-681 (West 2000); CAL. CORP. CODE § 17103 (WEST 1999); GA. CODE ANN. § 14-11-308 (1998).

[218] ALASKA STAT. § 10.50.150 (1998); ARK. CODE ANN. § 4-32-1202 (1997); DEL. CODE ANN. tit. 6, §§ 18-209, 18-210 (1998); CAL. CORP. CODE § 17551 (West 1999); D.C. CODE ANN. § 29-1340 (1998); FLA. STAT. ANN. §§ 608.4381, 608.4384 (West 1999).

[219] DEL. CODE ANN. tit. 6, § 18-603 (1999).

[220] *See id.* § 18-1101(c)(2).

[221] *See id.* § 18-202 (addressing the procedure for amending the certificate of formation but not containing special guidelines for amendment of the operating agreement).

[222] The lack of restrictions on amending the operating agreement raises the question of whether, for instance, a majority LLC owner could conceivably modify an operating agreement to remove a supermajority voting requirement that was initially contained in the

statute, a merger could be approved by a majority of the LLC members, leaving the minority without dissenter's rights, because the Delaware LLC statute offers only contractual appraisal rights, and not all minority owners may possess the foresight, resources, and/or negotiating power to obtain contractual appraisal rights in the operating agreement.[223]

The nature and scope of minority protections vary considerably among state LLC statutes, and the potential adverse impact of restrictions on buy-out rights will vary accordingly from state to state. For example, although California has eliminated the right to payment for a member's interest upon withdrawal unless the operating agreement provides otherwise, the California LLC statute contains several important minority protections.[224] In California, amendments to the articles of organization or the operating agreement require unanimous approval.[225] Actions approved at meetings other than by unanimous approval are valid only if the general nature of the proposal was stated in the notice of the meeting or in a written waiver of notice.[226] The California LLC statute also prohibits the indemnification of a person who has committed a breach of fiduciary duties[227] and provides dissenter's rights.[228] Finally, the California LLC statute expressly authorizes an action for dissolution in a variety of important circumstances, including when there is deadlock or internal dissension or when those in control have knowingly countenanced persistent and pervasive fraud, mismanagement, or abuse of authority.[229]

Like the California LLC law, New York's LLC statute contains several important features that are designed to deter member or manager misconduct. The New York statute establishes non-waivable standards of conduct,[230] contains specific requirements concerning meetings of mem-

---

agreement. A similar issue has been raised in a corporate context. *See* McNamara v. Frankino, 744 A.2d 988 (Del. 1999), *aff'g* Frankino v. Gleason, No. 17399, 1999 Del. Ch. LEXIS 218 (Del. Ch. Nov. 12, 1999) (upholding a Chancery Court opinion that permitted a 55% shareholder to use his majority consent to amend a bylaw requiring an 80% supermajority vote to expand the size of the board. The shareholder eliminated the supermajority requirement, expanded the board size, and thus gained loyal new board members).

[223] *See* DEL. CODE. ANN. tit. 6, § 18-210 (1998) (recognizing contractual appraisal rights); *id.* § 18-209 (1998) (providing rules for the majority approval of mergers and consolidations). Undoubtedly, courts will be called upon to rescind mergers accomplished in bad faith. *See* VGS Inc. v. Castiel, No. 17995, 2000 Del. Ch. LEXIS 122, at *15 (Del. Ch. Aug. 31, 2000) (rescinding a merger and holding that two of three LLC board members breached their duty of loyalty to the LLC, its investors, and a third board member because the two board members failed to give notice of the proposed merger to the third board member whose majoirty control was eliminated by virture of the merger).

[224] CAL. CORP. CODE § 17252 (West 1999).

[225] *Id.* § 17103(2).

[226] *Id.* § 17104(g). Notice requirements can be critically important in the context of a power struggle between majority and minority participants.

[227] *Id.* § 17155.

[228] *Id.* §§ 17600–17613.

[229] *Id.* § 17351.

[230] 32A N.Y. LTD. LIAB. CO. LAW § 417(a) (McKinney Supp. 2000).

bers and notices of meetings,[231] and also provides defined standards of care.[232] Under the New York LLC statute, the operating agreement cannot contain a provision that eliminates or limits the liability of a manager in cases involving bad faith, intentional misconduct, knowing violations of law, or other cases involving personal gain.[233] The statute permits members to dispense with meetings, prior notice, and actual votes only if written consents are obtained of not less than the minimum number of votes that would be necessary to authorize the underlying action.[234] A vote of a majority in interest is required to approve a dissolution of the LLC, approve the sale or other transfer of substantially all of the LLC's assets, or approve a merger or consolidation with another LLC.[235] In addition, a provision in the operating agreement that contains a specific voting requirement may not be amended without the vote of at least the same percentage in interest.[236] This restriction prevents majority owners from using their majority power to circumvent operating rules that require approval of more than a majority vote.[237]

In both New York and California, the protections provided to minorities may be sufficient to protect an LLC member even in the absence of a buy-out right. The potential adverse impact of restrictions on buy-out rights, then, must be assessed on a case-by-case basis, and the analysis must consider the extent to which each individual statute contains safeguards *other than* buy-out rights that are designed to deter majority abuse.[238]

---

[231] *See id.* § 403.
[232] *See id.* § 409.
[233] *Id.* § 417(a)(1).
[234] *Id.* § 407(a).
[235] *Id.* § 402(d).
[236] *Id.* § 402(e).
[237] *Cf.* McNamara v. Frankino, 744 A.2d 988 (Del. 1999) (permitting a majority shareholder to use his simple majority vote to change an article in the company's bylaws that required a supermajority vote of 80% to expand the size of the Board of Directors, thus enabling the 55% shareholder to amend the bylaws so that he could expand the size of the Board of Directors to appoint board members loyal to him).
[238] For example, although both Rhode Island and Oklahoma have eliminated default exit rights, the Rhode Island LLC statute lacks the important minority protections contained in the Oklahoma statute. *Compare* R.I. GEN. LAWS § 7-16-21 (Supp. 1999) (providing for majority of capital approval for dissolutions, sales and other transfers, and mergers or consolidations) *with* OKLA. STAT. tit. 18, § 2020 (1998) (indicating that unless otherwise provided in the articles of organization or written operating agreement, unanimous vote is required to approve: a dissolution; an amendment to the articles of organization or amendment to a written operating agreement; a term that reduces the existence of the LLC; a term that reduces the required vote of members to approve a dissolution, merger, sale, or other disposition of substantially all the assets; a term that permits a member to voluntarily withdraw; or a term that reduces the required vote of members to approve an amendment to the articles of organization or written operating agreement).

A. *Innovative Strategies that Restrict Default Exit Rights but Provide Additional Statutory Safeguards: Closing the Door but Opening a Window*

At least one state, Iowa, has taken an innovative approach to default exit rights by locking in LLC members, but simultaneously giving an exit right to dissenting members if the majority has changed the LLC articles or agreement in a manner that adversely affects the dissenting member's fundamental rights or preferences.[239]

The Iowa approach appears to close the exit door, while leaving a window open to the minority where the majority has adversely changed the minority's fundamental rights under the LLC articles or operating agreement. It represents one of several interesting strategies that could be considered in states where there is strong support for using the LLC as an estate-planning vehicle.

---

[239] IOWA CODE ANN. § 490A.704A (West 1999) provides in part:

2. A member may resign or withdraw from a limited liability company only at the time or upon the happening of an event specified in an operating agreement and pursuant to the operating agreement.

3. Unless an operating agreement provides otherwise, a member may not resign or withdraw from a limited liability company prior to the dissolution and winding up of the limited liability company. However, if the articles of organization or an operating agreement do not specify the time or the events upon the happening of which a member may resign or withdraw, a member may resign or withdraw from the limited liability company in the event any amendment to the articles of organization or operating agreement that is adopted over the member's written dissent adversely affects the rights or preferences of the dissenting member's membership interest in any of the ways described in paragraphs "a" through "e." A resignation or withdrawal in the event of such dissent and adverse effect is deemed to have occurred as of the effective date of the amendment, if the members give notice to the limited liability company not more than sixty days after the date of the amendment. In valuing the member's distribution pursuant to this subsection, any depreciation in anticipation of the amendment shall be excluded. An amendment that does any of the following is subject to this subsection:

a. Alters or abolishes a member's right to receive a distribution.

b. Alters or abolishes a member's right to voluntarily withdraw or resign.

c. Alters or abolishes a member's right to vote on any matter, except as the rights may be altered or abolished through the acceptance of contributions or the making of contribution agreements.

d. Alters or abolishes a member's preemptive right to make contributions.

e. Establishes or changes the conditions for or consequences of expulsion.

4. A member withdrawing under this section is not liable for damages for the breach of any agreement not to withdraw.

B. Other Statutory Protections: Unreasonable Reductions in Standards of Conduct and Remedies for Illegal, Fraudulent, or Unfairly Prejudicial Conduct

In light of the growing restrictions on the withdrawal and distribution rights of LLC members, the intended use of the LLC as a vehicle for small, informal business ventures, the considerable uncertainties surrounding the judicial interpretation of the duty of loyalty and standard of care, and the infancy of the procedural mechanisms for asserting a breach of fiduciary duty or standard of care, this Article advocates the enactment of at least two major statutory protections for minority LLC owners, including: (1) a prohibition on contractual provisions that unreasonably restrict or reduce fiduciary duties and the standard of care; and (2) a mechanism for seeking a dissolution or buy-out if there is deadlock or if the managers or members in control of the company have engaged in certain types of misconduct delineated in the statute, i.e., illegal or fraudulent conduct, unfairly prejudicial conduct, or possibly oppressive conduct. These two protections have been included in the ULLCA.[240] This dissolution right should be seriously considered particularly in states that are planning to eliminate default buy-out rights.

LLC statutes in a number of states already expressly prevent contractual provisions that unreasonably restrict the right to information or records or unreasonably eliminate the duty of loyalty. Although Delaware contains no limitation on the restriction or expansion of fiduciary duties,[241] several other states contain prohibitions upon contractual reductions in the duty of loyalty or duty of care, including Alabama,[242] Colorado,[243] and Florida.[244]

Further, a growing number of states provide the remedy of a judicial dissolution upon a showing of certain majority misconduct. For instance, Minnesota's LLC statute provides special remedies to closely held LLC owners where: (1) management is deadlocked; (2) "those in control of a limited liability company have acted fraudulently, illegally, or in a man-

---

[240] *See* UNIF. LTD. LIAB. CO. ACT § 103 (1996) (providing that the operating agreement may not unreasonably reduce the duty of loyalty or the duty of care); *id.* § 801 (enumerating the occurrence of events upon which the LLC is dissolved, including where "the managers or members in control of the company have acted, are acting, or will act in a manner that is illegal, oppressive, fraudulent, or unfairly prejudicial to the petitioner").

[241] *See* DEL. CODE ANN. tit. 6, § 18-1101 (1998).

[242] *See* ALA. CODE § 10-12-21 (Michie 1998) (effective January 1, 2001, an operating agreement may not unreasonably restrict the right to information or records or abolish the duty of loyalty).

[243] *See* COLO. REV. STAT. ANN. § 7-80-108 (1999) (prohibiting the operating agreement from unreasonably restricting access to books and records, unreasonably reducing the duty of care, or eliminating the requirements of good faith).

[244] FLA. STAT. ANN. § 608.423 (1999) (indicating that the operating agreement may not unreasonably restrict the right to information and records, abolish the duty of loyalty, or unreasonably diminish the duty of care).

ner unfairly prejudicial toward one or more members in their capacities as members, governors, or managers, or as employees of a closely held limited liability company"; or (3) "assets are being misapplied or wasted."[245] Similar judicial dissolution provisions for deadlock and/or majority misconduct are provided in the LLC statutes of Alaska, California, Florida, Idaho, Kansas, Maine, and Ohio.[246]

Dissolution remedies are increasingly important as states amend their LLC statutes to eliminate or otherwise restrict the LLC member's withdrawal and distribution rights. Corporate statutes have long contained the remedy of dissolution in the event of illegal, fraudulent, or oppressive conduct.[247]

Minority shareholder oppression statutes have attracted considerable criticism in the corporate context, and the adoption of oppression remedies in LLC statutes would undoubtedly be greeted with skepticism by some practitioners and scholars. Professor Larry Ribstein has argued in favor of retaining default exit rights in at least one business entity precisely to ward off the adoption of open-ended LLC oppression remedies.[248]

The criticism of existing dissolution remedies is well-founded in two major respects: (1) the statutes are vague with regard to what type of conduct constitutes "oppressive conduct"; and (2) most of the statutes do not identify remedies other than dissolution.[249]

A number of courts have attempted to define "oppressive conduct" in terms of the defeat of the minority shareholder's reasonable expectations.[250] Under this approach, courts have looked to the shareholder's rea-

---

[245] MINN. STAT. ANN. § 322B.833 (West Supp. 2000).

[246] *See* ALASKA CODE § 10.50.410 (1998); CAL. CODE § 17351 (1999); FLA. STAT. ANN. § 608.449 (West 1999); IDAHO CODE § 53-643 (1998); KANS. STAT. ANN. § 17-7629 (West 1997); ME. REV. STAT. ANN. tit. 31, § 702 (West 1997); OHIO REV. STAT. ANN. § 829-785 (West 2000).

[247] *See* MODEL BUS. CORP. ACT ANN. § 14.30 (3d ed. 1999) (indicating in the statutory summary that virtually all states provide for involuntary dissolutions in defined circumstances).

[248] Ribstein, *supra* note 125, at 340–41 (indicating that it was restricted exit in corporations that led to the problem of open-ended oppression remedies for close corporations and that at least one type of statute should be kept "safe" for non-family firms that do not have tax reasons for restricting member exit).

[249] *See* Sandra K. Miller, *How Should U.K. and U.S. Minority Shareholder Remedies for Unfairly Prejudicial or Oppressive Conduct Be Reformed?*, 36 AM. BUS. L.J. 579, 612–22 (1999).

[250] *See* Michaud v. Morris, 603 So. 2d 886, 888 (Ala. 1992) (stating that defeated expectations alone do not always show there has been oppressive conduct); Smith v. Leonard, 876 S.W.2d 266, 272–73 (Ark. 1994) (alleging breach of fiduciary duty and violation of minority shareholders' expectations); Pedro v. Pedro, 463 N.W.2d 285, 287 (Minn. Ct. App. 1990) (awarding plaintiff compensation for lost wages he reasonably expected to earn as an employee and stockholder of the company); Bonavita v. Corbo, 692 A.2d 119, 128 (N.J. Super. Ct. Ch. Div. 1996) (indicating that decedent's widow had reasonable expectations of obtaining some form of financial benefit or compensation and ordering the majority to purchase the widow's stock); Muellenberg v. Bikon Corp., 669 A.2d 1382, 1390 (N.J. 1996) (discussing the reasonable expectation test and ordering the majority to sell his

sonable expectations in acquiring an interest in the company, which sometimes entails being a key employee, having a voice in business decisions, or obtaining a reasonable return on the investment.[251] It may be argued, however, that the reasonable expectations test is vague and provides insufficient guidance to the business and legal communities.[252]

A major shortcoming of the reasonable expectations standard is its singular emphasis on the expectations of the dissatisfied participant, rather than upon the conduct of the majority. If the applicable oppression statute provides a judicial remedy for "oppression" or "unfairly prejudicial" conduct, the primary focus of the judicial inquiry should be upon the *conduct of the majority* rather than upon the *perceptions of the dissatisfied participant.*

While the definition of "oppression" is best left to judicial construction on a case-by-case basis, a court should consider a variety of factors in evaluating the defendant's conduct. In particular, this analysis should focus on those patterns of behavior that are typically indicative of op-

---

stock to the minority shareholder); Brenner v. Berkowitz, 634 A.2d 1019, 1020 (N.J. 1993) (involving a minority shareholder's allegations that the majority was violating the sales tax, underreporting income, and committing other wrongful acts); Kelly v. Axelsson, 687 A.2d 268, 274 (N.J. Super. Ct. App. Div. 1997) (adhering to the reasonable expectations approach and indicating that failure to maintain an accurate accounting system, resulting in the inability to verify the underreporting of income, could constitute oppressive conduct); Muscarelle v. Castano, 695 A.2d 330, 331 (N.J. Super. Ct. App. Div. 1997) (applying the reasonable expectations test to a dispute among family members who were partners); Exadaktilos v. Cinnaminson Realty Co., Inc., 414 A.2d 994, 995 (N.J. Super. Ct. App. Div. 1980), *aff'g* 400 A.2d 554 (N.J. Super. Ct. Law Div. 1979) (concluding that the plaintiff's reasonable expectations were not defeated in light of the fact that the plaintiff had not gotten along with others, had not learned the restaurant business, and had quit on more than one occasion); *In re* Matter of Kemp & Beatley Inc., 473 N.E.2d 1173, 1175 (N.Y. 1984) (involving two minority shareholders of a table linen manufacturer who sought judicial dissolution on the grounds that the conduct of the majority was fraudulent and oppressive); Foster v. Foster Farms, Inc., 436 S.E.2d 843, 847 (N.C. Ct. App. 1993) (ordering a dissolution of a corporation that owned a hog farm where two 50% shareholders could not agree on whether to borrow money and enter the futures market); Balvik v. Sylvester, 411 N.W.2d 383, 384 (N.D 1987) (involving a suit for dissolution by a minority shareholder who was fired from his position); Gee v. Blue Stone Heights Hunting Club, Inc., 604 A.2d 1141, 1144 (Pa. Commw. Ct. 1992) (indicating that oppressive actions refer to conduct that substantially defeats the reasonable expectations of the shareholder); Miller, *A Note on the Definition of Oppressive Conduct by Majority Shareholders: How Can the Reasonable Expectation Standard Be Reasonably Applied in Pennsylvania?*, *supra* note 123, at 73 (discussing advantages and limitations of the reasonable expectations test).

[251] *See* O'NEAL & THOMPSON, *supra* note 20, § 9.29.

[252] Some of the dissolution cases recognize that majority shareholders must be free to exercise business judgments that may adversely affect minority shareholders. *See* Exadaktilos v. Cinnaminson Realty Co., 400 A.2d 554, 562 (N.J. Super. Ct. Law Div. 1979), *aff'd*, 414 A.2d 994 (N.J. Ct. App. Div. 1980) (observing that the majority had not engaged in oppressive conduct where the minority was fired, since the minority shareholder had failed to learn the restaurant business, had alienated other employees, and was responsible for his own exclusion from the business). However, in some cases, the court has failed to place significant emphasis on the scope of the majority's discretion to make business decisions that may adversely affect the minority shareholder. *See, e.g.*, Muellenberg v. Bikon Corp., 669 A.2d 1382, 1383 (N.J. 1996) (directing that the majority shareholders sell their shares to the minority where the majority had oppressed the rights of the minority).

pressive conduct, such as the exclusion of the minority from management, the withholding of dividends or distributions, the payment of excessive salaries, the personal use or wasting of business assets, and the structuring of non-arms-length transactions. The reasonable expectations of the defendant could serve as an additional consideration. Judicial reference to explicit patterns of conduct clarifies the judicial decision-making process and helps to communicate the contours of acceptable conduct to the business community at large.[253]

A further shortcoming of oppression statutes concerns their failure to identify remedies other than a dissolution or buy-out that could be provided. When a dispute develops between the majority and minority, the payment of the fair market value of the minority's interest will not always be the desired remedy and may not sufficiently compensate for the injury suffered. The minority member who has spent years building a business may not want to relinquish the enterprise at any price. The LLC may own customer lists, intangible property, or other unique assets. Some or all of these assets may have been created or enhanced through the efforts of the minority. The minority who actively runs the business will not want to get ousted just as bright prospects appear on the horizon. A fair market value buy-out may not adequately compensate for lost opportunities where the business is on the brink of a major breakthrough, a favorable merger is around the corner, or a pivotal account has just been secured. In technological start-up companies, the organizers of the business may sustain losses for years, but may eventually become quite successful or sell out for a considerable sum of money. The minority who has been ousted just as his or her labors begin to bear fruit may need a remedy other than a "fair market value" buy-out, particularly in the rapidly developing, technologically driven sectors of our economy.

It would be useful for dissolution statutes to provide a list of remedies other than dissolution or buy-out that could be used to mediate a dispute among LLC members. For example, dissolution statutes could list remedies such as an order to perform, prohibit, alter, or set aside an action; the appointment of a receiver; the cancellation or alteration of a provision of the articles of organization or operating agreement; an award of damages; a partition; or an order to produce certain books and records. A variety of these remedies have already been suggested for use by LLC members or managers.[254] Interestingly, the little-used *Model Close Cor-*

---

[253] Kiriakides v. Atlas Food Systems, No. 25244, 2001 S.C. LEXIS 22, at *27–28 (S.C. Jan. 29, 2001) (applying a case-by-case analysis of oppressive conduct with reference to factors that reflect typical patterns of oppresive conduct).

[254] *See* JAMES R. BURKHARD, PARTNERSHIP & LLC LITIGATION MANUAL: ACCOUNTING & OTHER REMEDIES § 8.06–.07 (1995) (providing forms and discussing possible remedies for disputes among partners or LLC participants).

poration Supplement contains a list of a wide range of remedies short of dissolution.[255]

In the LLC context, it might be possible to improve existing dissolution remedies by incorporating ascertainable standards for determining when the remedies would be triggered. Perhaps the statute could provide a list of factors that are indicative of unfairly prejudicial or oppressive conduct. These factors could include opportunistic manipulations of the operating agreement as well as other conduct that typically signals a minority squeeze-out. For instance, the statute could list the following factors as being indicative of unfairly prejudicial conduct:

> a) the enactment of amendments to the operating agreement that unfairly alter or abolish the complainant's rights to receive a distribution;
> b) the enactment of amendments to the operating agreement that unfairly alter or abolish the complainant's rights to voluntarily withdraw or resign;
> c) the enactment of amendments to the operating agreement that unfairly alter or abolish the complainant's rights to vote;
> d) the enactment of amendments to the operating agreement that establish or change the conditions for or consequences of expulsion;
> e) the unreasonable withholding of distributions;
> f) the alteration or elimination of the complainant's role in management and/or as an employee;
> g) the payment of excessive salary or distributions to the majority;
> h) the structuring of transactions between the LLC and the controlling members on a non-arms-length basis;

---

[255] *See* 4 MODEL CLOSE CORPORATION SUPPLEMENT § 42 (3d ed. 1999) (providing for both extraordinary relief consisting of dissolution or buy-out, and ordinary relief including:

> 1) The performance, prohibition, alteration, or setting aside of any action of the corporation or of its share-holders, directors, or officers of or any other party to the proceeding;
> 2) the cancellation or alteration of any provision in the corporation's articles of incorporation or by-laws;
> 3) the removal from office of any director or officer;
> 4) the appointment of any individual as a director or officer;
> 5) an accounting with respect to any matter in dispute;
> 6) the appointment of a custodian to manage the business and affairs of the corporation;
> 7) the appointment of a provisional director (who has the rights, powers and duties of a duly elected) to serve for the term and under the conditions prescribed by the court);
> 8) the payment of dividends;
> 9) the award of damages to any aggrieved party.).

i) the appropriation or wasting of the LLC's assets.

Reference to such typical patterns of misconduct may prove to be quite useful in a case-by-case analysis of oppressive conduct in the context of the LLC.

While it is true that open-ended oppression remedies have created some uncertainty, we should not be deluded into believing that such remedies are not needed in the context of the LLC. As explained by J.A.C. Hetherington and Michael P. Dooley, when speaking of close corporations:

> The emphasis on contractual arrangements reveals a fundamental misunderstanding of close corporations. Whether the parties adopt [a] special contractual arrangement is much less important than their ability to sustain a close, harmonious arrangement over time. The continuance of such a relationship is crucial . . . .[256]

This observation applies with equal force to members of an LLC. No matter how highly negotiated a business deal is, invariably there will be breakdowns in personal relationships over time. LLC laws should anticipate disputes among members and/or managers and should provide mechanisms for their resolution.

IV. Eliminating Default Exit Rights in the Limited Partnership Rather than the LLC: Following the Lead of the National Conference of Commissioners on Uniform State Laws

The National Conference of Commissioners on Uniform State Laws ("NCUSL") has created an array of uniform business entities, including a uniform LLC.[257] Under NCUSL's proposed scheme, the Re-RULPA would eliminate default buy-out rights in the case of limited partnerships, thus making the limited partnership the entity of choice for achieving estate planning goals through the use of discounts resulting from minority ownership and a lack of marketability.[258] In contrast, according to the NCUSL's present framework, owners of LLCs under the ULLCA and partners in a general partnership under the RUPA have unrestricted default buy-out rights unless the LLC or partnership is for a term, in which

---

[256] Hetherington & Dooley, *supra* note 20, at 2 (discussing the problems posed by illiquidity in the close corporation).

[257] *See* Bishop, *supra* note 3, at 265–74 (providing an overview of the dissociation and dissolution rules of the uniform limited liability company).

[258] *See* Proposed Revisions of Unif. Ltd. P'ship Act (1976) with 1985 Amendments § 505 (completion targeted for 2001) (on file with author) (providing that a person has no right to a distribution on account of a dissociation).

case a distribution may not be payable until the passage of the relevant term or undertaking.

NCUSL's scheme of eliminating default exit rights in the limited partnership while retaining them for the LLC and the at-will general partnership is sensible in light of the practical use of these business entities. Limited partnerships tend to be subject to negotiated agreements and are typically the outgrowth of deliberate business and tax planning, as are term general partnerships and term LLCs. Default exit rights and other minority protections are arguably less critical in a negotiated legal environment than in the context of an informal business arrangement that has not been memorialized in a written agreement. The ULLCA and the RUPA strike a sensible balance between the need for minority investor protections and the desire for a flexible legal framework that can be controlled through private contracting with a minimum of judicial intervention.

The ULLCA contains a number of important legal protections for investors. The statute contemplates that some LLC agreements will not be in writing, thus paving the way for informal business relationships.[259] Consistent with the notion that the LLC should serve the needs of small informal businesses, the ULLCA authorizes the creation of "at-will" LLCs—LLCs without a specified term. If the LLC is an at-will company, a member who dissociates from the company is entitled to be paid the fair market value of his or her interest, subject to offsetting damages for wrongful dissociation.[260] Thus, the ULLCA retains traditional general partnership-like buy-out rules.

The ULLCA also contains an express duty of loyalty as well a duty to refrain from grossly negligent, reckless, or intentional misconduct or a knowing violation of law.[261] While the statute bestows the freedom to develop an individualized LLC agreement, it sets parameters regarding variations in standards of conduct and fundamental rights of members. The operating agreement may not unreasonably restrict rights to information or access to records, and may not eliminate the duty of loyalty or unreasonably reduce the standard of care. Finally, the statute sets forth events that can cause a dissolution of the company and provides that the LLC can be dissolved if the managers or members in control of the company have acted, are acting, or will act in a manner that is illegal, oppressive, fraudulent, or unfairly prejudicial to the petitioner.[262]

---

[259] *See* UNIF. LTD. LIAB. CO. ACT § 103 (1996).
[260] *See id.* §§ 701(a), 701(f).
[261] *See id.* §§ 409(b)(1)–(3), 409(c).
[262] *See id.* § 801.

V. CONCLUSION

There are many reasons to retain the default buy-out right in the LLC, including the illiquidity of private investments, the relative imbalance of power between majority and minority owners, the difficulty a minority has in achieving suitable contractual protection, the wide variety of enterprises conducted in the LLC form (some of which may lack operating agreements altogether), the substantive and procedural uncertainties regarding claims for violations of fiduciary duties and standards of care, the potential lack of built-in statutory protections other than buy-out rights, and, most of all, the likelihood of continued failures both in human relationships and in contractual documents. The current trend away from the broad judicial monitoring of private business enterprises creates an increased need for statutory direction and express statutory protections for minority LLC owners. It is suggested that states consider the approach taken by the National Conference of Commissioners on Uniform State Laws and eliminate default exit rights in the limited partnership, rather than in the LLC, as illustrated in the Re-RULPA. Consideration should be given to adopting at least two statutory protections for minority LLC members, including: (1) limitations on the right to reduce fiduciary duties and the standard of care unreasonably; and (2) the right to seek a dissolution in the event of certain specified majority misconduct, such as illegal, oppressive, or unfairly prejudicial conduct. The latter provision has long existed in corporate law. The ULLCA contains both of these protections and also provides for a fair market value buy-out in the case of LLCs without a stated term. The ULLCA provides a flexible framework for private investors, creatively balancing the interest in contractual freedom with that of establishing a stable mandatory core of law to prevent and deter fraud and opportunism.[263] It is therefore suggested that states carefully consider the ULLCA in connection with any modification of LLC member exit rights.

---

[263] *Cf.* John C. Coffee, Jr., *The Mandatory/Enabling Balance in Corporate Law: An Essay on the Judicial Role*, 89 COLUM. L. REV. 1618, 1621 (1989) (discussing the judicial role and the competing interests in mandatory corporate laws and contractual innovation).

SYMPOSIUM: REFORMING PUNITIVE DAMAGES

THE PUNITIVE DAMAGE DEBATE

On March 13, 2001, the Harvard Journal on Legislation *held a public symposium addressing the debate surrounding proposals to reform punitive damages, both at the state and national levels. This piece discusses the primary areas of disagreement with respect to punitive damages, followed by a brief summary of the opening remarks of each panelist at the symposium.*

Although punitive damages are awarded in only two to four percent of the civil cases in which plaintiffs prevail,[1] they are an oft-maligned part of the American judicial system. In recent years, the seemingly exorbitant and unjust nature of many punitive damage awards has raised the ire of the American public.[2] The $2.9 million punitive damage award rendered to a woman who was injured by spilling hot McDonald's coffee on herself, for example, became a "rallying cry" for tort reformers.[3] In addition to disturbing many American citizens, these high punitive damage awards have led many legal commentators to call for reform.[4] In light of the controversy surrounding punitive damages, the *Harvard Journal on Legislation* held a symposium entitled "Reforming Punitive Dam-

---

* The symposium was made possible, in part, with assistance from the following sponsors: The Center for Legal Policy at the Manhattan Institute; Exxon Mobil Corp.; Lexis Publishing; Ropes & Gray LLP; and Barbri, Inc.

[1] David Luban, *A Flawed Case Against Punitive Damages*, 87 GEO. L.J. 359, 360 (1998); *see also* Marc Galanter, *Shadow Play: The Fabled Menace of Punitive Damages*, 1998 WIS. L. REV. 1, 2 (indicating that punitive damages are awarded in three percent of jury verdicts).

[2] *See, e.g.*, Gregory Nathan Hoole, Note, *In the Wake of Seemingly Exorbitant Punitive Damage Awards America Demands Caps on Punitive Damages—Are We Barking Up the Wrong Tree?*, 22 J. CONTEMP. L. 459, 459–60 (1996). The large sum of punitive damage awards may feed into the reform movement as well. In 1992, for example, the seventy-five most populous counties in the United States awarded a total of $327,300,000 in punitive damages. Brian J. Ostrom, David B. Rottman & John A. Goerdt, *A Step Above Anecdote: A Profile of the Civil Jury in the 1990s*, 79 JUDICATURE 233, 239 (1996).

[3] This occurred even though the judge decreased the award significantly. Susanah Mead, *Punitive Damages and the Spill Felt Round the World: A U.S. Perspective*, 17 LOY. L.A. INT'L & COMP. L.J. 829, 830 (1995); *see also* Valerie P. Hans, *The Contested Role of the Civil Jury in Business Litigation*, 79 JUDICATURE 242, 242 (1996) (indicating that the McDonald's coffee verdict was the "natural vehicle" for the Republican Party to use in its tort reform efforts).

[4] *See, e.g.*, David Schkade, Cass R. Sunstein & Daniel Kahneman, *Deliberating About Dollars: The Severity Shift*, 100 COLUM. L. REV. 1139, 1144 (2000); *see also* Jonathan Hadley Koenig, Note, *Punitive Damage "Overkill" After* TXO Production Corp. v. Alliance Resources*: The Need for a Congressional Solution*, 36 WM. & MARY L. REV. 751, 751–52 (1995) (arguing that the "highly-publicized and seemingly unjust nature" of punitive damages made them an issue in the 1992 presidential elections). *But see* Luban, *supra* note 1, at 361–62 (arguing that punitive damages are generally modest and that only a few outlier juries award exorbitant sums).

ages," which included two panel discussions: "The Future of Punitive Damages" and "How Should We Decide Punitive Damages?"

BACKGROUND: THE PUNITIVE DAMAGE DEBATE

Punitive damages have led to disagreements at every level. First, scholars debate the theoretical purpose that punitive damages should serve. Second, scholars disagree about the empirical evidence regarding the pervasiveness of punitive damage awards. Finally, there is significant policy disagreement about whether punitive damage law should be reformed.

There are three prominent rationales for imposing punitive damages: deterrence, retribution, and compensation.[5] Deterrence is the most generally accepted theoretical justification for punitive damages.[6] According to this theory, punitive damages should be awarded in cases in which compensatory damages are insufficient to deter behavior that society deems illicit.[7] Compensatory damages alone may not sufficiently deter conduct for a number of reasons. First, because defendants are not always held liable for their actions, they are not required to internalize the costs of all the injuries they cause.[8] Additionally, certain legal rules inherently lead to less than complete recovery for plaintiffs.[9] As a result, punitive damages are necessary to force defendants to fully internalize the harms they cause.[10] Some proponents of punitive damages dispute the focus on deter-

---

[5] David F. Partlett, *Punitive Damages: Legal Hot Zones*, 56 LA. L. REV. 781, 792 (1996).

[6] *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 908(1) (1979); David Crump, *Evidence, Economics, and Ethics: What Information Should Jurors Be Given to Determine the Amount of a Punitive-Damage Award?*, 57 MD. L. REV. 174, 182 (1998); A. Mitchell Polinsky & Steven Shavell, *Punitive Damages: An Economic Analysis*, 111 HARV. L. REV. 869, 873–74 (1998); Cass R. Sunstein, David Schkade & Daniel Kahneman, *Do People Want Optimal Deterrence?*, 29 J. LEGAL STUD. 237, 237–38 (2000).

[7] Marc Galanter & David Luban, *Poetic Justice: Punitive Damages and Legal Pluralism*, 42 AM. U. L. REV. 1393, 1426–27 (1993) (indicating that the imposition of punitive damages can offset weak administrative remedies and alter corporate behavior).

[8] *See* Polinsky & Shavell, *supra* note 6, at 874; Dorsey D. Ellis, Jr., *Fairness and Efficiency in the Law of Punitive Damages*, 56 S. CAL. L. REV. 1, 52 (1982).

[9] The most important of these rules are the payment of attorney's fees by each party to its own counsel and the requirement in contract cases of certainty of damages before they are awarded. John A. Sebert, Jr., *Punitive and Nonpecuniary Damages in Actions Based Upon Contract: Toward Achieving the Objective of Full Compensation*, 33 UCLA L. REV. 1565, 1659–62 (1986). Although Sebert claims that he is concentrating upon full compensation as the justification for punitive damages, he also asserts that full compensation, including punitive damages, is necessary in order to ensure that parties only breach when it is efficient to do so. *Id.*

[10] Polinsky and Shavell present the classic economic model, arguing that when there is less than a 100% chance that defendants will be held liable, punitive damages should be used to offset the chance of non-detection. Polinsky & Shavell, *supra* note 6, at 887–90. It is frequently argued, however, that punitive damages do not actually deter. *See, e.g.*, E. Donald Elliott, *Why Punitive Damages Don't Deter Corporate Misconduct Effectively*, 40 ALA. L. REV. 1053, 1057–58 (1989) (arguing that the unpredictability of punitive damages

rence, arguing instead that punitive damages serve a retributive purpose.[11] According to retributive theory, punitive damages should be imposed if a defendant's actions are particularly heinous.[12] The purpose of imposing punitive damages under this rationale is to "vent the indignation of the victimized."[13] Under this theory, punitive damages should be imposed to punish the defendant regardless of whether they serve a deterrence function. Other authors assert that punitive damages are necessary to fully compensate plaintiffs.[14] Punitive damages are called upon to make up for the limitations that legal rules place on recovery,[15] including intangible damages that otherwise would go uncompensated.[16]

In addition to the theoretical disagreement about the purpose of punitive damages, there is significant disagreement about the empirical evidence regarding their prevalence. In one study, the RAND Corporation Institute for Civil Justice found that there had been a dramatic increase in punitive damage awards between 1960 and 1984.[17] Similarly, many critics have claimed that juries award punitive damages erratically and that these awards have gotten out of control.[18] Proponents of the continued vitality of punitive damages point to data indicating that the majority of punitive damage awards are for small sums[19] and are strongly correlated to the compensatory awards granted in the cases.[20] Theodore Eisenberg and Martin T. Wells argue that "[t]he available data suggest that businesses,

---

makes them an inefficient means of deterring corporate conduct).

[11] *See, e.g.*, Luban, *supra* note 1, at 378–79.

[12] *See, e.g.*, Galanter & Luban, *supra* note 7, at 1432 (arguing that retribution theory calls for a scaling of punitive damages to the heinousness of the behavior). *But see infra* note 69 (citing arguments in opposition to redistribution as a justification).

[13] Michael Rustad & Thomas Koenig, *The Historical Continuity of Punitive Damages Awards: Reforming the Tort Reformers*, 42 AM. U. L. REV. 1269, 1320–21 (1993) (quoting Note, *Punitive Damages and Libel Law*, 98 HARV. L. REV. 847, 851 (1985)).

[14] *See, e.g.*, Sebert, *supra* note 9, at 1570; *see also* Rustad & Koenig, *supra* note 13, at 1321–22 (asserting that some jurisdictions view compensation as yet another goal of punitive damages).

[15] *See* Sebert, *supra* note 9 (discussing how the legal rules of contract systematically under-compensate plaintiffs); Rustad & Koenig, *supra* note 13, at 1321–22.

[16] Partlett, *supra* note 5, at 793–95.

[17] Lynda A. Sloane, Note, *The Split Award Statute: A Move Toward Effectuating the True Purpose of Punitive Damages*, 28 VAL. U. L. REV. 473, 488 (1993) (summarizing the findings of the RAND study).

[18] *See, e.g.*, W. Kip Viscusi, *The Social Costs of Punitive Damages Against Corporations in Environmental and Safety Torts*, 87 GEO. L.J. 285, 333 (1998); Michael J. Saks, *Do We Really Know Anything About the Behavior of the Tort Litigation System—And Why Not?*, 140 U. PA. L. REV. 1147, 1254 (1992) (quoting Vice President Dan Quayle's claim of "arbitrary, even freakish application" of punitive damages).

[19] Ostrom, *supra* note 2, at 238.

[20] Eisenberg & Wells, *infra* note 21, at 393–94 (explaining that about half of the variance in punitive damage awards can be explained by the compensatory damage award and that once the outer five percent of awards are excluded, there are hardly any extreme awards); *see also* Michael Rustad, *In Defense of Punitive Damages in Products Liability: Testing Tort Anecdotes with Empirical Data*, 78 IOWA L. REV. 1, 50 (1992) (indicating that punitive damage awards in products liability cases are generally proportional to the actual damages).

insurance companies, and defense lawyers cannot support the claim that punitive awards are frequent, or that they follow a crazy pattern with little or no relation to compensatory awards .... The mass of punitive awards satisfy a surprisingly regular pattern."[21]

Finally, there is what might be termed "the policy debate" about whether punitive damages should be reformed. In many ways, the theoretical and empirical disagreements outlined above have fed into these policy discussions. The proposed reforms are most easily organized as those that would expand punitive damage awards and those that would contract these awards.

There is some support for the notion that punitive damages do not play a prominent enough role in American law. John A. Sebert, Jr., for example, argues that punitive damages should be expanded in contract cases because compensatory damages do not fully compensate plaintiffs for losses caused by contract breaches.[22] He asserts that the applicable legal rules lead to systematic under-compensation of plaintiffs in contract suits.[23] These rules include the requirements of certainty of damages, the refusal to require payment of interest, and the American rule for legal fees.[24] To overcome this problem, Sebert concludes that courts should extend the availability of punitive or super-compensatory damages in contract cases.[25]

The more prominent view, however, appears to be that punitive damages need to be limited. Proponents of this view have suggested numerous reforms, including capping punitive damages, taking punitive damage decisions away from juries, imposing more stringent burdens of proof for punitive liability, bifurcating trials so that punitive damages are considered only after the jury finds the defendant liable, and giving juries more specific instructions to aid them in setting punitive damages.

Many critics advocate for caps upon punitive damage awards,[26] and such caps have become a prominent reform enacted by state legislatures.

---

[21] Theodore Eisenberg & Martin T. Wells, *Punitive Awards After* BMW, *A New Capping System, and the Reported Opinion Bias*, 1998 WIS. L. REV. 387, 388–89; *see also* Stephen Daniels & Joanne Martin, *Myth and Reality in Punitive Damages*, 75 MINN. L. REV. 1, 4 (1990) (arguing that critics' empirical "propositions are based on scanty empirical data and highly questionable interpretations of those data").

[22] Sebert, *supra* note 9, at 1656–57. It is important to note, however, that the author makes this suggestion even though punitive damage awards are already most prevalent in contract cases. *See* Ostrom, Rottman & Goerdt, *supra* note 2, at 238–39 (asserting that the vast majority of punitive damage awards occur in contract cases and that the dollar value of punitive damages in contract cases exceeds the dollar value in tort cases).

[23] Sebert, *supra* note 9, at 1566–67.

[24] *Id.* at 1566–69.

[25] *Id.* at 1570. The proposed expansion of punitive damages in contract cases may be coming to fruition. Galanter & Luban, *supra* note 7, at 1415–16 (indicating that punitive damage awards are on the rise in contract cases).

[26] *See, e.g.*, Eisenberg & Wells, *supra* note 21, at 406-07 (urging a cap of ten times the compensatory award); Linda Babcock & Greg Pogarsky, *Damage Caps and Settlement: A Behavioral Approach*, 28 J. LEGAL STUD. 341, 343–44 (1999) (indicating that one of Vice

Some states have placed a dollar amount limitation upon punitive awards;[27] some states have limited punitive damages to a given multiple of compensatory damages;[28] some states have limited punitive damages to a percent of the defendant's profits;[29] and still other states have utilized a combination of these approaches.[30] These caps, however, have been criticized as irrational and contrary to the goals of punitive damages.[31]

Others have argued that punitive damage liability should be decided by judges rather than by juries. Traditionally, juries are charged with determining whether punitive damages will be imposed and with setting the level of those damages.[32] This system has come under attack, however, because of the alleged capricious nature of punitive damage awards.[33] Some critics argue that if punitive damages are to be awarded, they should be imposed by judges because juries are not competent to decide

---

President Dan Quayle's major reform proposals was a cap on punitive damages).

[27] *See, e.g.*, GA. CODE ANN. § 51-12-5.1(g) (1998) (limiting punitive damages to $250,000, except for product liability cases); VA. CODE ANN. § 8.01-38.1 (2000) (capping punitive damages at $350,000).

[28] *See, e.g.*, CONN. GEN. STAT. ANN. § 52-240b (1991) (limiting punitive awards in product liability cases to twice the compensatory award); FLA. STAT. ANN. § 768.73(1)(a) (West Supp. 2001) (capping punitive damages at three times the amount of the compensatory award or $500,000, whichever is greater).

[29] *See, e.g.*, KAN. STAT. ANN. § 60-3701(e)–(f) (Supp. 2000) (providing that punitive damages are generally limited to lesser of the defendant's highest annual income in the last five years or $5 million).

[30] *See, e.g.*, NEV. REV. STAT. ANN. § 42.005 (Supp. 1999) (capping punitive damages at three times the compensatory damages if those damages are above $100,000, and limiting punitive damages to $300,000 if the compensatory award is less than $100,000); TEX. CIV. PRAC. & REM. CODE ANN. § 41.008 (Vernon 1997) (limiting punitive damages to the greater of two times economic harm plus the non-economic harm or $200,000, unless the defendant's conduct was particularly egregious).

[31] *See, e.g.*, Crump, *supra* note 6, at 223–24 (arguing that caps might be irrational and that, if they are to be imposed, they should be set high enough to preserve the deterrence goal of punitive damages); Thomas M. Melsheimer & Steven H. Stodghill, *Due Process and Punitive Damages: Providing Meaningful Guidance to the Jury*, 47 SMU L. REV. 329, 347–48 (1994) (discussing various shortcomings of punitive damage caps); Partlett, *supra* note 5, at 824 (stating that imposing damage caps is contrary to the purpose of punitive damages); Sandra N. Hurd & Frances E. Zollers, *State Punitive Damages Statutes: A Proposed Alternative*, 20 J. LEGIS. 191, 199 (1994) (arguing that limiting punitive damages could undermine the deterrent purpose of punitive damages); Galanter & Luban, *supra* note 7, at 1432 (arguing that to achieve the goals of retribution, punitive damages should be scaled to the heinousness of the defendant's behavior rather than to the resulting harm).

[32] DAN B. DOBBS, LAW OF REMEDIES: DAMAGES—EQUITY—RESTITUTION § 3.11 (2d ed. 1993) (indicating that as a general rule punitive damages are set exclusively by juries and that judges only play a role in limiting damages).

[33] *See, e.g.*, Cass Sunstein, Daniel Kahneman & David Schkade, *Assessing Punitive Damages*, 107 YALE L.J. 2071, 2076 (1998). These critics follow in the long line of commentators who have asserted that juries are often not competent to decide the complex issues raised in civil litigation. *See, e.g.*, Joe S. Cecil, Valerie P. Hans & Elizabeth C. Wiggins, *Citizen Comprehension of Difficult Issues: Lessons From Civil Jury Trials*, 40 AM. U. L. REV. 727, 733–44 (1991) (discussing various attacks on jury competence but asserting that these attacks stand in contrast to the empirical research of legal scholars). Some have observed, for example, that while judges and juries often agree on punitive damage liability, they frequently differ with regard to the amount. *Id.* at 745–46.

whether punitive damages are appropriate.[34] Others have more modestly proposed that juries should decide whether punitive damages are to be imposed and should give the judge an indication of the level of outrage engendered by the defendant's behavior, but that the judge should be charged with actually setting the dollar award.[35] These critics assert that allowing the judge to set the dollar value would lead to greater predictability in awards.[36] This increased certainty would in turn further the deterrence rationale for punitive damages.[37] At least one state has adopted this approach in some cases, requiring that judges—not juries—set the punitive award.[38]

Some critics would also alter how the decision-maker sets the punitive award. First, some argue that the burden of proof should be heightened for purposes of determining punitive damages.[39] These critics reason that because punitive damages are intended in some respect to punish wrongdoers, more proof of culpability is required.[40] Though some scholars criticize this approach as irrational,[41] several states have already adopted a heightened standard of review.[42]

---

[34] *See, e.g.*, Viscusi, *infra* note 79, at 589–90 (arguing that juries do not properly consider cost-benefit analyses and thus the responsibility of evaluating the reprehensibility of corporate behavior should be shifted to administrative agencies).

[35] Sunstein, Kahneman & Schkade, *supra* note 33, at 2120 (comparing their proposal to criminal sentencing in which the jury decides liability but the judge sets the punishment); *see also* Malcolm E. Wheeler, *A Proposal for Further Common Law Development of the Use of Punitive Damages in Modern Product Liability Litigation*, 40 ALA. L. REV. 919, 948 (1989) (arguing that the jury should determine the values relevant to the law and economics deterrence formula through use of a special verdict form, and the judge should then calculate an acceptable range of values for punitive damages from which the jury may choose a final dollar amount).

[36] Sunstein, Kahneman & Schkade, *supra* note 33, at 2108; *see also* David Owen, *Punitive Damages in Products Liability Litigation*, 74 MICH. L. REV. 1257, 1320 (1976).

[37] *See* Sunstein, Kahneman, & Schkade, *supra* note 33, at 2111–15 (arguing that judges taking the decision of the dollar value away from juries will minimize unpredictability in punitive awards, which the authors view as the major problem with punitive damages).

[38] CONN. GEN. STAT. ANN. § 52-240b (1991) (stating that in products liability cases the court shall set punitive awards).

[39] Hurd & Zollers, *supra* note 31, at 201. There is some debate about whether to require clear and convincing evidence or proof beyond a reasonable doubt. *Compare* Dick Thornburgh, *America's Civil Justice Dilemma: The Prospects for Reform*, 55 MD. L. REV. 1074, 1086 (1996) (urging the use of a clear and convincing evidence standard) *with* COLO. REV. STAT. ANN. § 13-25-127 (2000) (requiring proof beyond a reasonable doubt).

[40] Hurd & Zollers, *supra* note 31, at 201; *see also* Wheeler, *supra* note 35, at 953–54 (arguing that there should be a presumption that punitive damages greater than twice the compensatory awards are excessive and that judges should impose a clear and convincing evidence standard).

[41] Melsheimer & Stodghill, *supra* note 31, at 346–47.

[42] These states usually require that the facts justifying punitive damages be proven by clear and convincing evidence. *See* ALA. CODE § 6-11-20 (Supp. 2000); ALASKA STAT. § 09.17.020 (Supp. 1999); GA. CODE ANN. § 51-12-5.1 (1999); IND. CODE § 34-51-3-2 (Supp. 2000); KAN. STAT. ANN. § 60-3701(c) (Supp. 1999); KY. REV. STAT. ANN. § 411.184 (Michie Supp. 2000); MINN. STAT. § 549.20(1) (2000); MONT. CODE ANN. § 27-1-221 (1999); OHIO REV. CODE ANN. § 2315.21 (West 2000); UTAH CODE ANN. § 78-18-1

Reformers have also proposed that trials involving punitive damage claims be bifurcated so that juries do not hear evidence regarding punitive damages until after they have found the defendant liable for the harm suffered.[43] These authors fear that evidence relevant only to determining punitive damages—such as a defendant's assets—will skew the jury findings of liability.[44] At least one empirical study confirms this fear.[45] At the same time, however, evidence indicates that bifurcating these trials might lead to higher average punitive damage awards.[46] Thus, bifurcation might come as a mixed blessing for defendants.

Other reformers suggest an alteration of punitive damage jury instructions. Some authors assert that juries are not given enough guidance for determining punitive damage awards.[47] They argue that juries are left with little understanding of the purposes of punitive damages, leading to awards that are nothing more than shots in the dark.[48] Given the disagreement about the purpose of punitive damages,[49] it is difficult to imagine what instructions should be given.[50] In addition, if punitive damages are meant to achieve all three goals discussed above, the legal system must determine how much weight each goal should be given. Ultimately, it is unclear what additional instructions, if any, judges could or should provide to juries.

Finally, there is sharp disagreement over what evidence should be put before the decision-maker. Some critics assert that the fact finder should consider punitive damage awards from similar cases because it would lead to increased certainty in awards.[51] The current consensus

---

(Supp. 2000). Colorado, however, has gone so far as to require proof beyond a reasonable doubt. COLO. REV. STAT. ANN. § 13-25-127 (2000).

[43] *See, e.g.*, Wheeler, *supra* note 35, at 947; Melsheimer & Stodghill, *supra* note 31, at 348–49.

[44] *See* Melsheimer & Stodghill, *supra* note 31, at 348–49 (arguing that bifurcation would eliminate the fear that the jury will consider the defendant's wealth in determining liability).

[45] Stephan Landsman et al., *Be Careful What You Wish For: The Paradoxical Effects of Bifurcating Claims For Punitive Damages*, 1998 WIS. L. REV. 297, 316 (reporting research results indicating that juries hearing evidence relevant to punitive damages during a combined trial are more likely to find a defendant liable).

[46] *Id.* at 325.

[47] *See* Melsheimer & Stodghill, *supra* note 31, at 330. Other critics more skeptically assert that jury instructions are altogether ineffective. *See, e.g.*, Sunstein, Kahneman & Schkade, *supra* note 33, at 2111; Alan Calnan, *Ending the Punitive Damage Debate*, 45 DEPAUL L. REV. 101, 102–03 (1995) (arguing that reformers often incorrectly claim that lack of guidance leads to exorbitant awards); Robert D. Cooter, *Punitive Damages, Social Norms, and Economic Analysis*, LAW & CONTEMP. PROBS., Summer 1997, at 73, 75–76.

[48] *See, e.g.*, Wheeler, *supra* note 35, at 947–48 (arguing that juries should be instructed that the purpose of punitive damages is to deter wrongful conduct, and that the jury should only award sufficient punitive damages to provide deterrence).

[49] *See supra* notes 5–16 and accompanying text.

[50] *See* Wheeler, *supra* note 35, at 947–48 (arguing that juries should not be instructed that punitive damages are intended to punish because adding the punishment language will confuse the jury).

[51] *See* Sunstein, Kahneman & Schkade, *supra* note 33, at 2117 (arguing that a jury

among states, however, is that juries should not be allowed to consider past awards.[52] Critics also argue that juries should not be allowed to consider the wealth of the defendant,[53] because it is not relevant to the magnitude of award necessary to deter the conduct.[54] Many states do allow the admission of evidence of a defendant's wealth.[55]

The preceding pages have presented some of the major debates among scholars of punitive damages. These authors disagree at all levels: theoretical, empirical, and policy. The Symposium was held to provide a forum to examine these and other arguments.

PANEL ONE: THE FUTURE OF PUNITIVE DAMAGES

MODERATOR:

Bruce Hay is professor of law at Harvard Law School, where he teaches civil procedure, professional responsibility, corporations, and environmental law. His research interests include legal procedure, legal ethics, and the economics of litigation. He has published articles on various issues relating to litigation.[56] Professor Hay received his B.A. from the University of Wisconsin at Madison and his J.D. from Harvard Law School.

PANELISTS:

Congressman Robert Barr (R-Ga.) is the United States Representative for the Seventh District of Georgia and is Chairman of the Judiciary Subcommittee on Commercial and Administrative Law. He also serves on the Government Reform Committee, on the Financial Services Committee, and as Assistant Majority Whip for the Republican Party. From 1986 to 1990 he was United States Attorney for the Northern District of Georgia. Congressman Barr received his B.A. from the University

---

should be provided with comparison factual scenarios and asked to decide which case most closely matches the egregiousness of a defendant's actions).

[52] *Id.* at 2118–19.

[53] *See* Crump, *supra* note 6, at 217. There is evidence indicating that the wealth of a defendant will affect the dollar amount of punitive awards. *See, e.g.*, Sunstein, Kahneman & Schkade, *supra* note 33, at 2105.

[54] *See, e.g.*, Crump, *supra* note 6, at 217.

[55] *See, e.g.*, Life Ins. Co. of Ga. v. Johnson, 684 So. 2d 685, 701–02 (Ala. 1996), *vacated on other grounds*, 117 S. Ct. 288 (1996); Wayte v. Rollins Int'l, Inc., 215 Cal. Rptr. 59, 72 (Cal. Ct. App. 1985); Lunsford v. Morris, 746 S.W.2d 471, 476 (Tex. 1988).

[56] For a sample of Professor Hay's contribution to the legal literature, see Bruce L. Hay, *The Theory of Fee Regulation in Class Action Settlements*, 46 AM. U. L. REV. 1429 (1997); Bruce Hay, *Optimal Contingent Fees in a World of Settlement*, 26 J. LEGAL STUD. 259 (1997).

of Southern California, his M.A. from George Washington University, and his J.D. from Georgetown University.[57]

Martha Chamallas is professor of law at the University of Pittsburgh Law School, where she teaches torts, employment discrimination, and feminist legal theory. Her work on tort law has focused on emotional injuries and the effects of race and gender bias in computing damages. Professor Chamallas received her B.A. from Tufts University and her J.D. from Louisiana State University.[58]

Marc Galanter is the John and Rylla Bosshard Professor of Law and South Asian Studies at the University of Wisconsin at Madison and the LSE Centennial Professor at the London School of Economics. He serves as chairman of Wisconsin's Institute for Legal Studies, one of the leading centers for empirical study of the legal system, as well as director of the Institute's Dispute Processing Research Program. Professor Galanter received his B.A., M.A., and J.D. from the University of Chicago.[59]

Jon Hanson is professor of law at Harvard Law School, where he teaches courses in torts, corporations, products liability theory, and law and behavioralism. Professor Hanson has testified before Congress on the issue of tobacco regulation, and has served as an expert witness in several of the tobacco lawsuits brought by the state attorneys general. Professor Hanson received his B.A. from Rice University and his J.D. from Yale Law School.[60]

Walter Olson is a senior fellow at the Manhattan Institute. He has written extensively on reforming the American legal system, and his commentaries frequently appear in national publications such as the *Wall Street Journal*, the *New York Times*, and the *National Review*. In 1991, Mr. Olson authored *The Litigation Explosion: What Happened When America Unleashed the Lawsuit*, and in 1999 he launched the Web site Overlawyered.com. Mr. Olson received his B.A. from Yale University.[61]

---

[57] For a sample of Congressman Barr's contribution to the legal literature, see Bob Barr, *High Crimes and Misdemeanors: The Clinton-Gore Scandals and the Question of Impeachment*, 2 TEX. REV. L. & POL. 1 (1997).

[58] For a sample of Professor Chamallas's contribution to the legal literature on the tort system, see Martha Chamallas, *The Disappearing Consumer, Cognitive Bias and Tort Law*, 6 ROGER WILLIAMS U. L. REV. 9 (2000); Martha Chamallas, *The Architecture of Bias: Deep Structures in Tort Law*, 146 U. PA. L. REV. 463 (1998).

[59] For a sample of Professor Galanter's contribution to the literature on punitive damages, see Galanter & Luban, *supra* note 7; Marc Galanter, *News from Nowhere: The Debased Debate on Civil Justice*, 71 DENV. U. L. REV. 77 (1993); Marc Galanter, *Shadow Play: The Fabled Menace of Punitive Damages*, 1998 WIS. L. REV. 1.

[60] For a sample of Professor Hanson's contribution to the literature on punitive damages, see Steven P. Croley & Jon D. Hanson, *The Nonpecuniary Costs of Accidents: Pain-And-Suffering Damages in Tort Law*, 108 HARV. L. REV. 1785 (1995); Steven P. Croley & Jon D. Hanson, *What Liability Crisis? An Alternative Explanation for Recent Events in Products Liability*, 8 YALE J. ON REG. 1 (1991).

[61] For a sample of Walter Olson's contribution to the legal literature see Walter Olson, *Tortification of Contract Law: Displacing Consent and Agreement*, 77 CORNELL L. REV. 1043 (1992).

Thomas M. Sobol is a partner with the law offices of Lieff, Cabraser, Heimann & Bernstein, LLP. His primary focus is on major class action litigation, and he is currently involved in class actions filed against the tobacco, gun, asbestos, healthcare, pharmaceutical, and industrial agriculture industries. He also assisted in antitrust litigation against Microsoft. Mr. Sobol received his B.A. from Clark University and his J.D. from Boston University School of Law.[62]

David Tuerck is the executive director of the Beacon Hill Institute at Suffolk University, where he also serves as chairman of the Department of Economics. His specialties include state tax policy, welfare reform, and the economics of regulation. Dr. Tuerck received his A.B. and A.M. from George Washington University, and his Ph.D. in economics from the University of Virginia.[63]

SUMMARY OF OPENING REMARKS:[64]

Congressman Robert Barr began the discussion by describing the current political climate toward punitive damages at the federal level. He argued that the election of President George W. Bush has greatly increased the chances of passing a comprehensive tort reform bill.[65] He indicated that he would favor a bill that included caps on punitive damage awards, and also suggested mandating that punitive damages be proven by clear and convincing evidence, rather than by a preponderance of the evidence.[66] Congressman Barr urged the passage of tort reform to limit the occurrence of punitive damage awards.

Professor Martha Chamallas advocated the continued use of punitive damages. Her argument began from the premise that general deterrence theory is not the only justification for punitive damages; punitive damages also serve a compensatory purpose. They were originally used in cases involving an affront to a person's honor.[67] In such cases, compen-

---

[62] For a sample of Mr. Sobol's contribution to the literature on punitive damages, see Elizabeth J. Cabraser & Thomas M. Sobol, *Equity for the Victims, Equity for the Transgressor: The Classwide Treatment of Punitive Damage Claims*, 74 TUL. L. REV. 2005 (2000).

[63] For a sample of Dr. Tuerck's contribution to the legal literature, see DAVID TUERCK & IN-MEE BAEK, AN ECONOMIC MODEL OF MASSACHUSETTS TAX POLICY (1994); DAVID TUERCK & WILLIAM F. O'BRIEN, JR., THE COMPASSION TAX CREDIT: A FAMILY ADVOCATE PILOT PROGRAM (1996).

[64] This summary is based on the author's recollection of the panelists' presentations.

[65] In 1996, Congress passed the Commonsense Legal Reform Act, which included a $250,000 cap on punitive damage awards; President Clinton vetoed the bill. *See* Babcock & Pogarsky, *supra* note 26, at 344.

[66] During the discussion portion, Congressman Barr acknowledged that the clear and convincing evidence standard might not have an enormous effect upon juries, but he maintained that it would symbolize the country's distaste of punitive damages.

[67] For a fuller explication of the history of punitive damages see Alan Calnan, *Ending the Punitive Damage Debate*, 45 DEPAUL L. REV. 101, 104–09 (1995) (discussing the historical roots of punitive damages in ancient cultures, the English common law, and early

satory damages were presumed to be insufficient to fully compensate plaintiffs for their injuries. The contemporary analogue to such cases is discrimination. An individual denied the opportunity to rent an apartment because of her race, for example, is not likely to be made whole by compensatory damages. Punitive damages, then, may serve an important compensatory role in this type of case.[68]

Professor Marc Galanter followed Professor Chamallas by describing some of the emerging empirical data regarding punitive damages. He indicated that we are just now beginning to understand punitive damages. They are most common in cases involving financial damages resulting from intentional torts and are more common in the South. Professor Galanter remarked that although there is a "folklore" that punitive damages are common and are awarded casually, in fact they are awarded infrequently. Despite this data, he observed that corporate decision-makers continue to overestimate the likelihood of punitive damage awards.[69]

Professor Jon Hanson presented his opinion that the scholarship calling for reform is flawed. First, he argued that this literature is based upon conventional law and economics without reference to attribution theory, which posits that harmed individuals do care about a person's motives. Professor Hanson gave the example of a waiter spilling a glass of tea on a patron. That patron might ask for a wide range of compensation. He might simply ask for an apology, or he might ask that the restaurant pay for his dry-cleaning, or he might ask that the waiter be fired. Professor Hanson argued that in deciding what compensation to ask for, the patron would care about why the waiter spilled the tea on him. Was it because the patron was African American? Was it an accident? The patron might also care about whether the waiter had spilled tea on anyone else before. The literature discussing punitive damage awards treats all spilled tea the same without regard to the details. Professor Hanson argued that this is a major shortcoming. The second problem that he articulated is the effect of money on the law and economics literature, a factor that is not considered by the academic literature itself. Professor Hanson observed that many of these articles are funded either directly or indirectly by business interests, and recommended that commentators engage in a more critical inquiry into the effects that this money might have on the literature as a whole.

Walter Olson next compared punitive damages to criminal sanctions. The distinction is that when the government attempts to impose criminal sanctions, it must provide certain procedural protections to the defendant.

America); *see also* Partlett, *supra* note 5, at 783–87.

[68] Her explanation of the compensatory purpose of punitive damages is similar to that in Parlett, *supra* note 6, at 793–95. Interestingly, however, Partlett contends that nonpecuniary damages other than punitive damages might be sufficient. *Id.*

[69] For a fuller exposition of these points see Marc Galanter, *Real World Torts: An Antidote to Anecdote*, 55 MD. L. REV. 1093, 1126–40 (1996).

Defendants are not given such protections in civil actions when they are sued by plaintiffs seeking punitive damages. Mr. Olson argued that the most controversial punitive damage awards are generally those in which some of the protections that a criminal defendant would have received were notably absent and impacted the course of the trial.[70] These include cases in which there was forum-shopping, the law was not clearly defined in advance, or there was a threat that the defendant would have to pay for the same conduct more than once. Cases in which the procedural protections would not have aided the defendant do not generally raise the ire of the public.

Thomas Sobol then provided the perspective of a practicing plaintiffs' attorney. He started by observing that the increase in the dollar value of punitive damages over the last decade is not surprising given the exponential growth of business enterprises. He suggested that one of the main problems with punitive damages is that some future plaintiffs may not be able to recover compensatory awards if large punitive damages awarded to other plaintiffs have exhausted the defendant's funds. To alleviate these distributive concerns, Mr. Sobol suggested joining punitive damage cases into class actions.[71]

David Tuerck began by defending the law and economics literature. He asserted that the authors of the seminal law and economics literature—Richard Posner, Steven Shavell, and Learned Hand—are hardly "lackeys" of the business community. Shavell, for example, argues that there are cases in which punitive damages are appropriate.[72] Nonetheless, these authors do believe that punitive damages should not be imposed when they might lead to overly cautious behavior. The primary problem that Dr. Tuerck sees with punitive damages is that they lead to uncertainty, which disturbs the presumptions underlying the economic model.

---

[70] This argument follows a line of literature focussing on the impropriety of using punitive damages to achieve the goal of punishment, which, it is argued, should be the exclusive domain of the criminal system. *See, e.g.*, David L. Walther & Thomas A. Plein, *Punitive Damages: A Critical Analysis:* Kink v. Combs, 49 MARQ. L. REV. 369 (1965); Jeffery W. Grass, *The Penal Dimensions of Punitive Damages*, 12 HASTINGS CONST. L.Q. 241, 314 (1985) (asserting that punitive damages are penal in nature and thus the government should not be allowed to pass the role off onto civil plaintiffs); Comment, *Criminal Safeguards and the Punitive Damages Defendant*, 34 U. CHI. L. REV. 408, 412–30 (1967); *see also* Smith v. Wade, 461 U.S. 30, 59 (1983) (Rehnquist, J., dissenting). *But see* Galanter & Luban, *supra* note 7, at 1456-60 (arguing that there is no need for defendants in actions for punitive damages to receive these protections, because the litigation is not between the state and an individual).

[71] *See generally* Cabraser & Sobol, *supra* note 62.

[72] *See* Polinsky & Shavell, *supra* note 6, at 887–90 (arguing that punitive damages are appropriate when there is less than a one hundred percent chance of liability).

PANEL TWO: HOW SHOULD WE DECIDE PUNITIVE DAMAGES?

MODERATOR:

Margo Schlanger is assistant professor at Harvard Law School, where she teaches constitutional law, torts, and civil rights law. Her research interests include constitutional torts and the reformation of government institutions through litigation. She is the author of *Injured Women Before Common Law Courts, 1860-1930.*[73] Professor Schlanger received her B.A. from Yale University and her J.D. from Yale Law School.

PANELISTS:

Carl T. Bogus is associate professor at the Roger Williams University School of Law, where he teaches courses in torts, products liability, administrative law, and evidence. He has written widely about products liability and tort reform, and his most recent work is *Why Lawsuits are Good for America: Disciplined Democracy, Big Business and the Common Law*, to be published in July of 2001. Professor Bogus received his B.A. and J.D. from Syracuse University.[74]

John E. Calfee is a resident scholar at the American Enterprise Institute. He has served in the Bureau of Economics at the Federal Trade Commission, and has taught courses on marketing and consumer behavior at the business schools of the University of Maryland and Boston University. Dr. Calfee's research has focused on the economics of tort liability. Dr. Calfee received his B.A. from Rice University, his M.A. from the University of Chicago, and his Ph.D. in economics from the University of California at Berkeley.[75]

Mary Rose is a research fellow at the American Bar Foundation, an institute devoted to the study of law and social science. Her research has focused on jury selection, jury decision-making, and the effect of jury reforms. She also has studied popular views on fair procedures and fair outcomes in legal and non-legal settings. Dr. Rose received her B.A. from Stanford University and her Ph.D. in social psychology from Duke University.[76]

---

[73] 21 HARV. WOMEN'S L.J. 79 (1998).

[74] For a sample of Professor Bogus's contribution to the literature on punitive damages and the tort system, see Carl T. Bogus, *War on the Common Law: The Struggle at the Center of Products Liability*, 60 MO. L. REV. 1 (1995).

[75] For a sample of Dr. Calfee's contribution to the literature on punitive damages, see Paul H. Rubin, John E. Calfee & Mark F. Grady, BMW v. Gore: *Mitigating the Punitive Economics of Punitive Damages*, 5 SUP. CT. ECON. REV. 179 (1997).

[76] For a sample of Dr. Rose's contribution to the literature on the tort system, see Neil Vidmar, Felicia Gross & Mary Rose, *Jury Awards for Medical Malpractice and Post-Verdict Adjustments of Those Awards*, 48 DEPAUL L. REV. 265 (1998).

David A. Schkade is the Herbert D. Kelleher/MCorp Regents Professor of Business at the University of Texas at Austin, where he teaches negotiation, the psychology of decision-making, and research methods. His recent work has focussed on the psychology of well-being, environmental resource valuation, jury decision-making, and decision support systems. Professor Schkade received his B.A. and M.B.A. from the University of Texas at Austin and his M.S. and Ph.D. in organizational psychology from Carnegie Mellon University.[77]

Neil Vidmar is the Russell M. Robinson II Professor of Law at Duke Law School. He also holds an appointment in the Duke Psychology Department. His research interests include civil jury performance. He has testified, consulted, and drafted amicus briefs as an expert on jury behavior for trials in the United States and abroad. Professor Vidmar received his B.A. from MacMurray College, and his M.A. and Ph.D. in social psychology from the University of Illinois.[78]

W. Kip Viscusi is the John F. Cogan, Jr. Professor of Law and Economics and the director of the Program on Empirical Legal Studies at Harvard Law School. His research has focused on individual and societal responses to risk and uncertainty, and he has published seventeen books and over 200 articles. Professor Viscusi is also the founding editor of the *Journal of Risk and Uncertainty* and has served on the editorial boards of eleven other journals, including the *American Economic Review* and the *Review of Economics and Statistics*. Professor Viscusi received his B.A. from Harvard College and his M.A., M.P.P., and Ph.D. in economics from Harvard University.[79]

SUMMARY OF OPENING REMARKS:[80]

The panel began with Dr. Mary Rose presenting a paper she co-authored with Professor Neil Vidmar.[81] Professor Vidmar and Dr. Rose evaluated many empirical claims that are made about punitive damages

---

[77] For a sample of Professor Schkade's contribution to the literature on punitive damages, see Schkade, Sunstein & Kahneman, *supra* note 4; Sunstein, Schkade, & Kahneman, *supra* note 6, and Sunstein, Kahneman & Schkade, *supra* note 33. Professor Schkade is also authoring an article that will appear in the next issue of the *Harvard Journal on Legislation*.

[78] For a sample of Professor Vidmar's contribution to the legal literature on torts and punitive damages, see Neil Vidmar, *The Performance of the American Civil Jury: An Empirical Perspective*, 40 ARIZ. L. REV. 849 (1998); Neil Vidmar, *Pap and Circumstance: What Jury Verdict Statistics Can Tell Us About Jury Behavior and the Tort System*, 28 SUFFOLK U. L. REV. 1205 (1994).

[79] For a sample of Professor Viscusi's contribution to the legal literature on punitive damages, see W. Kip Viscusi, *Corporate Risk Analysis: A Reckless Act?*, 52 STAN. L. REV. 547 (2000). Professor Viscusi is also authoring an article that will appear in the next issue of the *Harvard Journal on Legislation*.

[80] This summary is based on the author's recollection of the panelists' presentations.

[81] Neil Vidmar & Mary Rose, *Punitive Damages by Juries in Florida: In Terrorum and in Reality*, 38 HARV. J. LEGIS. 487 (2001).

by examining Florida cases reported in Westlaw's jury verdict database between 1988 and 2000. During the relevant time period there were, on average, twenty-two punitive damage awards per year in Florida. Excluding asbestos cases and the recent tobacco verdict, punitive damages were awarded in only one products liability case. Juries, the survey found, are unlikely to award punitive damages in premises liability cases. In non-products liability cases, punitive damages were most frequent in cases involving drunk driving, financial disputes, and assaults. The authors concluded that the Florida cases do not support the claim that there are significant problems with the current punitive damages system.

Professor David Schkade articulated what he considers the major shortcoming of punitive damages: their unpredictability.[82] Punitive damage awards are consistently erratic.[83] Jurors are given little guidance, largely because policymakers do not know what guidance to give. Professor Schkade indicated that one of the main causes of the unpredictability of jury awards is that different jurors use different "anchors." Anchors are the starting points that individual jurors use as their frames of reference. If, for example, two individuals are presented with identical facts but with two different anchors—one sees a plaintiff demand one million dollars in punitive damages while the other sees a plaintiff demand ten million dollars—the individuals come to different conclusions. Professor Schkade argued that in order to overcome the anchoring problem, the decision-maker should consider past cases in determining the punitive damage award.[84]

Professor Kip Viscusi discussed his recent research on jury competence. In one experiment, he tested jury responses to a defendant company's use of cost-benefit analysis. The use of cost-benefit analysis angered many of the jurors. When the hypothetical was changed by asserting that the company had placed a higher value on human life, this only increased the damage award. Professor Viscusi concluded that this was a result of the anchoring effect. He also tested the use of jury instructions based on the law and economics justification for punitive damages. The instructions and facts presented to the jurors provided a specific range of appropriate values for the punitive damage award. The jurors, however, generally failed to render judgments within the given ranges. This effect

---

[82] For a discussion of why uncertainty undermines the deterrence rationale for punitive damages, see Wheeler, *supra* note 35.

[83] All scholars do not agree with the assertion that punitive damages are erratic. One study, for example, found that punitive damages occur most frequently in cases "where breach of a legal duty suggests intentional or morally flawed behavior," and they are highly correlated with compensatory damage awards. Theodore Eisenberg et al., *The Predictability of Punitive Damages*, 26 J. LEGAL STUD. 623, 637 (1997); *cf.* Rustad, *supra* note 20, at 73 (arguing that punitive damages are generally awarded in cases where defendants knew of a product's risks and failed to remedy them or to warn consumers).

[84] Professor Schkade has made this proposal in previous articles. *See, e.g.*, Cass Sunstein, Daniel Kahneman & David Schkade, *supra* note 33, at 2114–26.

was exacerbated when the jurors were given anchors outside of the appropriate range. This led Professor Viscusi to the conclusion that juries are not necessarily competent to render punitive damages.[85]

Professor Carl Bogus presented a response to the criticism of juries. He started by pointing out that seventy-five to eighty percent of punitive damage awards are eliminated by judges.[86] This leads to two conclusions: (1) judges and jurors both play a role in determining punitive damages;[87] and (2) punitive damages are not as significant as a consideration of jury verdicts prior to remittitur might indicate. Professor Bogus also justified the failure of jurors to follow the instructions in experiments such as the one performed by Professor Viscusi. He asserted that jurors may understand that there is more to punitive damages than the factors contained in the jury instructions. For example, jurors may think that the focus of economists on the concept of the defendant's proper level of care is misapplied in cases in which the defendant holds other people's lives in its hands. Professor Bogus concluded that we should not remove the democratic element from punitive damage awards.

Dr. John Calfee joined other panelists in attacking the unpredictability of punitive damage awards. He offered the breast implant litigation as a paradigmatic example.[88] In that case, the defendant offered a four billion dollar settlement to all plaintiffs, from which thousands of plaintiffs opted out. Faced with the prospect of large jury verdicts, the company declared bankruptcy, despite the fact that some compensatory awards were later overturned. Dr. Calfee asserted that the defendant was fearful of punitive damages in the future. Because companies are unable to accurately assess either the risk of punitive damages or the likely magnitude of those damages, they are unable to make rational decisions. Dr. Calfee argued that supplying the decision-maker with a formula to determine punitive damages would avoid this unpredictability.

CONCLUSION

Overall, the Symposium provided a forum for scholars to debate the continued efficacy of punitive damages—a topic of particular significance given the recent criticism of punitive damages and accompanying pushes for reform. Critics have asserted that these supercompensatory awards are out of control and are not achieving the goals

---

[85] *Cf.* Sunstein, Kahneman & Schkade, *supra* note 33, at 2111 (indicating skepticism about the ability of jury instructions to produce damage awards based on principles of deterrence).

[86] *See also* Galanter & Luban, *supra* note 7, at 1408-09.

[87] This conclusion is supported in the literature. *See* Rustad, *supra* note 20, at 51–59 (indicating that judges frequently decrease runaway punitive damage awards).

[88] For a brief history of the breast implant litigation, see *In re* Dow Corning Corp., 1995 WL 495978 (Bankr. E.D. Mich. Aug. 5, 1995).

for which they were designed. The policy debate being fought in the legislatures and legal literature are symptoms of deeper empirical and theoretical disagreements. Until those disagreements are reconciled, the policy debate will likely continue.

—*Robert A. Klinck*

ARTICLE

PUNITIVE DAMAGES BY JURIES IN FLORIDA: *IN TERROREM* AND IN REALITY*

NEIL VIDMAR**
MARY R. ROSE***

In recent years there have been numerous proposals for punitive damage reform. Proponents of such reform have often asserted that punitive damages are both common and exorbitant. In this Article, Professor Vidmar and Dr. Rose examine the validity of these and other empirical claims about punitive damages. They do so by studying punitive damage awards reported in the Florida Jury Verdict Reporter. Ultimately, they conclude that there is no empirical support for the claims made by proponents of tort reform in Florida.

Tort reforms, particularly proposed changes to punitive damages, are on legislative agendas in a number of jurisdictions.[1] The calls for reform rest upon a multitude of empirical claims about the nature of awards and the risks of punitive damages that businesses currently bear. In this Article we examine more than a decade of actual punitive damage awards rendered by juries in Florida state courts. We closely examine such variables as the causes of action and the defendants against whom the punitive damages were assessed. We conclude that in Florida: (1) the frequency of punitive damages was strikingly low; (2) with the exception of asbestos cases, punitive damages were almost never given in products liability cases; (3) the relative amounts of punitive awards did not increase over the last decade; and (4) there is no evidence that juries award punitive damages capriciously and for minor forms of misconduct. These

---

* This research was supported, in part, by faculty research funds from Duke Law School. Preliminary results from the project were reported in Vidmar, Gross & Rose, *Jury Awards for Medical Malpractice and Post-Verdict Adjustment of Those Awards*, 48 DEPAUL L. REV. 265 (1998) and Rose, Vidmar & Gross, *Jury Awards in Products liability Cases and Post-Verdict Adjustments of Those Awards*. These results were orally presented at the Annual Law & Society Meetings, Miami, Florida, May 28, 2000. In July 2000 Professor Vidmar submitted an affidavit on behalf of The Coalition for Family Safety in *Florida Consumer Action Network et al. v. Bush*, No. 99-6689 (Fla. Cir. Ct.) that was partially based on data reported in this Article. He received remuneration for drafting the affidavit.

** Russell M. Robinson Professor of Law, Duke Law School. B.A., MacMurray College, 1962; M.A., Ph.D., University of Illinois, 1967.

*** Research Fellow, American Bar Foundation. B.A., Stanford University, 1991; M.A., Ph.D., Duke University, 1998.

[1] *This Should Be the Year for Product Liability Reform*, LANCASTER NEW ERA (Pa.), Feb. 19, 2001, at A6; Dan Morgan & Kathleen Day, *Early Wins Embolden Lobbyists for Business*, WASH. POST, Mar. 11, 2001, at A10; Bipin Avashiva, *Another Take on Medical Malpractice Reform*, CHARLESTON GAZETTE (W.V.), Mar. 17, 2001, at 5A.

findings challenge the rhetoric that accompanies much of the debate concerning the need for tort reform

I. JUSTIFYING REFORMS

Both the title and substance of this year's *Journal on Legislation* symposium reflect the fact that the attention of Congress and many states is once again focused on tort reform. Punitive damages play a central role in the debate. The reasons for attention to tort reform generally, and punitive damages in particular, are complex, but typically they are associated with concerted lobbying efforts by interested parties[2] or with media attention to particular cases, such as the now notorious McDonald's coffee burn case[3] and the $5 billion award against Exxon following the Exxon Valdez oil spill in Alaska.[4] Advocates of reform call for drastic changes in the current tort system, including statutory caps on the amounts that can be awarded[5] or removal of punitive award decisions from juries.[6]

The need for reform is typically premised upon claims about the widespread nature of the "problem" and its prevalence in the legislature's jurisdiction. For example, the legislative history of House Bill 775,[7] a tort reform bill passed by the Florida Legislature in 1999, includes testimony and position papers that made the following claims about punitive damages in Florida:[8]

- Punitive damages were being awarded frequently, particularly in products liability cases;
- Punitive damages in the latter part of the 1990s were awarded with greater frequency and in proportionately greater amounts than at the beginning of the 1990s;

---

[2] *See* STEPHEN DANIELS & JOANNE MARTIN, CIVIL JURIES AND THE POLITICS OF REFORM 29–59 (1995).

[3] Liebeck v. McDonald's Restaurants, 1995 WL 360309 (N.M. Dist. Ct. Aug. 18, 1994).

[4] *In re* Exxon Valdez, No. A89-0095-CV, 1995 U.S. Dist. LEXIS 12952 (D. Alaska Jan. 27, 1995).

[5] *See supra* note 1. *See generally* Theodore Eisenberg & Martin Wells, *Punitive Awards After BMW, a New Capping System, and the Reported Opinion Bias*, 1998 WIS. L. REV. 387.

[6] Reid Hastie & W. Kip Viscusi, *What Juries Can't Do Well: The Jury's Performance as a Risk Manager*, 40 ARIZ. L. REV. 901, 916 (1998); Cass Sunstein, Daniel Kahneman & David Schkade, *Assessing Punitive Damages (with Notes on Cognition and Valuation in Law)*, 107 YALE L.J. 2071, 2112–13 (2000). *See generally* Paul Mogin, *Why Judges, Not Juries, Should Set Punitive Damages*, 65 U. CHI. L. REV. 179 (1998).

[7] H.B. 775, 1999 Leg. (Fla. 1999).

[8] *February 3, 1999 Hearing Before Florida Senate Judiciary Committee*, 1999 Leg. (Fla. 1999); *February 2, 1999 Hearing Before Florida House Judiciary Committee*, 1999 Leg. (Fla. 1999); *Floor Debate on Conference Report on H.B. 775, Apr. 30, 1999* (transcript on file with authors); *Senate Floor Debate on Conference Report on H.B. 775, Mar. 9, 1999* (transcript on file with authors).

- Punitive damages were awarded in amounts that are both large at an absolute level and "disproportionate" when compared to compensatory awards, especially in products liability cases;
- Punitive damages repeatedly were being awarded against defendants for the same isolated and atypical course of conduct;
- Florida employers frequently were held vicariously liable for punitive damages for egregious acts of their employees or for criminal acts by third parties that occurred on business premises, even if the business owners had taken reasonable precautions;
- Florida juries were biased against "deep pocket" corporate defendants; and
- The "punitive damages crisis" put Florida businesses at a comparative disadvantage to their competitors in other states and stifled "innovation and economic development."

This last assertion involves an *in terrorem* claim; it is a claim that the very threat of punitive damages has led to negative economic consequences.[9]

The 1999 Florida tort reform legislation was declared unconstitutional on the grounds that the state constitution bans laws that embrace more than one subject, but efforts to save the reform continue.[10] In the meantime, tort reform has arisen anew with Florida nursing home interests asserting the need for reform. The claims in this latest debate again involve punitive as well as compensatory damages and *in terrorem*

---

[9] *See February 3, 1999 Hearing Before Florida Senate Judiciary Committee*, 1999 Leg. (Fla. 1999).

> The fact that juries award punitive damages relatively infrequently is of no consequence. The mere threat of punitive damages inflates the cost of settlement and undermines notions of fairness and judicial efficiency. Even worse, the imposition of excessive [punitive damages] awards threatens to remove safe products from the market and to deter innovation in the market place.

George N. Meros, Jr. & Chanta G. Hundley, *Florida's Tort Reform Act: Keeping Faith with the Promise of* Hoffman v. Jones, 27 FLA. ST. U. L. REV. 461, 478 (2000); *see also* George N. Meros, Jr., *Toward a More Just and Predictable Civil Justice System*, 25 FLA. ST. U. L. REV. 141, 154, 155 (1998) (describing punitive damages as out of control). Mr. Meros's views are significant as he has stated that he represented the Florida Chamber of Commerce, *id.*, and he served as "Special Counsel, T.R.U.E. Coalition and affiliated business interests in tort reform efforts before the Florida Legislature." Meros & Hundley, *supra*, at 461.

[10] *See* Susan Strother Clarke, *Shop Talk Column*, ORLANDO SENTINEL (Fla.), Mar. 5, 2001, *available at* 2001 WL 15016347.

claims.[11] Although our research is centered on Florida, we note that similar claims about the terror punitive damages allegedly create in businesses have been made elsewhere.[12]

The central question is whether these claims have an empirical foundation. If the claims are valid in whole, or even in part, there is serious cause for concern and remedial measures would seem to be in order. However, if the claims are without foundation, there would seem to be no rational basis for change. Even the alleged threat of punitive damages perceived by businesses must be reevaluated if data do not support these fears because the proper remedy would be education of businesses about the actual risks.

We now turn directly to the task of providing data from Florida that reflect on the claims made regarding punitive damages. Later, we will place our findings in the context of other empirical studies of punitive damages and discuss the broader implications of our findings.

II. THE DATA

Our data set was obtained from the Florida Jury Verdict Reporter archived on Westlaw. For each case there is a basic summary of the claims involved at trial. In all but a few instances, the reports provide a breakdown of the award into its compensatory and punitive components, and, if there was more than one defendant, the amounts assessed against each defendant. Verdict reporters may exclude some cases. Research on other jury verdict reporters, however, indicates that to the extent that cases are omitted by reporter compilers, the missing cases tend to be those in which the defendant prevails or in which the plaintiff's award is small.[13] Thus, if there is a bias in the data, it is not likely to be one of omission of large punitive awards.

To draw our sample we searched the data base using the search cues "punitive" and "verdict" for the years 1988 through June 2000. The case summaries in the 1988 reports were more truncated than in subsequent

---

[11] Sarah Skidmore, *Today's Topic: Tort Reform: Legislators Consider Insurance Debate*, FLA. TIMES UNION (Jacksonville, Fla.), Mar. 18, 2001, at A1; Marcia Mattson, *Florida's Nursing Home Industry Says It's Near Collapse*, FLA. TIMES UNION, Mar. 18, 2001, at A1; Matthew Pinzur, *Lobbyists in Furious Battle: Big Business, Trial Lawyers Turn Debate From Care to Tort Reform*, FLA. TIMES UNION, Mar. 25, 2001, at A1.

[12] *See generally* Herbert M. Kritzer & Frances Kahn Zemans, *The Shadow of Punitives: An Unsuccessful Effort to Bring It into View*, 1998 WIS. L. REV. 157; Thomas Koenig, *The Shadow Effect of Punitive Damages on Settlements*, 1998 WIS. L. REV. 169; Steven Garber, *Product Liability, Punitive Damages Business Decisions and Economic Outcomes*, 1998 WIS. L. REV. 237.

[13] *See, e.g.*, A. Russell Localio, *Variations on $962,258: The Misuse of Data on Medical Malpractice*, 13 LAW MED. & HEALTH CARE 126, 126 (1985); Deborah Jones Merritt & Kathryn Ann Barry, *Is the Tort System in Crisis? New Empirical Evidence*, 60 OHIO ST. L.J. 315, 326 (1999); Neil Vidmar, *Pap and Circumstance: What Jury Verdict Statistics Can Tell Us About the Tort System*, 28 SUFFOLK U. L. REV. 1205, 1209–10 (1994).

years, suggesting that the reporting system was in the process of development and raised concerns that the compilation of cases may not have been as comprehensive as later years. Additionally, there were fewer cases for 1999 and 2000, which may indicate incomplete data for those years at the time of our search. Below, we report verdicts for all years (1988 through 2000) when the data are aggregated across years; however, any longitudinal analyses—i.e., examinations of change over time—are restricted only to years 1989 through 1998. In this way, we do not inadvertently report fewer punitive awards per year than actually may have occurred.

The search produced a substantial number of cases that we eliminated from the analysis after a careful reading of the summaries. Some cases were removed because they were decided in federal courts, and the present study focuses on cases adjudicated in state courts.[14] In some instances, the plaintiff asked for punitive damages in the pleadings, but the verdict on liability was for the defendant. In other instances, the plaintiff asked for punitive damages, but the judge concluded that punitive damages were inappropriate at the beginning of trial or at the conclusion of the evidence. Our data set includes only those cases in which the issue of punitive damages was put to the jury—that is, cases in which the jury either awarded or denied punitive damages.

Both authors read each case summary and independently developed classification systems for the cases, examining the main basis on which punitive damages were awarded. There was substantial agreement as to which categories were reflected in the cases; classification disagreements were resolved through discussion.[15]

---

[14] In the total data set there were twenty-two cases in Florida's federal courts. Fourteen of these cases involved civil rights claims related to employment; three involved suits against law enforcement for violations of civil rights; two concerned charges of slander or defamation that involved federal laws; one was a premises liability case at an amusement park; one was a legal malpractice claim involving diversity jurisdiction; and one consisted of a suit against the Cuban government. The median total award for federal cases was $913,078; the median punitive award was $346,806. The median ratio of punitive to compensatory award (discussed below for the state data) was 0.82:1.

[15] Each author initially developed his or her own classification scheme and coded the cases into categories. The separate coding schemes were, in fact, remarkably similar. For example, the first author had only two major categories that the second author had not considered; likewise the second author had only one category not considered by the other. In addition, we developed slightly different sub-categories (e.g., one had a separate sub-category for sexual harassment under employment, whereas the other coded these cases into a general discrimination sub-category). As there are often multiple causes of action and multiple defendants, categorization was focused upon the basis for which the jury awarded, or was to consider awarding, punitive damages. Thus, a case might involve an assault, as well as a premises liability claim; however, if punitive damages were not assessed against the owner of the premises, the case was categorized as an assault case. The final classification system was developed through discussion of the two systems, considering cases in the light of: (1) the injury involved in the case (e.g., financial damage, violations of privacy); (2) a unique setting or legal issue shared by the cases (e.g., employment-related disputes); and (3) the claims made in the legislative history of House Bill 775 (e.g.,

III. RESULTS: A PROFILE OF PUNITIVE DAMAGES IN FLORIDA

The Florida data set does not provide support for the claims made by tort reformers. In fact, the data indicate that punitive damages were rarely awarded, punitive damages were almost never given in product liability cases,[16] punitive awards are not on the rise, and cases involving punitive awards report instances of serious misconduct.

TABLE 1. NUMBER OF PUNITIVE DAMAGE CASES INVOLVING JURIES, BY YEAR (1989–1998)

Year	Number of Cases	Number with Non-Zero Awards	Median Ratio
1989	32	27 (84%)	0.46:1
1990	27	26 (96%)	0.17:1
1991	28	25 (89%)	0:83:1
1992	22	19 (86%)	0.52:1
1993	21	19 (90%)	0.55:1
1994	27	26 (96%)	0.93:1
1995	15	13 (87%)	0.92:1
1996	17	17 (100%)	1.13:1
1997	21	17 (81%)	0.40:1
1998	22	19 (86%)	0.90:1
As of 1998	23.2/year	20.8/year	0.67:1

A. Per Annum Overview

Table 1 reports by year the total number of punitive damage claims between 1989 and 1998 that were put to a jury (Column 1), the number

---

we ensured that there was a distinct category for products and premises liability), *see supra* note 8. A third person, a lawyer, also examined a sample of the cases and arrived at classifications that were highly similar to those of the authors. Because we developed categories while simultaneously coding cases into these categories, traditional estimates of coding reliability (e.g., kappa) are not applicable. However, in all only about 20% of the cases were shifted around following our resolution of the two coding systems, a remarkably low number given our independent approach to the cases. We retain a list of the cases in the data set on file so that our classifications can be checked by other interested parties.

[16] As is discussed below, the only cases involving a non-zero punitive award involved asbestos or tobacco.

and percentage of times that the jury returned a punitive award (Column 2), and the median ratio of punitive to compensatory damages (Column 3).[17]

Given the rhetoric contained in the legislative history of Florida House Bill 775, Table 1 shows a number of surprising things. First, whether we consider the total number of cases allowing punitive damages or just the number in which the jury actually awarded damages, the average (or mean) number of punitive awards annually was very low: on average there were 23.2 cases per year in which punitive damage claims were put to the jury and 20.8 cases per year in which the jury rendered a punitive award. Between 1989 and 1998, there was a significant downward trend in the number of punitive damage cases per year that were put to juries.[18] Considering the data another way, there were fewer punitive awards during the three years from 1996 through 1998 (an average of twenty cases per year) than for the three years from 1989 through 1991 (an average of twenty-nine cases per year). These figures are striking when we consider that the Florida Legislature's Office of Economic and Demographic Research reports that Florida's population rose 13.7% between the years 1989 to 1991 and the years 1996 to 1998.[19] Thus, the number of punitive awards per capita in the 1989 to 1991 period was 2.0 per 100,000 persons, whereas the per capita number of punitive awards in the 1996 to 1998 period was only 1.1 per 100,000 persons.

Column 3 shows the median ratio of punitive to compensatory damages. Over the period from 1989 through 1998, the median for this ratio was 0.67:1, meaning that the punitive damage component was two thirds of the compensatory award. Because it is the median value, half the awards in the data set had smaller ratios—i.e., the punitive portion was a smaller percentage of the compensatory portion. The table indicates that this median ratio varied substantially from year to year, with no clear pattern of increase. A closer examination of the cases suggests that this variability in the median across years is best explained by the fact that the small numbers of punitive damage cases per year differ markedly in their distributions of case-type. As we discuss in the next section, the

---

[17] It bears repeating that Table 1 represents only cases with verdicts for the plaintiff. For instance, in 1989 juries returned with a non-zero punitive award in 84% of cases in which the defendant was liable for damages, not in 84% of all punitive damage cases, which would include verdicts for the defense.

[18] A regression model examined whether the year of the case predicted the count of cases per year. The parameter reflecting a decline in the frequency of cases across years was statistically significant, $b = -1.22$, $t = -2.82$, $p < .05$. This result should be treated with caution however, because it is sensitive to the exclusion of the year 1988. If the years 1988 through 1998 are considered, the downward trend does not achieve conventional levels of statistical significance, $p < .07$.

[19] In the 1989 to 1991 period the average population was estimated at 12,893,333 and in the 1996 to 1998 period it was estimated to be 14,708,300. *See* FLA. LEGISLATURE OFFICE OF ECON. & DEMOGRAPHIC RESEARCH, STATE POPULATION BY AGE, RACE, SEX (on file with authors).

ratio of punitive to compensatory awards differs across our categorization of cases; for example, it is very low for motor vehicle accidents and comparatively high for cases involving financial fraud or employment discrimination. To the extent that many of the cases in a given year involve motor vehicle accidents, the median ratio will likewise be low. For example, in 1990 44% of cases were motor vehicle accidents and the median ratio was 0.17:1. The reverse is true of a year such as 1996, in which 47% of the cases concerned discrimination claims and the median ratio was at its highest, 1.13:1. Thus, a perceived change in the ratio of punitive damages to compensatory damages most likely does not reflect a change in jury behavior but rather a change in the types of cases juries hear.[20]

In short, the data in Table 1 lend no support to the claim that the frequency of punitive damages by juries was high at an absolute level or that juries became more likely to award punitive damages. To the contrary, the data indicate that levels were low—indeed, strikingly so when we make a per capita comparison—and the trend was in the opposite direction. Additionally the data on median punitive award ratios lend no support to the claim of an increase in the ratios of punitive to compensatory damages over the years. Indeed, the median punitive damage award was a modest two-thirds of the median compensatory damage award.

B. Case Types by Causes of Action

Table 2 analyzes the total set of data (1988 through June 2000), considering our categories for the main causes of action in which punitive damages were awarded.[21] We developed eleven different types of case categories, and Table 2 orders them in terms of their frequency. The most frequent case type is motor vehicle accidents caused by impaired or reckless drivers and accounts for 23.3% of all punitive damage award cases. The second most frequent type involves fraud or financial loss (17.4%), which is a broad category involving breach of contract cases (e.g., with respect to business transactions or insurance coverage), fraud and deceptive trade practices, tortious interference, trademark infringement, and property damage. Third is a category involving sexual and physical assaults, including civil suits related to homicides (15.9%). Almost 57% of punitive damage awards result from these three case types, with the remaining awards distributed over the other eight categories.

---

[20] Vidmar, *supra* note 13, at 1213–16.
[21] *See supra* note 15.

TABLE 2. DISTRIBUTION OF CASES BY CASE TYPE, 1988–2000 (N = 270 TOTAL)

Category	Number of Cases	Percent of Total
Motor vehicle accidents/impaired drivers	63	23.3
Fraud, financial losses	47	17.4
Assaults (physical and sexual)	43	15.9
Products liability	20	7.4
Information violations	20	7.4
False imprisonment/false arrest	20	7.4
Premises liability	17	6.4
Discrimination/harassment	13	4.8
Professional negligence (medical care)	12	4.4
Workplace injuries/failure to pay benefits	11	4.1
Improper treatment of dead persons	4	1.5

Although product liability and premises liability cases were the subject of much discussion in the legislative history of House Bill 775, there were just twenty products liability cases and seventeen premises liability cases between 1988 and 2000. Taken together these cases account for less than 14% of the total (7.4% + 6.3% = 13.7%). We discuss these two categories in more detail in the next section. Two other categories, information violations and false arrest/imprisonment, each had twenty cases. The information violations category reflects both harms to privacy through improper releases of confidential information,[22] as well as slander, defamation, or libel. The false imprisonment/false arrest category covers, for example, cases in which a person was wrongfully detained on suspicion of shoplifting or other thefts; accusations that someone maliciously pursued wrongful criminal charges against the plaintiff; police misconduct; and in one instance, a case in which a South American airline responded to disgruntled passengers by removing them from the plane during a stop-over, detaining and strip-searching them, and leaving them stranded in a foreign country.[23]

[22] In one case, a business shared a credit report with another business in violation of the law governing releases of this information. Murphy v. Elebash, No. 95-1221-CA-01, 1998 WL 775524 (Fla. Cir. Ct. Sept. 1, 1998).

[23] Herrera v. Zuliana de Aviacion, No. 94-15999-CA-01, 1998 WL 1059885 (Fla. Cir.

The remaining categories described in Table 2 each represent less than five percent of the cases in the data set. The discrimination/harassment category involves thirteen cases of employment discrimination or harassment (4.8%).[24] The professional negligence category includes three cases concerning professional negligence in the provision of medical care, which encompasses a small number of medical malpractice claims, one case of negligent and abusive care at a drug treatment facility, and eight nursing home negligence cases (two of which resulted in no punitive damage award). There were eleven cases of workplace injuries or failure to pay benefits following work-related accidents.

Finally, a small number of cases necessitated the creation of the "improper treatment of dead persons" category. All of these cases involved emotional distress claims brought by families of people whose bodies were mishandled after death. In two instances, a crematorium was found to have mixed together the ashes of several people, along with dirt from the ground, and then to have represented the motley concoction as a single deceased person's ashes.[25] In another case, a husband discovered that a funeral home misplaced his wife's amputated legs which were to be kept in cold storage until she died so that her entire body could be buried in accordance with Orthodox Jewish tradition.[26] Lastly, the parents of a deceased Haitian boy claimed that a hospital harvested his organs without their consent.[27]

C. Products and Premises Liability Cases

As the above review makes clear, punitive damage cases before juries in Florida involved a wide variety of misconduct allegations, most often concerning drunk driving, financial improprieties, or assaults. However, the Florida legislature focused much of its justification for punitive damage reforms on harms to businesses, especially burdens associated with products or premises liability. Because of this focus, these cases deserve further examination and elaboration.

Punitive damages were awarded in only sixteen of the twenty products liability cases in which the jury was allowed to consider punitive damages. More interestingly, however, is the paucity of cases that do not involve asbestos. All but one of the product liability cases with a punitive

---

Ct. Dec. 19, 1998).

[24] There were six sexual harassment cases in state courts. All but one were against a single doctor accused of mistreating several employees over many years.

[25] Read v. Cremation Sys. Int'l, No. 95-6994-CI-11, 1998 WL 735112 (Fla. Cir. Ct. May 11, 1998); Smith v. Cremation Sys. Int'l, No. 96-6141-CI-19, 1998 WL 735113 (Fla. Cir. Ct. Apr. 15, 1998).

[26] Dresin v. Menorah Gardens & Funeral Chapels, No. 94-2125-09, 1997 WL 817902 (Fla. Cir. Ct. May 15, 1997).

[27] Vernet v. Univ. of Miami, No. 93-3386-CA, 1998 WL 1059895 (Fla. Cir. Ct. July 24, 1998).

award involved asbestos injury claims, with the remaining case involving manufacturers' liability for cigarettes.[28] The jury granted no punitive award in three asbestos cases; the remaining case involving no punitive award was a suit against General Motors after a child died from the explosion of a car in an accident. Given all of the claims concerning the role of punitive damages in product liability cases in Florida and elsewhere, these low frequencies are surprising yet apparently correct.[29]

Next, consider premises liability, a category that includes instances of employers being held accountable for the acts of their employees. Our categorization scheme resulted in seventeen premises liability cases, fourteen of which resulted in a punitive award. The details of plaintiffs' allegations provided in the case summaries from the Florida Jury Verdict Reporter provide some context for the awards.

In one case, the plaintiffs alleged that a manufacturing plant had been illegally dumping toxic chemicals into the trash dumpster on their property, which was located near a residential neighborhood.[30] The company had been cited previously for this behavior and claimed to have ceased such activities.[31] When a nine-year-old boy and his friend entered the property and climbed in the dumpster to play, he and the friend were overcome by toxic fumes and died.[32] In another case involving exposure to toxins, the captain of a cargo ship created a dangerous situation by ignoring warnings about using a pump in an unventilated area.[33] When his own crewmen were overcome by carbon monoxide, the captain called the Coast Guard for help but did not inform them that highly toxic materials were involved.[34] The fumes killed one of the rescuers as he tried to help the injured crewmen.[35]

---

[28] Even this latter case was not a straightforward victory. In *Maddox v. Brown & Williamson Tobacco Corp.*, No. 97-03522-CA, 1998 WL 933419 (Fla. Cir. Ct. June 10, 1998), an appellate court ruled that the trial court abused its discretion by failing to change venue.

[29] We returned to Westlaw and searched the Florida Jury Verdict Reporter for all jury verdicts involving products liability for the years 1989 through 1998. On average Florida had thirty-nine cases per year that reached the level of litigation necessary to be included in the verdict reporter. This figure may underestimate the actual number of filed cases since some may disappear through settlement or dismissal early in the litigation process. There was an average of thirty-four cases per year that went to trial during the period we investigated. There was a trend toward fewer cases near the end of the decade as compared to the beginning. As with rates of punitive damages, these figures should be considered in light of the fact that Florida's population increased substantially over this period, leading to the conclusion that on a per capita basis product liability lawsuits actually declined. Our search for products cases produced no other awards for punitive damages, apart from the asbestos and cigarette cases reported above.

[30] Perez v. William Recht Co., No. 92-8983, 1995 WL 861061 (Fla. Cir. Ct. Sept. 28, 1995).

[31] *Id.*

[32] *Id.*

[33] Taibl v. Juno Marine Agency, Inc., No. 88-45327-CA-22, 1998 WL 355212 (Fla. Cir. Ct. Feb. 23, 1998).

[34] *Id.*

[35] *Id.*

In *Montalvo v. Rancho El Nuevo Mundo, Inc.*,[36] defendants allowed a group of minors to throw a "rage party" on their horse ranch. Alcohol was served during the event, and one attendee was attacked and killed during the party by another group of youths.[37] The jury in *Palank v. CSX Transportation, Inc.*[38] found the CSX railway system liable for a fatal train derailment and assessed punitive damages for failing to maintain the railroad tracks. Although the company claimed in written reports to have inspected the tracks twice a week, physical evidence suggested they had not been examined in as long as a year.[39] A furniture company faced punitive damages in *Harrison v. Tallahassee Furniture Co., Inc.*[40] when one of its deliverymen stabbed a female customer multiple times. The delivery man had a long criminal record as well as a history of drug abuse and hospitalizations for paranoid schizophrenia.[41] In *Kinder v. International Union of Operating Engineers, Local 765*,[42] a union local was held responsible for an attack on a man who refused to organize his non-union company.

A plaintiff was awarded punitive damages against an apartment complex owner in *Clavel v. E.G. Goldsmith*.[43] The plaintiff was raped in her apartment after having made repeated requests that the apartment managers fix broken window locks.[44] In another sexual assault case, a church was assessed punitive damages when its pastor induced a mentally retarded teenage boy to smoke crack with him, then sexually assaulted the teen and worked with a deacon in the church to have the young man publicly recant accusations of the multiple episodes.[45]

Three premises liability cases involved dog bites, one of which was to a babysitter in a private home[46] while the other two involved customers being attacked at a place of business.[47] In *Lycans v. Moore*, the defendant, the business's owner, told the customer that the dog was secure and to ignore the "Bad Dog" sign on the property.[48] The customer was then attacked by a German Shepherd.[49]

---

[36] No. 96-16035-CA-22, 1998 WL 555151 (Fla. Cir. Ct. Apr. 28, 1998).
[37] *Id.*
[38] No. 92-22407(13), 1995 WL 683935 (Fla. Cir. Ct. July 30, 1997).
[39] *Id.*
[40] No. 86-1860, 1989 WL 527555 (Fla. Cir. Ct. Sept. 1989).
[41] *Id.*
[42] No. 86-09959-CV, 1988 WL 502870 (Fla. Cir. Ct. July 1988).
[43] No. 88-22407, 1989 WL 527479 (Fla. Cir. Ct. Oct. 1989).
[44] *Id.* The jury awarded punitive damages although it also found the plaintiff 30% responsible for the injury. *Id.*
[45] Doe v. Masters, No. CL-92-6875-AD, 1998 WL 735101 (Fla. Cir. Ct. May 28, 1998).
[46] Winbigler v. Commander, No. 87-15767, 1988 WL 504496 (Fla. Cir. Ct. Dec. 1988).
[47] Greely v. Twidy, No. 92-21746 (13), 1993 WL 806978 (Fla. Cir. Ct. Mar. 16, 1993); Lycans v. Moore, No. 91-3268, 1991 WL 713035 (Fla. Cir. Ct. Oct. 1991).
[48] No. 91-3268, 1991 WL 713035 (Fla. Cir. Ct. Oct. 1991).
[49] *Id.*

Two premises liability cases involved the consumption or presence of alcohol. In *Niesen v. McNally*,[50] a bouncer and other employees allegedly battered and ejected a female patron after she screamed about unwanted sexual advances from another customer; this customer, also ejected, then physically assaulted the woman in the bar's parking lot.[51] Last, in *Estate of Fields v. Mills*,[52] jurors awarded punitive damages against a company when its employee drank excessively at a business function in the presence of company executives, and where the company paid for and expensed the drinks. That same night, the employee killed someone while driving under the influence of alcohol.[53]

In sum, although punitive damage awards are indeed assessed against businesses, the circumstances typically suggest serious misconduct on the part of defendants or their employees. The Florida legislature was concerned that businesses are held responsible even when they have taken reasonable precautions.[54] Notably, however, cases with non-zero punitive damage awards against businesses all tended to involve allegations of either a knowing and active disregard for the law—e.g., illegal toxic dumping, allowing minors to have alcohol, falsely reporting inspections—or misconduct by people in senior positions—e.g., a ship's captain, a business owner, corporate executives. In short, these cases rarely involved businesses engaging in commonly accepted practices or those having taken normal precautions for safety.

Additionally, it bears mentioning that businesses are not the most frequent targets of punitive damage awards. Juries awarded punitive damages against businesses in 129 cases, which is 48% of all cases.[55] This is likely due to the fact that for two of the three most frequent punitive damage categories in our classification—impaired/reckless driving and assaults—individual offenders predominate. Further, when a business and an individual are jointly named in a lawsuit, it does not follow that the business will necessarily be liable for the punitive portion of the award. Of the 78 cases in which both an individual and a business were named as being responsible for an injury, 28 cases (36%) resulted in individuals being solely responsible for the punitive portion of the award. There are several reasons why a business may not be exposed to punitive damages in such cases: as a matter of law the business's conduct may not rise to a level warranting punitive damages;[56] the jury may find no liabil-

---

[50] No. 95-1793 AN, 1996 WL 901937 (Fla. Cir. Ct. Oct. 21, 1996).

[51] *Id.*

[52] No. 86-11216 01, 1991 WL 712868 (Fla. Cir. Ct. Nov. 1991).

[53] *Id.*

[54] *See supra* notes 10–11 and accompanying text.

[55] It is, however, 74% of those in which businesses were named

[56] Mercury Motors Express, Inc v. Smith, 393 So. 2d 545, 549 (Fla. 1981) (holding that while "an employer is vicariously liable for compensatory damages resulting from the negligent acts of employees committed within the scope of their employment even if the employer is without fault," vicarious liability for punitive damages requires "some fault"

ity on the part of the business; or the jury may find the business liable for compensatory damages but decide not to hold them responsible for punitive damages. In short, a claim against a business for punitive damages does not automatically mean the business will be liable for punitive damages.

D. Variability of Awards by Causes of Action

Table 3 presents information, disaggregated by type of case, on the median total award, the median punitive awards, and the median ratio of punitive to compensatory damages.[57] All these measures differed markedly across categories. The lowest median punitive award, along with the lowest ratio of punitive to compensatory damages, appeared in the motor vehicles category, suggesting that punitive damages function as a sort of civil fine levied by the jury as a punishment, but not an unduly harsh one. In contrast, both in absolute terms and relative to compensatory awards, punitive damages were highest in the four cases involving improper treatment of dead persons (6.3:1). This result should not be surprising given that American law has long recognized that important cultural and religious values are associated with mistreatment of the dead.[58]

For the majority of categories, the median ratio of punitive damages to compensatory awards was at or below 1:1, indicating that punitive damages did not exceed compensatory damages. This majority includes both product and premises liability—two areas often invoked as proof of the need for reform. In four of the categories, compensatory damages were approximately twice the punitive damages.

Discrimination/harassment and professional negligence cases yielded median ratios of 2.3:1 and 2.5:1, respectively, which means that punitive damages were, on average, slightly more than twice compensatory damages. In addition, these categories, both of which represent violations of professional norms, had total awards in excess of $1 million, which was well above the overall median of $612,028.

---

on the part of the employer).

[57] With few exceptions, median and mean ratios across categories did not differ substantially.

[58] *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 868 (1977); Arnold v. Spears, 63 So. 2d 850 (Fla. 1953).

TABLE 3. DISTRIBUTION OF TOTAL, PUNITIVE AWARDS, AND THE MEDIAN RATIO OF PUNITIVE TO COMPENSATORY AWARDS, BY CASE TYPE (1988–2000)

Type of case	N cases	Median Total Award	Median Punitive Award	Pun:Comp Ratio (median)
Motor vehicle accidents	63	284,736	21,579	0.1:1
Fraud, contract violation, and other financial damage cases	47	392,158	318,055	1.0:1
Assaults	43	221,461	59,832	0.4:1
Products liability	20	2,245,635	666,936	0.8:1
Information violations (privacy, slander, defamation, libel)	20	191,264	108,530	1.1:1
False imprisonment/ false arrest	20	234,752	139,814	0.4:1
Premises liability	17	933,660	200,081	0.5:1
Discrimination/ harassment	13	1,344,841	1,030,530	2.3:1
Professional negligence	12	3,078,133	1,006,172	2.5:1
Workplace injuries/ failure to pay benefits	11	317,260	71,820	0.5:1
Other: Improper treatment of dead persons	4	3,434,572	3,052,075	6.3:1
Overall	270	612,028	151,871	0.7:1

Note: Awards adjusted to 1999 dollars.

We looked more closely at the professional negligence category of cases in light of the most recent debate in Florida regarding nursing home business concerns about their liability exposure and about punitive damages.[59] There were only six professional negligence cases in our sample that involved jury awards of punitive damages. In *Estate of Collins v. Beverly Enterprises-Florida, Inc.*,[60] a $10 million punitive damage award was assessed against a nursing home, resulting in a punitive to compensatory ratio of 4.9:1. The patient's family claimed that their mother had been physically assaulted once by a demented patient who was permitted to roam the halls; that she had been sexually assaulted on two separate occasions; that there were other injuries that were consistent with physical assaults; and that her treating physician did not properly treat a massive infection that ultimately led to her death.[61] There were two 1998 cases. In *Canady v. Manor Healthcare, Inc.*,[62] a case involving a resident who suffered serious infections, punitive damages were assessed in the amount of $567,000. The punitive to compensatory ratio was 2.7:1. In the second case, *Estate of Barnes v. First Healthcare Corp.*,[63] a patient with dementia and a history of wandering away left the facility, fell into a pond and drowned. This resulted in a punitive award of $4,500,000, which represents a 2.4:1 punitive to compensatory ratio.[64]

There were two 1993 cases. In *Estate of Spilman v. Beverly Enterprises-Florida, Inc.*,[65] $2 million in punitive damages were assessed against a nursing home that left the plaintiff, a former resident, unattended and in his own waste for extended periods of time, resulting in bedsores and malnutrition. This award represented a 2.8:1 ratio.[66] In the other case, *Jones ex rel. Clark v. Clearwater Convalescent Center, Inc.*,[67] the resident suffered pressure ulcers, a hip fracture, and amputation of a leg below the knee. The punitive award was $300,000, representing a 1:1 punitive to compensatory ratio.[68]

Finally, *Wanderon v. Unicare Healthcare Facilities, Inc.*[69] involved a resident who suffered ulcers, scabies, infections, and fractures which led to amputation of a leg. There were also accusations that leaving the patient to lie in her own waste contributed to her medical problems.[70] The

---

[59] *See supra* note 11.

[60] No. 98-433-CA, 2000 WL 1203896 (Fla. Cir. Ct. Mar. 31, 2000). The dollar figures presented for this case and those that follow are in actual dollars.

[61] *Id.*

[62] No. 97-04349-CA, 1998 WL 1021519 (Fla. Cir. Ct. Dec. 14, 1998).

[63] No. CL 97-1621 AF, 1998 WL 355241 (Fla. Cir. Ct. Feb. 10, 1998).

[64] *See id.*

[65] No. 92-1345-CA-01 Div. B, 1993 WL 813877 (Fla. Cir. Ct. Dec. 8, 1993).

[66] *See id.*

[67] No. 91-7612-15, 1993 WL 807055 (Fla. Cir. Ct. Feb. 5, 1993).

[68] *See id.*

[69] No. CA-90-1645, 1992 WL 737242 (Fla. Cir. Ct. Jan. 1992).

[70] *Id.*

punitive award was $1,250,000, representing a punitive to compensatory ratio of 0.9:1.[71]

These summaries are not a substitute for the evidence heard by the juries, but they present a prima facie case that the awards could have been reasonable. In any event it is worth reiterating that over a period of twelve and a half years there were only six punitive damage awards in medical negligence cases.

E. "Mega" Awards

Although the ratio between the punitive and compensatory awards is probably the best yardstick to gauge the excessiveness of punitive damage awards, or lack thereof, it can obscure some large awards and variability within award categories. Despite the average punitive award being a modest sixty-eight percent of the compensatory award, there are some very large punitive awards in Florida. We identified the twenty largest of these awards (converted to 1999 dollars), and these are reported in Table 4, which provides the case citations, the total award, the punitive award, and a description of the cause of action.

In six of the cases (listed in Table 4 with a superscripted "a"), there is some indication that the defendant was not represented by a lawyer at trial.[72] In a seventh case the over $300 million award was against a company that had filed for bankruptcy and was facing criminal charges.[73] Thus, for these seven cases, it would appear that the chances of a plaintiff recovering the punitive award were nil or nearly so.

---

[71] *See id.* Another case related to elder care was coded in the employment category. In *Carr v. Orlando Regional Healthcare System, Inc.*, No. CI 95-142, 1996 WL 901962 (Fla. Cir. Ct. Nov. 13, 1996), an employee was terminated from a home health care business after filing a report of elder abuse. The case resulted in a punitive award of $40,000 to the employee, representing a 0.6:1 punitive to compensatory ratio.

[72] In four cases, the verdict reporter explicitly stated that the defendant either defaulted or was unrepresented at trial; in two others, the reporter put "n/a" in the space for defense counsel's name.

[73] Estate of Perez v. William Recht Co., No. 92-8983, 1995 WL 861061 (Fla. Cir. Ct. Sept. 28, 1995).

TABLE 4. HIGHEST 20 PUNITIVE AWARDS IN FLORIDA SAMPLE

Plaintiff	Total Award	Punitive Award	Description
Perez[74]	542,650,919	325,590,551	Illegal disposal of toxic chemicals; child died
Chipps[75]	77,097,948	76,018,150	Insurance coverage breach disabling child
Palank[76]	57,855,300	51,526,480	Train accident/ tracks not maintained
Intl. Ship Repair[77]	43,676,246	41,221,184	Insurance coverage breach, floating dry dock sunk
Ballard[78]	33,847,745	31,946,417	Asbestos
Read[79]	33,993,252	30,441,718	Crematorium gave back mix of ashes[a]
Scheller[80]	25,621,292	25,489,506	Interference with a doctor's medical practice
Van Dyk[81]	28,615,917	22,892,734	Escaped convict shot patron in bar (award is against shooter only)[a]
Rawson Food[82]	32,996,112	22,370,245	Investment banking negligence

[74] Estate of Perez v. William Recht Co., No. 92-8983, 1995 WL 861061 (Fla. Cir. Ct. Sept. 28, 1995).

[75] Chipps v. Humana Health Ins. Co. of Fla., Inc., No. CL 96-00423 AE, 2000 WL 730640 (Fla. Cir. Ct. Jan. 4, 2000).

[76] Estate of Palank v. CSX Transp., Inc., No. 92-22407 (13), 1995 WL 683935 (Fla. Cir. Ct. July 30, 1997); Case Updates, 1999 WL 1289390 (1999) (indicating that the jury in Palank awarded $50,000,000 in punitive damages).

[77] Int'l Ship Repair & Marine Serv., Inc. v. St. Paul Fire & Marine Serv. Ins. Co., No. 94-1368-CIV-T-17C, 1997 WL 862900 (Fla. Cir. Ct. Nov. 21, 1997).

[78] Ballard v. Owens Corning, No. 93-10817 AD, 1997 WL 335698 (Fla. Cir. Ct. Jan. 21, 1997).

[79] Read v. Cremation Sys. Int'l, No. 95-6994-CI 11, 1998 WL 725112 (Fla. Cir. Ct. May 11, 1998).

[80] Scheller v. Am. Med. Int'l, Inc., No. 80-5519 CA(L) 01 AF, 1989 WL 527424 (Fla. Cir. Ct. Mar. 1989).

[81] Van Dyk v. Newell, No. 92-01893-CA, 1993 WL 807893 (Fla. Cir. Ct. Aug. 3, 1993).

[82] Rawson Food Serv., Inc. v. Pantry Pride Enter., Inc., No. 86-2324, 1988 WL 504543 (Fla. Cir. Ct. Nov. 1988).

Montenegro[83]	23,073,421	18,215,859	Civil suit following a murder conviction
Dudley[84]	21,495,304	16,740,891	Asbestos[b]
Ferguson[85]	17,683,535	16,150,962	Tortious interference in a van line business[b]
Goldberg[86]	12,509,680	12,123,008	A doctor sexually harassed an employee[a]
Caron[87]	13,596,320	11,789,024	Employee fell to his death during roof installation
Wheeland[88]	20,603,460	11,446,367	A woman infected her husband with HIV[a]
Lowell[89]	13,129,242	10,123,948	Asbestos
Collins[90]	11,641,556	9,683,841	Patient-on-patient assault/ sexual abuse in a nursing home[c]
Anderson[91]	7,719,510	7,592,961	Maritime accident/ exposure to toxic gas.
Montalvo[92]	13,244,682	6,595,706	Horse ranch agreed to party in which minors were served alcohol; plaintiff beaten/died[a]
Cruz[93]	12,950,431	6,475,216	Civil suit following a murder conviction[a]

[83] Montenegro v. Balmaseda, No. 90-20979 CA 29, 1991 WL 712194 (Fla. Cir. Ct. Apr. 1991).

[84] Dudley v. Owens-Corning Fiberglas Corp., No. 93-9601, 1994 WL 864863 (Fla. Cir. Ct. Apr. 29, 1994).

[85] Ferguson Transp., Inc. v. N. Am. Van Lines, Inc., No. 87-7567 AN, 1992 WL 736930 (Fla. Cir. Ct. June 1992), *verdict reduced*, 639 So. 2d 32 (Fla. Dist. Ct. App. 1994), *aff'd*, 687 So. 2d 821 (Fla. 1996).

[86] Goldberg v. Callahan, No. 92-4018 CI 19, 1996 WL 901963 (Fla. Cir. Ct. Oct. 15, 1996).

[87] Caron v. Goldberg, No. 90-09996 21, 1992 WL 737452 (Fla. Cir. Ct. May 27, 1992).

[88] Wheeland v. Wheeland, No. 93-9072 CA 15, 1993 WL 807885 (Fla. Cir. Ct. Aug. 26, 1993).

[89] Lowell v. Celotex Corp., No. 85-4140, 1990 WL 628490 (Fla. Cir. Ct. Mar. 1990).

[90] Collins v. Beverly Enter.-Fla., Inc., No. 98-433-CA, 2000 WL 1203896 (Fla. Cir. Ct. Mar. 31, 2000).

[91] Andersen v. Sky Cruises Ltd., Inc., No. 87-8892, 1990 WL 630243 (Fla. Cir. Ct. Jan. 1990).

[92] Montalvo v. Rancho El Nuevo Mundo, Inc., No. 96-16035 CA 22, 1998 WL 555151 (Fla. Cir. Ct. Apr. 28, 1998).

[93] Cruz v. Gailfoil, No. CL 90-10950 AE, 1991 WL 712328 (Fla. Cir. Ct. Aug. 1991).

Notes to Table IV: Awards adjusted to 1999 dollars. [a] Defendant apparently not represented by counsel. [b] Punitive portion reduced or overturned on appeal. [c] Case settled for an undisclosed amount.

Another set of cases were subject to post-verdict reviews. On appeal of *Ferguson Transportation, Inc.*, the punitive damage award for tortious breach of contract was denied as a matter of law.[94] In one case the punitive award was reduced on *remittitur* from $15 million to slightly over $5 million,[95] and in another case the dispute was settled following the verdict.[96] In *Scheller*[97] the award for tortious interference was upheld on appeal.[98] The Florida Supreme Court denied review of a District Court of Appeals decision upholding the punitive damage award in *Palank.*[99] Other cases had motions pending following the award or the intermediate appellate court decision.

In summary, the mega awards look quite different when we consider the cases more closely. It seems likely that about half of the twenty awards resulted in either no payment or a reduced payment. In two of the remaining cases, the large punitive awards were upheld by appellate courts, but we have no information as to whether the money was paid or how much was paid. Similarly, in the other cases we have no information about the ultimate outcome. We now consider these mega-awards in the light of other findings.

In recent research we have examined post-trial adjustments of awards in samples of medical malpractice cases.[100] The data set included samples of Florida cases as well as samples from New York and California. One of our central conclusions was that a substantial number of high-end, or outlier, awards—even when they exclusively involve only compensatory damages—are reduced downwards because of apportionment of comparative negligence to the plaintiff, *remittitur* by the judge, or post-verdict negotiations by the parties.

Viscusi reached a similar conclusion in his book on products liability reform.[101] He concluded that in products liability cases involving punitive damages "plaintiffs received only 29% of the original punitive award. Courts often reduce punitive damages on appeal, and defendants may

---

[94] 639 So. 2d 32 (Fla. Dist. Ct. App. 1994), *aff'd* 687 So. 2d. 821 (Fla. 1997).

[95] Dudley v. Owens-Corning Fiberglas Corp., No. 93-9601, 1994 WL 864863 (Fla. Cir. Ct. Apr. 29, 1994).

[96] Collins v. Beverly Enter.-Fla., Inc., No. 98-433-CA, 2000 WL 1203896 (Fla. Cir. Ct. Mar. 31, 2000).

[97] No. 80-5519 CA(L) 01 AF, 1989 WL 527424 (Fla. Cir. Ct. Mar. 1989).

[98] Am. Med. Int'l, Inc. v. Scheller, 590 So. 2d 947 (Fla. Dist. Ct. App. 1991).

[99] CSX Transp., Inc. v. Palank, 743 So. 2d 556 (Fla. Dist. Ct. App. 1999), *review denied*, 760 So. 2d 946 (Fla. 2000).

[100] *See* Neil Vidmar, Felicia Gross & Mary Rose, *Jury Awards for Medical Malpractice and Post-Verdict Adjustment of Those Awards*, 48 DEPAUL L. REV. 265 (1998).

[101] W. KIP VISCUSI, REFORMING PRODUCTS LIABILITY (1991).

negotiate a reduction in this amount in return for prompt payment of the damages amount."[102]

The RAND Corporation's Institute for Civil Justice also studied punitive damage awards in Cook County (Chicago, Illinois) and San Francisco that were awarded between 1979 and 1983.[103] The author of the RAND report concluded: "Jury verdicts are not the last word in cases involving punitive damages. Remittitur, post trial motions, appeals and even settlements may reduce the award amount."[104] Roughly one out of two punitive awards were reversed or reduced in the post-verdict period, with the largest awards having the highest rates of downward adjustment.[105]

Rustad and Koenig also conducted research on the aftermath of punitive damage awards.[106] In a nationwide sample of punitive damages awarded in non-asbestos products liability cases, 25% of the time the award was upheld and 23% of the time the award was reversed or reduced by an appellate court.[107] In a sample of all products liability cases, 36% of the time the case settled after the verdict; in 10% of cases, the award was not collectible; and in 6% of cases, an appeal was still pending.[108] The largest punitive awards were the least likely to survive intact.[109] When awards exceeded $10 million, on average only 11% of the initial punitive damage award was collected.[110]

Thus, while jury awards are occasionally high, these outlier awards must be viewed in the context of the other mechanisms that the legal system has developed to control for excesses. These mechanisms are utilized with regularity and play an important role in bringing litigation to an end, preventing a prolonged appeals process.

IV. Comparison with Other Studies

A. Infrequency of Punitive Awards

The data from Florida present a very different picture of punitive damages than the claims made by proponents of tort reform in Florida.

---

[102] *Id.* at 94.

[103] Mark Peterson et al., RAND Corp., Punitive Damages: Empirical Findings (1987).

[104] *Id.* at 26.

[105] *Id.* at 27–28.

[106] Michael Rustad, *In Defense of Punitive Damages in Products Liability: Testing Tort Anecdotes with Empirical Data*, 78 Iowa L. Rev. 1 (1992); Michael Rustad & Thomas Koenig, *Reconceptualizing Punitive Damages in Medical Malpractice: Targeting Amoral Corporations, Not "Moral Monsters,"* 47 Rutgers L. Rev. 975 (1995).

[107] Rustad, *supra* note 106, at 55 tbl.11.

[108] *Id.* at 57, tbl.14.

[109] *Id.* at 59–60.

[110] *Id.*

They are, however, quite consistent with the findings of other empirical studies.

Two studies of civil trials were conducted during the 1990s by the Bureau of Justice Statistics ("BJS") of the United States Department of Justice in conjunction with the National Center for State Courts.[111] In 1992 and again in 1996, the BJS studied the seventy-five largest state court jurisdictions in the United States.[112] The data included the Florida counties of Dade, Orange, and Palm Beach. In 1992 there were 367 cases in which plaintiffs prevailed in these three counties.[113] Punitive damages were awarded in just seven of the 367 cases.[114] In 1996 the BJS researchers found 420 cases in which plaintiffs prevailed in Dade, Orange, and Palm Beach counties.[115] Once again, however, punitive damages were awarded in just a small fraction of these cases, that is, in just fifteen of the 420 successful lawsuits.[116]

The first BJS Study found that, nationwide, in 1992 successful plaintiffs in tort cases were awarded punitive damages 4% of the time.[117] By contrast, in contract cases prevailing plaintiffs received punitive awards 12.2% of the time.[118] The figures for 1996 were comparable with punitive damages awarded in, 4% of successful tort cases and 12.7% of successful contract cases.[119] The BJS studies did not distinguish between tort and contract cases when they reported punitive damages by county. However, the combined percentages of punitive awards to plaintiffs who won their jury trials in the three Florida counties in 1992 and 1996, as described above, were 2% and 3.6% respectively.[120] In short, in both 1992 and 1996 Florida fell below the national average in percentages of cases in which punitive damages were awarded to plaintiff winners.

The second BJS Study indicated that in 1996 the median punitive damage award for all jury trial cases nationwide was $50,000.[121] The comparison figure for Dade County, which had thirteen of the fifteen pu-

---

[111] CAROL J. DEFRANCES ET AL., U.S. DEPT. OF JUSTICE, CIVIL JURY CASES AND VERDICTS IN LARGE COUNTIES (Bureau of Justice Statistics Bulletin 1995) [hereinafter 1995 BJS Study]; CAROL J. DEFRANCES & MARIKA F. X. LITRAS, U.S. DEPT. OF JUSTICE, CIVIL TRIAL CASES AND VERDICTS IN LARGE COUNTIES, 1996 (Bureau of Justice Statistics Bulletin 1999) [hereinafter 1999 BJS Study].

[112] 1995 BJS Study, *supra* note 111, at 701; 1999 BJS Study, *supra* note 111, at 1.

[113] 1995 BJS Study, *supra* note 111, at 713 app. tbl.2.

[114] *Id.*

[115] 1999 BJS Study, *supra* note 111, at 21 app. C.

[116] *Id.* at 22 app. D. In 1996 there were a total of 649 civil jury trials of all types in the three counties (including cases in which civil plaintiffs lost). Thus, using the total of all cases tried to juries in 1996 as the denominator and the 15 punitive awards as the numerator, we may conclude that punitive damages were awarded in 2.3% of all civil lawsuits tried by juries.

[117] 1995 BJS Study, *supra* note 111, at 708 tbl.8.

[118] *Id.*

[119] 1999 BJS Study, *supra* note 111, at 16.

[120] *See supra* notes 113–116 and accompanying text.

[121] 1999 BJS Study, *supra* note 111, at 10.

nitive damage award cases in 1996, was $57,000.[122] Dade County appears to have been within close range of the national median. Palm Beach, however, had a much higher median award—$225,000[123]—but this is based on only two cases.

Eisenberg and his colleagues analyzed the data from the 1992 BJS Study as well as twenty-five years of awards from Cook County, Illinois and California.[124] They applied sophisticated statistical analyses to these data sets and arrived at several conclusions. There was no evidence that punitive damage awards were more likely when individuals sued businesses than when individuals sued individuals.[125] Juries rarely awarded punitive damages and appeared to be "especially reluctant" to do so in medical malpractice and products liability cases.[126] Punitive damages were most likely to result in business/contract cases and intentional tort cases.[127] The authors went so far as to state: "Unless the case involves an intentional tort or a business related tort (such as employment claims) punitive damages will almost never be awarded."[128]

The most recent report from the RAND Institute for Civil Justice examines fifteen state courts of general jurisdiction in a period from 1985 to 1994, fourteen of which allow for punitive damages under state law.[129] Researchers in that organization observed: "Perhaps the most striking finding that emerges from the jury verdict data in this study is that punitive damages are awarded very rarely."[130] The report also concluded: "The discussion about punitive damages focuses primarily on products liability, but in jurisdictions we examined, most punitive damages were awarded in intentional tort and business cases .... In contrast, products liability was the underlying cause of action in only 4.4 percent of the punitive damage awards made."[131]

A recent article arising out of RAND data by Moller identified and examined a large sample of cases involving what were identified as financial injury cases involving punitive damages.[132] He concluded that fifty percent of all punitive damage awards are made in cases in which

---

[122] *Id.* at 22.

[123] *Id.*

[124] *See* Theodore Eisenberg et al., *The Predictability of Punitive Damages*, 26 J. LEGAL STUD. 623 (1997).

[125] *Id.*

[126] *Id.* at 623, 634–37.

[127] *Id.* at 634–37.

[128] *Id.* at 659. Daniels and Martin also analyzed punitive damages from a large sample of cases from jurisdictions around the nation for the years 1988 through 1990 and reached a similar conclusion. Daniels & Martin, *supra* note 2, at 217.

[129] ERIK MOLLER, RAND CORP., TRENDS IN CIVIL JURY VERDICTS SINCE 1985, at 33 (1996) (stating that Washington state does not permit punitive damages).

[130] *Id.* at 33.

[131] *Id.* at 34.

[132] Erik Moller et al., *Punitive Damages In Financial Injury Jury Verdicts*, 28 J. LEGAL STUD. 283, 327–31 (1999).

the plaintiff alleges a financial injury.[133] Moller specifically examined a set of data from Alabama because of the widespread belief that Alabama juries are excessively generous in awarding damages.[134] Contrary to this belief, however, the data indicated that Alabama juries generally award smaller punitive damage awards than other jurisdictions.[135] Alabama juries, however, did render punitive awards that were larger relative to the compensatory awards than other jurisdictions.[136] There is no clear explanation for this difference, which could be due to peculiarities of Alabama law on punitive damages or the tendencies of Alabama jurors or some other reason.[137]

Finally, Eisenberg and his colleagues compared judge versus jury trial outcomes in 1996 in forty-five of the nation's largest counties.[138] After controlling for differences in the types of cases heard by each set of decision-makers, the authors concluded that there was no substantial evidence that judges and juries differed in the rate at which they awarded punitive damages nor in the basic ratio of punitive to compensatory damages. Jury trials did have a greater range of punitive damages for a given level of compensatory damages but in the end, the authors concluded, there were only a "trivially" few cases in which the jury award would have exceeded what a judge might have awarded.

B. Pretrial Effects of Punitive Damages on Settlements

Despite the fact that our data and that reviewed from other studies strongly contradict the claim that there is a tort crisis in Florida with regard to products liability cases, or a crisis regarding awards of punitive damages in any other type of tort claim, proponents of tort reform argue that claims for punitive damages are as bad as actual awards of punitive damages.[139] Such claims can be summarized as follows: because insurance is not available for punitive awards, fears of irrational jury awards drive corporate defendants to settle claims instead of going to trial and to settle for amounts larger than they would otherwise offer. While our data from the Florida verdict reporter do not allow us to assess this issue directly, other researchers have addressed it. Koenig conducted an extensive review of pre-verdict settlements and found no evidence of the alleged phenomenon.[140]

---

[133] *Id.* at 332.
[134] *Id.* at 327.
[135] *Id.* at 330.
[136] *Id.* at 331.
[137] *Id.* at 332.
[138] Theodore Eisenberg et al., Juries, Judges, and Punitive Damages: An Empirical Study (Unpublished Draft of Nov. 1, 2000) (copy on file with authors).
[139] *See, e.g.*, Meros & Hundley, *supra* note 9.
[140] Thomas Koenig, *The Shadow Effect of Punitive Damages on Settlements*, 1998 WISC. L. REV. 169, 208–09.

Koenig's review of pre-verdict settlement practices in other jurisdictions yields a similar conclusion. Claims adjusters settle cases around predicted compensatory damages rather than around punitive damages.[141] This conclusion does not necessarily contradict the notion that claims for punitive damages affect the litigation process in complex ways that are not well understood.[142] The hypothesis that punitive damage awards (or the prospect of such awards) may have some effect on the settlement process, including an inflationary effect, is not inherently implausible. Nevertheless, the important point is that despite frequent claims by tort reform proponents in Florida, and around the country, that punitive damages claims and punitive damages awards produce an *in terrorem* effect on corporate defendants, there is no systematically documented evidence that this is so. Indeed, there is evidence that such predicted effects are minimal or non-existent. In particular, inasmuch as the Florida jury verdict data yielded almost no products liability cases other than those involving asbestos in which punitive damages were awarded, it seems extremely improbable that punitive damages or the fears of punitive damages would have a significant effect on settlement rates in products liability cases other than possibly asbestos and cigarette cases. If corporate defendants are, as was claimed in the legislative history of Florida House Bill 775, "scared to death" of punitive damages awards because they believe that punitive damages awards are "skyrocketing," "exploding," and otherwise "running wild,"[143] there is no empirical evidence to support their fear.[144]

V. CONCLUSION

Our study joins a substantial literature showing that punitive damage awards by juries are infrequent. In addition, the present study goes beyond existing work to provide a highly detailed portrait of the types of cases in which punitive damages are awarded. Both through our account of the categories of cases involving punitive damage awards, as well as the closer inspection of large awards, this study provides a greater context for understanding the circumstances in which juries are asked to consider punitive damage awards.

---

[141] *Id.*

[142] It is noteworthy, for instance, that studies of claims adjusters, *e.g.*, *id.*, may fail to find attention to punitive damages in settlement decisions because insurance companies are not generally financially responsible for these awards.

[143] Meros & Hundley, *supra* note 9, at 477–78.

[144] Of note, the BJS data do show a slight increase in punitive awards between 1992 and 1996, from 2% to 3.5%, in the three largest Florida counties. *See supra* note 120. It is possible that an increase in large counties helped to create an impression that punitive damages were on the rise in Florida as a whole; however, our data dispute the notion that this increase was a statewide phenomenon.

This systematic examination of actual jury awards of punitive damages in Florida gives no indication of a crisis regarding punitive damages, either in their frequency or in their amounts as assessed by the ratio of punitive damages to compensatory damages. Drunken or reckless driving cases, assaults, and financial fraud constituted approximately fifty-seven percent of all cases. Individuals, not businesses, were quite often the defendants in these cases.

In particular, there is no support for the claim that punitive damages are frequent in product liability cases. Except for asbestos cases, punitive damages simply were not given in product liability cases. Similarly, punitive damages were rare in premises liability and *respondeat superior* cases. The factual conditions involved in the few premises liability cases present claims of very serious misconduct, including criminal actions.

There were some mega awards in the data, albeit many fewer than would be expected from the rhetoric involved in the tort reform debate. But closer examination of these cases also revealed that it was highly unlikely that the plaintiff would ever collect any punitive damages; in fact, in some cases even compensatory damages were unlikely to be collected. Additional research uncovered the fact that in two of the mega award cases appellate courts concluded that the amounts awarded by the jury were appropriate under the law. These findings are consistent with our previous research indicating that apparently excessive awards, even though infrequent, need to be viewed in the context of remedial mechanisms and processes of the legal system, such as *remittitur*, appellate review, and settlement.[145]

Finally, we turn to the implication of these data for the claims of *in terrorem* effects of punitive damages. The data indicate that in Florida there is no basis for fears of punitive damages from legitimate business practices. Of course, *in terrorem* effects involve subjective perceptions, and the argument has been made that fear is more important than actual threat. Combined with experimental research that has focused on jury variability in setting punitive awards,[146] the argument is that juries are to be feared and, therefore, reforms are necessary. There are some striking ironies here. First, proponents of tort reform accuse jurors of awarding extreme remedies based on an exaggerated sense of risk and a focus on atypical events.[147] The Florida data suggest that tort reform advocates themselves are engaging in the same type of exaggeration by asking for drastic change without clear evidence of a system-wide problem. Second,

---

[145] Vidmar, Gross & Rose, *supra* note 100, at 298.

[146] *See* David Schkade, Cass Sunstein & Daniel Kahneman, *Deliberating About Dollars: The Severity Shift*, 100 COLUM. L. REV. 1139 (1996); Sunstein, Kahneman & Schkade, *supra* note 6.

[147] *See* Hastie & Viscusi, *supra* note 6; Reid Hastie, David Schkade & John Payne, *Juror Judgments in Civil Cases: Hindsight Effects on Judgments of Liability for Punitive Damages*, 23 LAW & HUM. BEHAV. 597, 612 (1999).

courts have sometimes looked askance at social scientists who argue for changes in the legal system on the basis of results of experimental studies that tend to ignore real-world dynamics and contexts.[148] Here we, both social scientists, suggest that the legal literature should consider the recent spate of experimental studies of punitive damages[149] in light of real-world dynamics and contexts before jumping to policy generalizations.[150]

Being properly cautious, we limit our conclusions to the state of Florida, although, as we have noted, research involving other states seems consistent with the Florida findings. A decade ago, Sanders and Joyce wrote an insightful article about the tort reforms that were enacted in many states in the 1980s following the insurance crisis that caused concern about rates and even the availability of insurance.[151] Those authors observed that in Texas and elsewhere the problems perceived and cited by legislatures, specifically growth in size and uncertainties within the tort law system, were not well understood and that the reforms that were passed did not address the perceived problems. To a very substantial extent, attention to anecdote and unsupported assertions characterized reform legislation passed in Illinois and subsequently overturned by the Illinois Supreme Court.[152] When contrasted with data about actual punitive damage awards, the portion of Florida's House Bill 775 dealing with punitive damages appears to fall within this same genre of legislative decision-making. We do not take a position on whether tort reform is needed in Florida or elsewhere—there may be problems. However, our research indicates rather convincingly that, in Florida at least, current justifications for reforming punitive damages lack empirical support.

---

[148] *See, e.g.*, Lockhart v. McCree, 476 U.S. 162, 169 (1986) (rejecting lower court's reliance on some 15 empirical studies as sufficient evidence of juror-bias in death penalty cases).

[149] *See* Hastie & Viscusi, *supra* note 6; Hastie, Schkade & Payne, *supra* note 147; Schkade, Sunstein & Kahneman, *supra* note 149.

[150] This is a particularly important consideration given that funding for many of these studies comes from parties highly interested in reforms, *see* Elizabeth Amon, *Exxon Bankrolls Critics of Punitives: Then It Cites the Research in Appeal of $5.3 B Valdez Award*, NAT'L L.J., May 17, 1999, at A1, and that some of the studies have been heavily criticized as containing major conceptual and methodological flaws, *see* Neil Vidmar, *Juries Don't Make Legal Decisions! And Other Problems: A Critique of Hastie et al. on Punitive Damages*, 23 LAW & HUM. BEHAV. 705 (1999); Richard Lempert, *Juries, Hindsight, and Punitive Damage Awards: Failures of a Social Science Case for Change*, 48 DEPAUL L. REV. 867 (1999); Robert J. MacCoun, *Epistemological Dilemmas in the Assessment of Legal Decision Making*, 23 LAW & HUM. BEHAV. 723 (1999).

[151] Joseph Sanders & Craig Joyce, *"Off to the Races": The 1980s Tort Crisis and the Law Reform Process*, 27 HOUST. L. REV. 207 (1990).

[152] Best v. Taylor Machine Works, 689 N.E.2d 1057 (Ill. 1997).

RECENT DEVELOPMENTS

AIDS CRISIS

Over 36 million people in the world live with HIV/AIDS[1]—a number greater than the population of California.[2] While the number of AIDS-related deaths in the United States has been falling since 1996,[3] the AIDS crisis continues to escalate in the developing world.[4] Over two-thirds of all people infected with HIV/AIDS live in sub-Saharan Africa, where more than 8.8% of adults are infected[5] and a lack of access to health care and information about the disease hinders treatment and prevention.[6] While the success of antiretroviral drugs may have reduced the sense of urgency associated with the disease in the United States among lawmakers and the media,[7] most of the world's HIV/AIDS victims cannot afford treatment that could substantially increase their life spans.[8] Despite recent signs that the number of new infections in the region may be

---

[1] JOINT UNITED NATIONS PROGRAMME ON HIV/AIDS, AIDS EPIDEMIC UPDATE 3 (2000) [hereinafter UNAIDS], *available at* http://www.unaids.org/wac/2000/wad00/files/WAD_epidemic_report.htm.

[2] U.S. DEP'T OF COMMERCE, STATISTICAL ABSTRACT OF THE UNITED STATES 28 (119th ed. 1999).

[3] *Researchers Sound Alarm About HIV Complacency*, AIDS POL'Y & L., Sept. 17, 1999, at 1 [hereinafter *Researchers Sound Alarm*]. The number of Americans estimated to be living with AIDS has continued to increase annually, but AIDS-related deaths began to decline in 1996 when antiretroviral treatments first became available. *Id.*; CENTERS FOR DISEASE CONTROL AND PREVENTION, HIV/AIDS SURVEILLANCE REPORT 35 tbl.26 (Mid-year ed. 2000), *available at* http://www.cdc.gov/hiv/stats/hasr1201.pdf. The percentage declines have been smaller in subsequent years. *See Researchers Sound Alarm*, *supra*, at 1.

[4] *See* UNAIDS, *supra* note 1, at 4.

[5] *See* UNAIDS, *supra* note 1, at 5.

[6] *See* H.R. REP. No. 106-548, at 6 (2000) (citing poorly developed health care infrastructure as a factor contributing to the spread of AIDS in Africa). AIDS drugs are far beyond the means of most of the disease's developing-world victims, many of whom do not have access to basic health care. *See* David Finkel, *Few Drugs for the Needy*, WASH. POST, Nov. 1, 2000, at A1. In addition, limited communication channels make preventive efforts more difficult. Information about how the virus is transmitted has spread slowly to isolated rural communities. *See* Karl Vick, *Disease Spread Faster Than the Word*, WASH. POST, July 7, 2000, at A1.

[7] *See generally* Barton Gellman, *The Global Response to AIDS in Africa*, WASH. POST, July 5, 2000, at A1 (describing initial failure of United States and international policymakers to react to the African AIDS crisis); Hawlan Ng, *AIDS in the Media*, HARV. AIDS REV., Fall 1999/Winter 2000, at 18, 18–19 (observing a cooling of media interest in AIDS following the introduction of antiretroviral drugs).

[8] In the United States, antiretroviral treatment costs between $10,000 and $15,000, and the average per capita GNP in sub-Saharan Africa was only $509 in 1998. Sheryl Gay Stolberg, *Africa's AIDS War*, N.Y. TIMES, Mar. 10, 2001, at A1; WORLD BANK, AFRICAN DEVELOPMENT INDICATORS 2001, at 33 (2001). Several pharmaceutical companies have announced that they will sell AIDS drugs at sharply reduced prices in Africa. *See* Melody Petersen and Donald G. McNeil, Jr., *Maker Yielding Patent in Africa for AIDS Drugs*, N.Y. TIMES, Mar. 15, 2000, at A1. Even at reduced prices, treatment remains unaffordable for many AIDS victims. Rachel Swarns, *AIDS Obstacles Overwhelm a Small South African Town*, N.Y. TIMES, Mar. 29, 2001, at A1.

stabilizing,[9] the human suffering caused by HIV/AIDS and its impact on African economies and societies are expected to continue well into the twenty-first century.[10]

Although the AIDS virus has been present in sub-Saharan Africa for over twenty years,[11] the rest of the world only recently began to focus attention on the region's epidemic.[12] The increasing interest was heralded by two international gatherings: the United Nations Security Council held a special session to discuss AIDS in January 2000[13] and in July 2000 the world AIDS conference in Durban, South Africa called more attention to the urgency of the epidemic in Africa.[14]

The United States's wealth, influence, and technological expertise place it in a unique position to respond to the AIDS crisis.[15] Furthermore, it is in the self-interest of the United States to combat the spread of AIDS: a report released by the National Intelligence Council concludes that the prevalence of AIDS and other infectious diseases worldwide poses a major national security threat.[16]

The United States has begun to direct greater economic resources to the task of fighting AIDS worldwide. AIDS-specific funding to the United States Agency for International Development increased from $140 million in fiscal year 1999 to $330 million in fiscal year 2001.[17] The United States also has leveraged its resources to support international cooperation. The Clinton administration launched the government-wide

---

[9] UNAIDS, *supra* note 1, at 4.

[10] South African economic growth is expected to fall .3% to .4% annually due to AIDS. Its GDP in 2010 is expected to be 17% lower than what it would have been without the disease. UNAIDS, *supra* note 1, at 4; *see also* Peter Wehrwen, *The Economic Impact of AIDS in Africa*, HARV. AIDS REV., Fall 1999/Winter 2000, at 12.

[11] UNAIDS, *supra* note 1, at 5 (stating that the AIDS epidemic started in sub-Saharan Africa in the late 1970s or early 1980s).

[12] *See A Turning-Point for AIDS?*, ECONOMIST, July 15, 2000, at 77, 77 [hereinafter *A Turning Point*]; Barton Gellman, *World Shunned Signs of the Coming Plague*, WASH. POST, July 5, 2000, at A1.

[13] U.S. Ambassador to the United Nations Richard Holbrooke spearheaded the session. Gellman, *supra* note 12. Vice President Gore delivered remarks calling for "the world's wealthier, healthier nations to match America's increasing commitment to a worldwide crusade against AIDS." Press Release, Office of the Vice President, Remarks as Prepared for Delivery by Vice President Al Gore to the U.N. Security Council (Jan. 10, 2000), *available at* http://clinton4.nara.gov/ONAP/pub/vp_un_sc.html.

[14] *A Turning Point*, *supra* note 12, at 77.

[15] *See* WHITE HOUSE OFFICE OF NAT'L AIDS POLICY, ACTION AGAINST AIDS: A LEGACY OF LEADERSHIP AT HOME AND AROUND THE WORLD 40 (2000) [hereinafter WHITE HOUSE].

[16] NAT'L INTELLIGENCE COUNCIL, THE GLOBAL INFECTIOUS DISEASE THREAT AND ITS IMPLICATIONS FOR THE UNITED STATES, ENVTL. CHANGE AND SECURITY PROJECT REP., Summer 2000, at 33, 34. The report details how several diseases including HIV/AIDS are expected to impact U.S. national security. *Id.* at 37. It states that the social and economic impact of AIDS will foster political instability abroad, and that the disease's spread through foreign militaries and peacekeeping forces will negatively impact U.S. security interests. *Id.*

[17] WHITE HOUSE, *supra* note 15, at 30.

Leadership and Investment in Fighting an Epidemic ("LIFE") initiative in July 1999 to increase funding and demonstrate American leadership in the fight against AIDS.[18] The United States has provided more funding to the Joint United Nations Programme on HIV/AIDS ("UNAIDS") than any other country.[19]

Members of the 106th Congress introduced several bills aimed at fighting the global AIDS epidemic.[20] House Bill 3519, sponsored by Representative James Leach (R-Iowa), incorporated several of these initiatives.[21] The House considered the bill under suspension of the rules and passed it on May 15, 2000.[22] The Senate version passed by unanimous consent on July 26, 2000,[23] and the House agreed to adopt that version without objection on July 27.[24] President Clinton signed the bill into law on August 19, 2000 as the Global AIDS and Tuberculosis Relief Act.[25] The Act authorizes a total of $510 million in United States contributions in 2001 and an equal amount in 2002 toward programs designed to prevent AIDS and treat individuals who are already infected in the developing world.[26] Of those contributions, $150 million is authorized for a new World Bank AIDS Trust Fund ("the Fund"),[27] which would accept contributions from public and private donors and provide grants to countries with high AIDS prevalence rates and those likely to develop high prevalence rates.[28] The legislation also authorizes $300 million in foreign assistance for combating the spread of AIDS.[29] The remaining $60 million in contributions is targeted to the Global Alliance for Vaccines and Immunizations ($50 million) and the International AIDS Vaccine Initiative ($10 million)[30] with the goals of increasing the availability of various vaccinations worldwide and promoting the development and manufacture of new vaccines.[31]

---

[18] *Id.* at 30.

[19] *Id.* at 28.

[20] *See, e.g.*, H.R. 2765, 106th Cong. (1999); S. 2026, 106th Cong. (2000); H.R. 4038, 106th Cong. (2000); S. 2032, 106th Cong. (2000); S. 2132, 106th Cong. (2000); H.R. 3812, 106th Cong. (2000); H.R. 5219, 106th Cong. (2000). Representatives of the 107th Congress have also introduced legislation on the global AIDS epidemic. *See, e.g.*, H.R. 502, 107th Cong. (2001).

[21] 146 CONG. REC. H47 (daily ed. Jan. 27, 2000).

[22] 146 CONG. REC. H3016-25 (daily ed. May 15, 2000).

[23] 146 CONG. REC. S7624 (daily ed. July 26, 2000).

[24] 146 CONG. REC. H7177-81 (daily ed. July 27, 2000).

[25] Global AIDS and Tuberculosis Relief Act of 2000, Pub. L. No. 106-264, 114 Stat. 748.

[26] *Id.* §§ 111(a), 112(a), 141(a).

[27] *Id.* § 141(a).

[28] *See id.* §§ 121(a), 122(a), 123(b), 141(a).

[29] *Id.* § 111(a).

[30] *Id.* § 112(a).

[31] *Id.* § 112(b). The legislation additionally authorizes the federal government to implement a comprehensive tuberculosis control program with the help of the World Health Organization. *Id.* § 203. The logic of including AIDS and tuberculosis initiatives in the same bill arises from the fact that the two epidemics are linked. The AIDS epidemic is

House Bill 3519 is noteworthy for its broad scope and for the bipartisan support it received.[32] However, the legislation leaves much work to be done. The funding authorized by the bill covers a mere fraction of the enormous cost of AIDS prevention initiatives and basic care in the developing world, estimated at $3 billion in Africa alone.[33] Furthermore, the 106th Congress failed to appropriate a large percentage of the funds authorized by the bill; specifically, the World Bank Trust Fund received $20 million from the United States for fiscal year 2001, rather than the $150 million authorized by House Bill 3519.[34] Additionally, the legislation only briefly mentions the importance of making drugs available to developing countries at affordable prices and lacks concrete proposals to address this crucial problem.[35] Recent drug pricing developments both increase the potential impact of U.S. funding for drug purchases and highlight the need for greater investment in basic health infrastructure in order to provide such treatment in an effective manner.[36]

The creation of the Fund serves as one of the centerpieces of House Bill 3519. The House Committee on Banking and Financial Services described the Fund as a Marshall Plan-like device to help developing countries, and African nations in particular, recover from the economic and

---

frustrating efforts to halt the spread of tuberculosis, since people with AIDS are especially susceptible to the pulmonary disease. *See* Editorial, *New Weapons Against Tuberculosis*, N.Y. TIMES, Mar. 17, 2001, at A10. Senator Barbara Boxer (D-Cal.) added tuberculosis provisions to House Bill 3519 after first introducing them in 1999 as the International Tuberculosis Control Act, S. 1497, 106th Cong. (1999). *See* 146 CONG. REC. S7624 (daily ed. July 26, 2000) (statement of Sen. Boxer).

[32] Although House Bill 3519 benefited from the support of key members of both parties, the Clinton administration did not give the Fund its full endorsement at the Banking Committee hearings. *Global AIDS Crisis and Pandemic in Africa: Hearing on H.R. 3519 Before the House Comm. on Banking and Fin. Servs.*, 106th Cong. 52–53 (2000) [hereinafter *Hearing*]. Citing the need to secure "broad support and ownership internationally," Treasury Undersecretary for International Affairs Timothy Geithner stated that the administration wanted to "stop just one step short of being specifically committed to the vehicle." *Id.* Furthermore, the broader question of how best to address the global AIDS crisis is not immune from partisan debate. *See AIDS: Epidemic Without Borders*, L.A. TIMES, July 9, 2000, at M4 (noting Sen. Trent Lott's (R-Miss.) disagreement with the Clinton administration's characterization of AIDS as a threat to national security).

[33] WHITE HOUSE, *supra* note 15, at 40.

[34] Global AIDS and Tuberculosis Relief Act of 2000 § 141(a); Foreign Operations, Export Financing, and Related Programs Appropriations Act of 2001, Pub. L. 106-429, tit. 2, 114 Stat. 1900, 1900A-5 (2000). Before the federal government can spend money, Congress must pass and the President must sign both an authorization and an appropriation. The authorization sets an upper limit on the amount that can be allocated, while the appropriation determines the *actual* amount allocated. The Global AIDS and Tuberculosis Relief Act of 2000 served as authorizing legislation, while the Foreign Operations, Export Financing, and Related Programs Appropriations Act of 2001 is appropriating legislation. *See* STANLEY E. COLLENDER, THE GUIDE TO THE FEDERAL BUDGET: FISCAL 1998, at 1 (1997).

[35] Global AIDS and Tuberculosis Relief Act of 2000 §§ 111(a), 122(a).

[36] *See* Barry R. Bloom, *AIDS: The Drugs Won't Be Enough*, WASH. POST, Mar. 9, 2001, at A27.

social consequences of the AIDS epidemic.[37] The Fund is designed to help countries prevent the spread of AIDS and provide care for those who are already infected.[38] It also seeks to provide care and education for children who have been orphaned by AIDS.[39]

House Bill 3519 directs the Secretary of the Treasury to negotiate the establishment of a Board of Trustees to govern the Fund, consisting of representatives from donor countries.[40] The United States representative would be appointed by the President and confirmed by the Senate.[41] The legislation also envisions an Advisory Board consisting of "individuals with experience and leadership in the fields of development, health care (especially HIV/AIDS), epidemiology, medicine, biomedical research, and social sciences," as well as representatives from the United Nations and nongovernmental organizations.[42]

The legislation envisions that the Fund and its advisory board will be established within the World Bank (the "Bank").[43] Several factors make the Bank a logical trustee. The Bank has been engaged in health financing since 1980[44] and started funding HIV/AIDS programs six years later.[45] It has gained experience financing AIDS-related initiatives through its co-sponsorship of UNAIDS.[46] Bank involvement in addressing the global AIDS problem is particularly appropriate given the disease's devastating effect on the economies of developing countries.[47] Because developing countries are accustomed to receiving economic direction from the Bank, the Bank's administration of the fund may impress upon governments the vital connection between their economic future and their ability to address health issues.[48]

---

[37] The version of House Bill 3519 reported by the House Committee on Banking and Financial Services names the Fund the World Bank AIDS Marshall Plan Trust Fund. H.R. REP. NO. 106-548, at 1 (2000). Representative Barbara Lee (D-Cal.) introduced House Bill 2765, the AIDS Marshall Plan Fund for Africa Act, on August 5, 1999. 145 CONG. REC. E1772–73 (daily ed. Aug. 5, 1999); H.R. 2765, 106th Cong. (1999).

[38] Global AIDS and Tuberculosis Relief Act of 2000 § 121(b)(1).

[39] *Id.* § 121(b)(2).

[40] *Id.* § 121(a). President Clinton's statement upon signing the bill notes that these provisions may raise constitutional concerns by directing executive branch policy on international affairs, and adds that he will "treat them as precatory." Statement on Signing of the Global AIDS and Tuberculosis Relief Act of 2000, 36 WEEKLY COMP. PRES. DOC. 1906 (Aug. 19, 2000).

[41] *Id.* § 121(c).

[42] *Id.* § 124.

[43] *Id.* § 121(a).

[44] Joseph Brunet-Jailly, *Has the World Bank a Strategy on Health?*, 51 INT'L SOC. SCI. J. 347, 347 (1999).

[45] Over $950 million in Bank funding has been committed to HIV/AIDS-related projects. AFRICA REGION, WORLD BANK, INTENSIFYING ACTION AGAINST HIV/AIDS IN AFRICA: RESPONDING TO A DEVELOPMENT CRISIS 35 (2000).

[46] *See* H.R. REP. NO. 106-548, at 8 (2000).

[47] *See* Wehrwen, *supra* note 10, at 12; UNAIDS, *supra* note 1, at 4–5.

[48] Finance ministries do not often consider health issues as falling within their purview. H.R. REP. NO. 106-548, at 9 (2000) (quoting testimony by Treasury Undersecretary Timothy Geithner).

However, Bank administration of the Fund opens the Fund to attack by individuals who are opponents of the Bank itself rather than critics of its administration of the Fund *per se*. In the past, members of Congress have criticized the Bank for a lack of transparency in its operations and for making loans that were later misappropriated by corrupt government officials.[49] Citing these abuses, congressional critics have called for decreased funding of the Bank and other international financial institutions.[50] Concerns about misappropriation led the Senate Foreign Relations Committee to cut back the House authorizations for the Fund from five years to two in order to give Congress an opportunity to evaluate the transparency of the Fund and the uses to which the aid is put.[51] Placing the Fund under the Bank's trusteeship may jeopardize future United States contributions if Congress remains dissatisfied with bank reform efforts.[52]

House Bill 3519 seeks to address these concerns by requiring that the Bank develop criteria for selecting programs to be funded that are acceptable to the Secretary of the Treasury.[53] House Bill 3519 requires the Secretary of the Treasury to submit annual reports on the Fund's activities[54] and mandates a General Accounting Office report on the Fund's effectiveness within two years of enactment.[55] It also requires the Fund's Board of Trustees to "ensure full and prompt public disclosure of the proposed objectives, financial organization, and operations of the Trust Fund."[56] The trust fund structure is intended to isolate grant-making decisions from other bank activity, and the bill specifically prohibits the Fund from making grants for project development associated with Bank loans.[57] However, Bank policy would influence the Fund to the extent that "the trust fund policy objectives should . . . be in harmony with interna-

---

[49] *See, e.g.*, H.R. REP. NO. 106-548, at 17–18 (2000) (dissenting view of Rep. Ron Paul (R-Tex.) and Rep. Bob Barr (R-Ga.) arguing that governments misuse Bank aid and that Bank aid is ineffective, and opposing additional Bank funding to establish an AIDS Trust Fund). *But see Hearing*, *supra* note 32, at 34–35 (statement of Rep. Leach (R-Iowa), Chairman, House Comm. on Banking and Fin. Servs.) (urging that discussion of the Fund be separated from the debate on international financial institutions reform). Banking Committee hearings on the bill took place during a period of intense discussion about reforming international financial institutions, including the World Bank. For example, in March 2000 a congressional commission headed by Professor Allan H. Melzer of Carnegie Mellon University issued a highly politicized report urging an overhaul of the World Bank and International Monetary Fund. Joseph Kahn, *Report Seeks Big Changes in I.M.F. and World Bank*, N.Y. TIMES, Mar. 8, 2000, at C4.

[50] *See supra* note 49 and accompanying text.

[51] 146 CONG. REC. S7619 (daily ed. July 26, 2000) (statement of Sen. Jesse Helms (R-N.C.)).

[52] *Id.* S7619-20.

[53] Global AIDS and Tuberculosis Relief Act of 2000, Pub. L. No. 106-264, § 123(c), 114 Stat. 748, 756.

[54] Global AIDS and Tuberculosis Relief Act of 2000 § 131(a).

[55] *Id.* § 131(b).

[56] *Id.* § 123(e).

[57] *Id.* § 122(a)(3)(d).

tional development priorities and the Bank's country assistance strategies."[58] If the Fund applies definitions of poverty that the Bank has previously applied in conjunction with its lending programs, certain countries may be ineligible for assistance.[59] Furthermore, prior Bank practice in financing health care initiatives raises the possibility that loan proceeds may be channeled through nations' existing—and often inadequate—health care infrastructure, thus preventing aid from reaching many of those infected.[60]

The bill gives the Fund broad leeway in determining grant recipients and selecting projects to support. The Fund may make grants to both governments and nongovernmental organizations, with priority given to countries "that have the highest HIV/AIDS prevalence rate or are at risk of having a high HIV/AIDS prevalence rate."[61] Supported activities include prevention programs; initiatives ensuring uncontaminated blood supplies; testing and counseling; programs to prevent vertical transmission; programs to support and provide education for orphans; initiatives to deter sexual assault and programs to provide treatment to sexual assault victims; and "incentives to promote affordable access to treatments against AIDS and related infections."[62] The Fund may also provide technical assistance by subsidizing the research and development and implementation of treatment and care services, including providing access to affordable drugs.[63]

Strong United States financial support for the Fund would give it a fighting start and encourage future donations. But again, while House Bill 3519 authorizes $150 million for each of the fiscal years 2001 and 2002,[64] Congress appropriated only $20 million to the Fund for fiscal year 2001.[65] This unwillingness to appropriate such a large percentage of the authorized funds may reflect conflict between those who believe that multilateral institutions such as the World Bank provide a valuable op-

---

[58] H.R. REP. No. 106-548, at 8–9 (2000).

[59] *See, e.g., Hearing, supra* note 32, at 91 (testimony of Dr. James M. Sherry, Director, UNAIDS Programme Dev. and Coordination) (noting that some countries that are badly in need of assistance in treating and preventing the transmission of AIDS are ineligible for Bank loans because the Bank does not consider the countries to be impoverished, and arguing that these countries should not be prohibited from receiving Fund grants).

[60] *See* Brunet-Jailly, *supra* note 44, at 352–53 (criticizing the World Bank's health care financing on the grounds that those with access to the existing health care infrastructure are more likely to receive the benefits of the loans, confining aid to a limited portion of the urban population).

[61] Global AIDS and Tuberculosis Relief Act of 2000 § 122(a)(3). The goal of this provision is to target the grants to sub-Saharan Africa. Spencer Rich, *Panel Targets Aid to African and Other HIV-Afflicted Nations*, NAT'L J. NEWS SERV. (Mar. 15, 2000), *at* http://web.lexis-nexis.com/congcomp.

[62] Global AIDS and Tuberculosis Relief Act of 2000 § 122.

[63] *Id.* § 122(a)(1).

[64] *Id.* § 141(a).

[65] Foreign Operations, Export Financing, and Related Programs Appropriations Act of 2001, Pub. L. 106-429, tit. 2, 114 Stat. 1900, 1900A-5 (2000).

portunity to leverage American contributions[66] and those who believe that United States agencies could make better use of the funds than the Bank.[67]

In addition to amounts authorized to support the Fund, House Bill 3519 also authorizes funding for multilateral assistance to countries impacted by AIDS. It amends section 104(c) of the Foreign Assistance Act of 1961[68] to declare that control of the AIDS epidemic is a goal of American foreign aid, placing particular emphasis on vertical transmission of the disease from mother to child.[69] The legislation authorizes $300 million to fund AIDS prevention programs in coordination with international organizations, national and local governments, and other organizations.[70] It requires that any agency administering foreign assistance "undertake a comprehensive, coordinated effort to combat HIV and AIDS" by fostering programs that provide prevention, education, counseling, testing, medication to prevent vertical transmission, and care for people with HIV/AIDS.[71]

House Bill 3519 states that at least 8.3% of the $300 million authorized for foreign assistance should be made available to implement vertical transmission prevention programs.[72] Recent medical breakthroughs make it possible to prevent mother-to-child transmission of AIDS cheaply and effectively.[73] Vertical transmission can occur through contact with infected fluids during birth or subsequently through breast feeding.[74] Less frequently, the virus is transmitted across the placenta.[75] Substituting formula in place of breast milk can help avert vertical transmission, but it is expensive and risks exposing babies to diarrheal disease through the use of contaminated water mixed with powdered formula.[76] Transmission at birth can often be prevented by administering antiretroviral drugs to the mother before the baby is born.[77] Until re-

---

[66] *Hearing*, *supra* note 32, at 2 (statement of Rep. Leach) (indicating that the Fund seeks to leverage U.S. contributions in order to raise the enormous amounts needed to address the AIDS epidemic).

[67] *See, e.g.*, 146 CONG. REC. H5941–2 (daily ed. July 12, 2000) (statement of Rep. Brad Sherman (D-Cal.)).

[68] Foreign Assistance Act of 1961, 22 U.S.C. § 2151.

[69] Global AIDS and Tuberculosis Relief Act of 2000 § 111(a).

[70] *Id.*

[71] *Id.*

[72] *Id.*

[73] *Id.* § 103.

[74] *A Turning Point*, *supra* note 12, at 77.

[75] *Id.*

[76] *Id.* at 78. Some women may fear that if they use formula, others in their community will identify them as having HIV/AIDS and will ostracize them. If women had better access to the formula, then more women might use it instead of breast feeding, thus making mothers with AIDS less identifiable. *Hearing*, *supra* note 32 (prepared testimony of Mary Fisher, founder of the Family AIDS Network), *available at* http://www.house.gov/financialservices/3800fis.htm.

[77] *A Turning Point*, *supra* note 12, at 78.

cently, this approach was not practical in developing nations because AZT, the drug most commonly given to expectant mothers, is too expensive to be used widely.[78] However, recent trials revealed that AZT can be effectively substituted with nevirapine at a cost of four dollars per birth.[79] Following this discovery, members of Congress introduced several bills to fund vertical transmission prevention programs.[80] The Senate-adopted Amendment 4018 reflects these suggestions.[81]

Even if children born to infected mothers escape infection, they face the heightened risk of being orphaned. Worldwide, 15.6 million children under the age of fifteen have lost either their mother or both parents to AIDS, a number that is expected to reach 24.3 million in 2010.[82] In addition to suffering the loss of parental care and love, AIDS orphans often quit school in order to support themselves and their siblings and lack access to health care.[83] Their vulnerability may place them at an increased risk of contracting AIDS themselves, especially if they engage in prostitution as a source of income.[84]

Helping AIDS orphans has been a central concern of United States foreign policy since December 1998, when President Clinton called on Sandra Thurman, Director of the White House Office of AIDS Policy, to lead a fact-finding mission in Africa and issue recommendations on how the United States could alleviate the plight of orphans.[85] The trip increased public awareness of this aspect of the AIDS epidemic.[86] The orphan crisis also influenced the lawmakers who supported House Bill 3519.[87] At least twenty percent of the $300 million authorized by the bill for foreign assistance must be directed toward the support and education of sub-Saharan African orphans, including those orphaned by AIDS.[88] These authorizations supplement other portions of the statute addressing the plight of orphans, including a provision directing the President to coordinate the development of a multidonor strategy for providing support and education for orphans.[89] In addition, the Act states that Trust

---

[78] *Id.*

[79] *Id.*; *Hearing*, *supra* note 32 (prepared testimony of Catherine M. Wilfert, M.D., Scientific Dir., Elizabeth Glaser Pediatric AIDS Foundation), *available at* http://www.house.gov/financialservices/3800wil.htm.

[80] *See, e.g.*, H.R. 4038, 106th Congress (2000); S. 2032, 106th Congress (2000).

[81] 146 CONG. REC. S7625, 7695 (daily ed. July 26, 2000).

[82] SUSAN HUNTER & JOHN WILLIAMSON, U.S. AGENCY FOR INT'L DEV., CHILDREN ON THE BRINK: EXECUTIVE SUMMARY 1 (2000).

[83] *Id.* at 4.

[84] *Id.* at 8.

[85] Remarks Announcing AIDS Initiatives, 2 PUB. PAPERS 2104–05 (Dec. 1, 1998).

[86] HUNTER & WILLIAMSON, *supra* note 82, at 6.

[87] While support for AIDS prevention programs often highlights lawmakers' differing judgments regarding the morality of sexual behavior, the plight of orphans elicits broad-based concern. *See, e.g.*, 146 CONG. REC. S7619-24 (daily ed. July 26, 2000).

[88] Global AIDS and Tuberculosis Relief Act of 2000, Pub. L. No. 106-264, § 111(a), 114 Stat. 748, 751–52.

[89] *Id.* § 113.

Fund grants should include funding for the support and education of orphans.[90]

While the bill addresses a number of the problems caused by the AIDS epidemic, the high cost of drugs and treatment leads some to argue that a vaccine is the only real solution to the AIDS epidemic.[91] The creation of an AIDS vaccine, however, presents significant challenges. Approaches commonly used to develop vaccines are ineffective against HIV.[92] In addition to the scientific challenges faced by researchers, market forces discourage many drug companies from developing vaccines. Because vaccines are administered only a few times, they earn less revenue than drugs that must be taken frequently.[93] The costs of administering the vaccine will be a heavy burden for impoverished countries, so market demand for an AIDS vaccine is uncertain in spite of the dire need.[94] It is impossible to determine how many years of research will be necessary to develop a successful vaccine.[95] However, some progress on a vaccine has already been made, and results from the first advanced-stage clinical trial of a vaccine against HIV are expected in the fall of 2001.[96]

House Bill 3519 supports vaccine development by authorizing $10 million annually in contributions to the International AIDS Vaccine Initiative ("IAVI") for fiscal years 2001 and 2002.[97] IAVI is a charity that funds the development of AIDS vaccines by pharmaceutical companies. Instead of taking a share of the profits from vaccines developed by fund

---

[90] *Id.* § 122(a)(2)(E).

[91] *See, e.g.*, John Carey & Amy Barrett, *An AIDS Vaccine Is No Longer a Dream*, BUS. WK., Sept. 6, 1999, at 94; Richard Marlink, *Lessons from the March of Dimes*, HARV. AIDS REV., Spring 1997, at 15.

[92] HIV is less vulnerable to antibodies than are other diseases because the virus mutates rapidly and antibodies cannot bind to the sugar-coated surfaces of its proteins. David Baltimore, *Can We Make an AIDS Vaccine?*, NAT'L FORUM, Summer 1999, at 35, 36.

[93] Carey & Barrett, *An AIDS Vaccine Is No Longer a Dream*, *supra* note 91, at 94.

[94] *See* Amy Barrett & John Carey, *'We Have to Find a Solution,'* BUS. WK., Sept. 6, 1999, at 98. This problem might be alleviated by the creation of a trust fund that would pay to distribute a vaccine once it has been developed. *Id.* Legislation sponsored by Sen. John Kerry (D-Mass.) and Rep. Nancy Pelosi (D-Cal.) during the 106th Congress would have created a vaccine purchase fund within the U.S. Treasury and directed the President to enter into negotiations to create a similar fund to receive international donations. S. 2132, 106th Cong. §§ 7–8 (2000); H.R. 3812, 106th Cong. §§ 7–8 (2000). The proposal would have created additional incentives for vaccine development by providing tax credits for vaccine research and sales. S. 2132 §§ 5–6; H.R. 3812 §§ 5–6.

[95] Baltimore, *supra* note 92, at 35.

[96] Carey & Barrett, *An AIDS Vaccine Is No Longer a Dream*, *supra* note 91, at 94. It may be necessary to develop separate vaccines for different strains of HIV found in different parts of the world. For example, vaccines based on strains that are prevalent in the United States may not protect against strains found in Africa. Carey & Barrett, *'We Have to Find a Solution,' supra* note 94, at 98. The first human trials of a possible vaccine based on HIV subtype A, which is the subtype most prevalent in eastern Africa, began in March of this year. Press Release, Int'l AIDS Vaccine Initiative, Trials of First Vaccine Candidate Designed for Africa Officially Begin in Nairobi (Mar. 6, 2001) [hereinafter IAVI], *available at* http://www.iavi.org.

[97] Global AIDS and Tuberculosis Relief Act of 2000, Pub. L. No. 106-264, § 112(a), 114 Stat. 748, 753.

recipients, IAVI requires firms to offer their new products at low profit margins.[98] Human trials on a vaccine sponsored by IAVI began in Nairobi earlier this year.[99]

When and if a vaccine is developed, distribution will pose a further challenge. Poor infrastructure—such as a lack of efficient transportation systems and problems with refrigeration—have hindered past immunization programs, and even today many children do not have access to basic vaccinations.[100] An organization working to remedy this problem is the Global Alliance for Vaccines and Immunizations ("GAVI"), a coalition of national governments, private foundations, and international organizations that promotes immunization programs around the world.[101] House Bill 3519 authorizes the appropriation of $50 million annually in contributions to GAVI for 2001 and 2002.[102] The bill specifically appropriates the funds to the Global Fund for Children's Vaccines, which develops vaccination programs in impoverished nations.[103]

While a vaccine ultimately may provide an inexpensive and effective solution to the AIDS crisis, it will come too late for the millions of people who are already infected or will become so before it has been developed.[104] Antiretroviral drugs could help African AIDS victims cope with the disease, but with a price of over $10,000 per year the treatment has been far too costly for most people.[105] While recent commitments by drug companies to reduce AIDS drug prices are raising hopes that treatment will become more affordable in the developing world, cost may still remain an obstacle at the anticipated discounts.[106]

America's pharmaceutical industry and public sector have played a large role in global AIDS treatment by taking the lead in developing new AIDS drugs,[107] so whether the United States pursues strong international patent protection for these technologies will have a direct effect on both the development of such drugs and their availability to AIDS sufferers abroad. On the one hand, well-enforced patent laws foster the development of these drugs by offering incentives to private research firms.[108] On

---

[98] *A Turning Point*, *supra* note 12, at 77.

[99] *See supra* note 96.

[100] John Donnelly, *Immunizations Plummet in Poorest Nations; Wars, Funding Cuts Blamed for Decline*, BOSTON GLOBE, Nov. 13, 2000, at A1; Gellman, *supra* note 12.

[101] Press Release, Global Alliance on Vaccines and Immunizations, Global Fund for Children's Vaccines Names President (July 17, 2000), *available at* http://www.vaccinealliance.org/press/press_president.html.

[102] Global AIDS and Tuberculosis Relief Act of 2000, § 112(a).

[103] *Id.* The Bill and Melinda Gates Foundation created the Global Fund for Children's Vaccines with a grant of $750 million. *Id.*

[104] *See Hearing*, *supra* note 32 (prepared testimony of Mary Fisher, founder of the Family AIDS Network), *available at* http://www.house.gov/financialservices/3800fis.htm.

[105] Finkel, *supra* note 6.

[106] Stolberg, *supra* note 8.

[107] *See* Exec. Order No. 13,155, 65 Fed. Reg. 30,521 (May 12, 2000).

[108] *See Science and Profit*, ECONOMIST, Feb. 17, 2001, at 21, 21–22 (arguing that wealthy countries should provide AIDS drugs to the developing world by increasing for-

the other hand, this protection gives the developer a monopoly on the drug, allowing it to price many of the poorest AIDS victims worldwide out of the market.[109]

During the House Banking Committee debate on House Bill 3519, the committee adopted by voice vote two amendments regarding drug affordability.[110] Representative Jan Schakowsky (D-Ill.) offered an amendment clarifying that funds could be used for the research and development of affordable drugs.[111] Representatives John LaFalce (D-N.Y.) and Maxine Waters (D-Cal.) offered an amendment calling for the Treasury Secretary to direct the United States Executive Director of the World Bank to "use his or her voice vote to promote the availability of affordable HIV/AIDS drugs."[112] However, the final version of House Bill 3519 contains no concrete proposals regarding the role of the United States in the enforcement of patent protections for AIDS drugs.[113]

Despite drug industry opposition, some governments have already produced generic versions of AIDS drugs.[114] Nevertheless, when faced

---

eign aid for drug purchases rather than by reducing drug company profits).

[109] For a discussion of how U.S. trade policy has affected AIDS treatment in Thailand, see Rosemary Sweeney, Comment, *The U.S. Push for Worldwide Patent Protection for Drugs Meets the AIDS Crisis in Thailand: A Devastating Collision*, 9 PAC. RIM L. & POL'Y J. 445 (2000).

[110] H.R. REP. NO. 106-548, at 10–11 (2000).

[111] *Id.* at 10.

[112] *Id.* at 11.

[113] Representative Waters's original proposal sought the World Trade Organization's nullification of the intellectual property rights agreements that prevent countries from providing generic versions of AIDS drugs. 146 CONG. REC. H2385 (daily ed. May 2, 2000) (statement of Rep. Waters).

[114] For example, in 1998 the Brazilian government began manufacturing AIDS drugs as part of its highly successful treatment and prevention program. Tina Rosenberg, *Look At Brazil*, N.Y. TIMES, Jan. 28, 2001, § 6 (Magazine), at 26. Governments can facilitate the availability of generic versions of AIDS drugs through compulsory licensing, which involves granting a license to manufacture a patented product without permission from the patent holder. Sara M. Ford, Comment, *Compulsory Licensing Provisions Under the TRIPs Agreement: Balancing Pills and Patents*, 15 AM. U. INT'L L. REV. 941, 945 (2000). The Trade Related Aspects of Intellectual Property ("TRIPS") agreement allows for compulsory licensing. *See id.* at 949. In March 2001, the Indian generic drug manufacturer Cipla requested such a license from the South African government to sell AIDS drugs. Rachel Swarns, *AIDS Drug Battle Deepens in Africa*, N.Y. TIMES, Mar. 8, 2001, at A1. After Cipla's announcement, several major pharmaceutical companies announced that they would dramatically decrease AIDS drug prices in Africa. Swarns, *supra* note 8, at A1. Some pharmaceutical companies oppose compulsory licensing, and a number of drug manufacturers filed suit in South Africa challenging the law that allows the government to grant Cipla's request. Stolberg, *supra* note 8, at A1. The companies dropped their suit in April, three years after filing. Rachel L. Swarns, *Drug Makers Drop South Africa Suit Over AIDS Medicine*, N.Y. TIMES, Apr. 20, 2001, at A1.

Parallel importing provides an additional channel through which generic AIDS drugs can be made available cheaply in developing countries. Third parties purchase drugs abroad, import them, and then sell them in another market at prices below those offered by official distributors. David Benjamin Snyder, Comment, *South Africa's Medicines and Related Substances Control Amendment Act: A Spoonful of Sugar or a Bitter Pill to Swallow?*, 18 DICK. J. INT'L L. 175, 180 (1999). Parallel importing is not forbidden by the TRIPS agreement. *Id.* at 188.

with potential United States trade retaliation, other countries have changed their patent policies to comply with pharmaceutical companies' demands.[115]

In response to this problem, President Clinton signed an executive order in May 2000 that declared that the United States would not seek changes in patent laws in sub-Saharan African countries that promoted access to HIV/AIDS drugs.[116] Based on a proposal by Senators Diane Feinstein (D-Cal.) and Russell Feingold (D-Wis.) that was deleted from the African Growth and Opportunity Act,[117] the order refers to patent law in sub-Saharan Africa only[118] and does not address the broader problem of drug access worldwide. Because it is an executive order, it is also more vulnerable to future attacks than legislation would be.[119] However, President Bush has announced that he will not overturn the policy.[120]

Although the funds authorized by House Bill 3519 would provide only a fraction of the $3 billion that UNAIDS estimates will be necessary for implementing prevention and care programs in Africa alone,[121] the legislation nonetheless provides a rallying cry in the fight against AIDS. This effort, however, is severely dampened by Congress's unwillingness to appropriate the majority of the funds authorized by House Bill 3519.[122] The legislation, and in particular the World Bank AIDS Trust Fund it establishes, could play an important leveraging function by using American contributions to encourage donations from other sources.[123] Funding spent for AIDS treatment will have an even greater impact as pharmaceutical companies follow through on pledges to reduce drug prices.[124] Congress's failure to adequately endow the Fund squanders this opportunity. Furthermore, the programs authorized in House Bill 3519 will require sustained funding over many years.[125] The Bush budget proposal would fund the provisions at their 2001 appropriations, including only

---

[115] *See supra* note 109 and accompanying text; *see also* Rosenberg, *supra* note 114, at 26 (describing the United States's aggressive efforts to dissuade South Africa from permitting compulsory licensing).

[116] Exec. Order No. 13,155, 65 Fed. Reg. 30,521 (May 12, 2000).

[117] Pub. L. No. 106-200, 114 Stat. 251 (2000).

[118] Exec. Order No. 13,155, 65 Fed. Reg. 30,521 (May 12, 2000).

[119] The President can overturn an executive order without Congressional approval. *See* Robert Pear, *The New Administration: The Regulations*, N.Y. TIMES, Jan. 23, 2001, at A16.

[120] Donald G. McNeil, Jr., *Bush Keeps Clinton Policy on Poor Lands' Need for AIDS Drugs*, N.Y. TIMES, Feb. 22, 2001, at A9. Legislation introduced during the 107th Congress addresses access to AIDS drugs and the United States's enforcement of intellectual property laws. *See, e.g.*, S. 463, 107th Cong. (2001).

[121] WHITE HOUSE, *supra* note 15, at 40.

[122] *See* Foreign Operations, Export Financing, and Related Programs Appropriations Act, 2001, Pub. L. 106-429, tit. 2, 114 Stat. 1900, 1900A-5 (2000).

[123] *See* Statement on Signing the Global AIDS and Tuberculosis Relief Act of 2000, 36 WEEKLY COMP. PRES. DOC. 1906 (Aug. 19, 2000) (stating that international cooperation is necessary to effectively fight AIDS).

[124] *See* Jeffrey D. Sachs, *A Modest Proposal*, NEWSWEEK, Mar. 19, 2001, at 20.

[125] *See* Tom Malinowski, *The Epidemic and the Administration*, WASH. POST, Feb. 9, 2001, at A29.

$20 million of the $150 million authorized in House Bill 3519 for contributions to the Fund,[126] leaving the future of the United States' commitment to fighting the AIDS epidemic in doubt.[127]

—Anna-Marie Tabor

---

[126] OFFICE OF MGMT. & BUDGET, BUDGET OF THE UNITED STATES, FISCAL YEAR 2002 app. at 1018 (2001).

[127] On May 11, 2001, President Bush announced that the United States would pledge $200 million to a new United Nation's fund for combating AIDS, malaria, and tuberculosis. David E. Sanger, *Bush Says U.S. Will Give $200 Million to World AIDS Fund*, N.Y. TIMES, May 12, 2001, at A4.

CYBERSQUATTING

The proliferation of access to the Internet has resulted in a new form of trademark abuse, known as "cybersquatting." Cybersquatting, as defined by a recent Senate report, is "the deliberate, bad-faith, and abusive registration of domain names[1] in violation of the rights of trademark owners."[2] Individuals can easily obtain such domain names because registrars assign the addresses for a flat fee on a first-come, first-served basis.[3]

The practice of cybersquatting can be carried out in a number of different ways.[4] Methods include the registration of another's mark as a domain name,[5] the registration of a misspelling of another's mark,[6] the registration of another's mark for use with a suffix other than ".com,"[7] the registration of another's mark as part of a domain name,[8] the registration of another individual's name as a domain name,[9] the registration of the mark of a competitor,[10] and the registration of the mark of an organization opposed by the registrant.[11]

Until recently, trademark owners had limited recourse against such practices. On November 29, 1999, however, President Clinton signed into law the Anticybersquatting Consumer Protection Act ("ACPA"), which offers a greater level of protection for trademark owners.[12]

---

[1] Congress has defined "domain name" as "any alphanumeric designation that is registered with or assigned by any domain name registrar, domain name registry, or other domain name registration authority as part of an electronic address on the Internet." 15 U.S.C. § 1127 (Supp. V 1999).

[2] S. REP. No. 106-140, at 4 (1999).

[3] *See* Panavision v. Toeppen, 141 F.3d 1316, 1318–19 (9th Cir. 1998) ("Domain names with the .com designation must be registered on the Internet with Network Solutions, Inc. ('NSI'). NSI registers names on a first-come, first-served basis for a $100 registration fee. NSI does not make a determination about a registrant's right to use a domain name."). *But see* Avery Dennison Corp. v. Sumpton, 189 F.3d 868, 872 n.1 (9th Cir. 1999) (indicating that NSI is no longer the exclusive registrar of domain names).

[4] Steven R. Borgman, *The New Federal Cybersquatting Laws*, 8 TEX. INTELL. PROP. L.J. 265, 266–67 (2000).

[5] *Id.* at 266–67; S. REP. No. 106-140, at 5 ("Some [individuals] register well-known brand names as Internet domain names in order to extract payment from the rightful owners of the marks, who find their trademarks 'locked up' and are forced to pay for the right to engage in electronic commerce under their own brand name.").

[6] Borgman, *supra* note 4, at 266–67; S. REP. No. 106-140, at 6 ("[C]ybersquatters often register well-known marks to prey on customer confusion by misusing the domain name to divert customers from the mark owner's site to the cybersquatter's own site, many of which are pornography sites that derive advertising revenue based on the number of visits, or 'hits,' the site receives.").

[7] Borgman, *supra* note 4, at 266–67.

[8] *Id.* at 266–67; 145 CONG. REC. S9749 (daily ed. July 29, 1999) (statement of Sen. Orrin Hatch (R-Utah)).

[9] Borgman, *supra* note 4, at 266–67.

[10] *Id.* at 266–67.

[11] *Id.*

[12] *See* 145 CONG. REC. H12908 (daily ed. Dec. 3, 1999). The Anticybersquatting Consumer Protection Act was introduced as S. 1255, 106th Cong. (1999) and later incorpo-

The Lanham Act,[13] the legal tool traditionally employed to protect trademarks, provides a civil cause of action against those who use a word, term, name, symbol, or device in a way that may cause confusion or mistake as to the affiliation between the user and another person or company, or that misrepresents the origin of the user's goods, services, or commercial activities.[14] The Lanham Act restricts only commercial uses of another's mark.[15]

Under the Lanham Act, when the alleged infringer's product differs from that of a mark holder:

> [T]he prior owner's chance of success is a function of many variables: the strength of his mark, the degree of similarity between the two marks, the proximity of the products, the likelihood that the prior owner will bridge the gap, actual confusion, and the reciprocal of defendant's good faith in adopting its own mark, the quality of defendant's product, and the sophistication of the buyers.[16]

In an attempt to show that the Lanham Act could cure cybersquatting, one commentator, Kevin Murch, applied these factors to an early cybersquatting dispute.[17] Murch argued that the factors could support a finding of liability in a case where one-time MTV personality Adam Curry registered and used the domain name mtv.com.[18] According to Murch, six of the eight factors favored MTV's position, while the other two were not yet capable of determinative application.[19] The strength of MTV's mark

---

rated into Title III (§§ 3001–3010) of S. 1948, 106th Cong. (1999). The latter bill was in turn enacted into law by Pub. L. No. 106-113, 113 Stat. 1501 (2000), an omnibus appropriations package. The provisions of the ACPA have been codified in chapter 22, subchapter III of 15 U.S.C., and all references to the provisions of the ACPA in this piece will be cited to the relevant section of the United States Code.

[13] Trademark Act of 1946, ch. 540, 60 Stat. 427 (1946) (codified as amended in scattered sections of 15 U.S.C.).

[14] *See* 15 U.S.C. § 1125(a)(1) (Supp. V 1999).

[15] The Act can only be employed against one who "uses in commerce" the mark of another. 15 U.S.C. § 1125(a)(1). Use of a mark is deemed "use in commerce" only when (i) it is applied to goods that are "sold or transported in commerce," or (ii) when used "in the sale or advertising of services" that are "rendered in commerce," or, in the case of interstate and international services, when "the person rendering the services is engaged in commerce in connection with the services." 15 U.S.C. § 1127.

[16] Polaroid Corp. v. Polarad Electronics Corp., 287 F.2d 492, 495 (2d Cir. 1961) (holding that plaintiff was barred from bringing a claim of infringement because it had delayed bringing the claim for 11 years, though it had known of the allegedly infringing use, and because the defendant's use was far removed from plaintiff's primary areas of activity).

[17] *See* Kevin L. Murch, *Cybercourt: Copyright and Trademark Law on the Information Superhighway*, 24 CAP. U. L. REV. 809, 819–22 (1995).

[18] MTV Networks v. Curry, 867 F. Supp. 202 (S.D.N.Y. 1994). The parties ultimately settled the dispute. *See* Murch, *supra* note 17, at 809.

[19] Murch, *supra* note 17, at 819–22.

could be shown by the company's reliance on the mark in its advertising, and by the fact that the company has never gone by any other name.[20] The marks used by MTV and Curry had a high degree of similarity, since each used the same three-letter symbol.[21] Although MTV's television product and Curry's Web site were not in close proximity, MTV had recently set up a bulletin board accessible through AOL. Having bridged this gap, the proximity factor would likely cut in MTV's favor.[22] Curry's prior knowledge of the MTV mark and the similarity between his mark and that of MTV make it likely that the "good faith" factor would favor plaintiff MTV.[23] The "sophistication of the buyers" factor would likewise favor MTV, for, although users of Curry's electronic bulletin board were more likely than the average person to be familiar with the underlying technology, they were also likely to know of Curry only because of his previous affiliation with MTV. In that sense, the relevant consumers cannot be said to be sufficiently sophisticated to avoid being confused by Curry's use of mtv.com, since they would probably assume that Curry was still affiliated with MTV and that MTV was involved in the administration of mtv.com.[24] As for the remaining two factors, neither actual confusion nor the quality of Curry's product had been empirically established.[25]

Murch's analysis is cogent; however, the Curry case was not a typical cybersquatting dispute. Two details set Curry apart from other cybersquatters. Adam Curry actually worked for MTV for a number of years,[26] and he employed the domain name for commercial purposes.[27] The former fact made it likely that MTV viewers and mtv.com visitors would associate Curry with MTV and be confused as to who actually administered the Web site. This brings the case within the purview of the Lan-

---

[20] *Id.* at 820.
[21] *Id.*
[22] *Id.* at 820–21.
[23] *Id.* at 821.
[24] *Id.* at 821–22.
[25] *Id.* at 822.
[26] MTV Networks v. Curry, 867 F. Supp. 202, 203 (S.D.N.Y. 1994).
[27] *Id.* at 204. In drafting the Anticybersquatting Consumer Protection Act, the Senate Committee on the Judiciary enumerated the most common forms of cybersquatting. *See* S. REP. No. 106-140, at 5–7 (1999). Nowhere does the Committee mention a concern with the misuse of well-known brand names by former employees likely to be identified by the public with their former employer's mark. Rather, the Committee explains that a major goal of the legislation is to create a legal tool applicable to infringers who avoid the commercial use requirement of pre-existing law:

> [C]ybersquatters have become increasingly sophisticated as the case law has developed and now take the necessary precautions to insulate themselves from liability. For example, many cybersquatters are now careful to no longer offer the domain name for sale in any manner that could implicate liability under existing trademark dilution case law.

S. REP. No. 106-140, at 7.

ham Act, which applies anytime the use of another's trademark is likely to cause confusion or mistake or to deceive consumers as to the "affiliation, connection or association of such person with another person, or as to the origin, sponsorship, or approval of his or her goods, services, or commercial activities by another person."[28] Most cybersquatters, however, will not have any affiliation with the owner of the mark in question.

The applicability of the Lanham Act's "use in commerce" clause, which states that the Act only applies to one who uses the mark of another "in commerce,"[29] further sets Curry's conduct apart from typical cybersquatting. Millions of users accessed Curry's Web site,[30] providing Curry with the possibility of obtaining advertising revenue. This made it possible for MTV to claim that Curry had used the company's mark "in commerce."[31] Other trademark holders, however, may not be able to bring such claims, since a cybersquatter who registers a domain name solely for the purpose of holding it ransom against its rightful owner will not technically be using that domain name to run a business.

Another remedy available to mark owners is the Federal Trademark Dilution Act ("FTDA").[32] While Congress designed the law to prevent "dilution"—the weakening of a mark's distinctiveness through commercial use by non-owners[33]—at least one senator hoped the new law would apply to instances of cybersquatting.[34] Indeed, some mark owners have used the remedies provided by the bill to win judgments against parties who registered their marks as domain names.[35]

Under the FTDA, the owner of a "famous mark" is entitled to an injunction against another person's commercial use of a mark or trade name if that use "causes dilution of the distinctive quality of the mark."[36]

---

[28] 15 U.S.C. § 1125(a)(1)(A) (Supp. V 1999).

[29] 15 U.S.C. § 1125(a)(1).

[30] *MTV Networks*, 867 F.Supp. at 204.

[31] The *MTV Networks* opinion does not make clear whether Curry actually profited from mtv.com. *See id.*

[32] 15 U.S.C. § 1125(c).

[33] *See* 141 CONG. REC. S19,310–11 (daily ed. Dec. 29, 1995) (statement of Sen. Orrin Hatch upon introducing the FTDA) ("[T]his bill is designed to protect famous trademarks from subsequent uses that blur the distinctiveness of the mark or tarnish or disparage it, even in the absence of a likelihood of confusion. Thus, for example, the use of DuPont shoes, Buick aspirin, and Kodak pianos would be actionable under this bill.").

[34] "Although no one else has yet considered this application, it is my hope that this anti-dilution statute can help stem the use of deceptive Internet addresses taken by those who are choosing marks that are associated with the products and reputations of others." 141 CONG. REC. S19,312 (daily ed. Dec. 29, 1995) (statement of Sen. Patrick Leahy (D-Vt.)).

[35] *See, e.g.*, Panavision v. Toeppen, 141 F.3d 1316 (9th Cir. 1998) (holding defendant liable under the FTDA for registering panavision.com and attempting to sell it for $13,000 to plaintiff, owner of the trademark Panavision); Intermatic Inc. v. Toeppen, 947 F. Supp. 1227 (N.D. Ill. 1996) (issuing a permanent injunction under the FTDA enjoining defendant from using intermatic.com, which defendant had previously registered and used).

[36] 15 U.S.C. § 1125(c)(1).

In order to prevail in a trademark dilution claim the plaintiff must demonstrate the existence of dilution, commerce, and famousness.[37] The FTDA defines dilution as "the lessening of the capacity of a famous mark to identify and distinguish goods or services, regardless of the presence or absence of—(1) competition between the owner of the famous mark and other parties, or (2) likelihood of confusion, mistake or deception."[38] Courts are encouraged to consult a list of eight factors in determining whether a mark is famous: (1) the mark's distinctiveness; (2) the duration and extent of the mark's use at issue; (3) the duration and extent of advertising and publicity of the mark; (4) the geographical extent of the trading area in which the mark is used; (5) the channels of trade for the goods or services with which the mark is used; (6) the level of recognition of the mark at issue, measured in the trading areas and channels of trade of both the mark's owner and the alleged dilutor; (7) third parties' use of the mark; and (8) whether the mark was registered.[39]

Unlike the Lanham Act, the FTDA does not require the mark holder to show that the defendant's practice would cause confusion. Instead, the plaintiff need only prove that the defendant's practice dilutes the mark. Courts have held that cybersquatting satisfies the dilution element. For instance, in *Intermatic Inc. v. Toeppen*,[40] a federal court ruled that "Toeppen's action in registering and using 'intermatic.com' as a domain name violate[d] [the FTDA] and the Illinois Anti-Dilution Act because it lessens the capacity of a famous mark, Intermatic, to identify and distinguish goods or services as a matter of law."[41] The court reached this decision even though Toeppen's site displayed only a map of Champaign-Urbana for most of the time it was active.[42] Magistrate Judge Denlow wrote:

> Dilution of Intermatic's mark is likely to occur because the domain name appears on the web page and is included on every page that is printed from the web page .... Attaching Intermatic's name to a myriad of possible messages, even something as innocuous as a map of Urbana, Illinois, is something that the [Federal Trademark Dilution] Act does not permit .... The fact that "intermatic.com" will be displayed on every aspect of the web page is sufficient to show that Intermatic's mark will likely be diluted.[43]

---

[37] *See, e.g.*, Westchester Media v. PRL USA Holdings, Inc., 214 F.3d 658, 670–71 (5th Cir. 2000); Viacom Inc. v. Ingram Enterprises, Inc., 141 F.3d 886, 888 (8th Cir. 1998); Nabisco, Inc. v. PF Brands, Inc., 50 F. Supp. 2d 188, 197 (S.D.N.Y. 1999).

[38] 15 U.S.C. § 1127.

[39] *Id.*

[40] 947 F. Supp. 1227 (N.D. Ill. 1996).

[41] *Id.* at 1236.

[42] *Id.* at 1232.

[43] *Id.* at 1240–41.

This decision suggests that any registration and use of a trademarked domain name could constitute dilution. Thus, the FTDA's "dilution" standard could be more effective in curtailing cybersquatting than the older Lanham Act's requirement of "confusion."

The FTDA, like the Lanham Act, requires the mark holder to show that the defendant made commercial use of the mark.[44] In *Intermatic*, defendant Toeppen did not (except for a very short period which ended prior to the effective date of the FTDA)[45] use the site to promote any sort of goods or services.[46] Nonetheless, Magistrate Judge Denlow held that Toeppen's behavior satisfied the "commercial use in commerce"[47] requirement of the FTDA. He reasoned that "Toeppen's intention to arbitrage the 'intermatic.com' domain name constitutes a commercial use"[48] and that "Toeppen's use of the Internet satisfies the 'in commerce' requirement" of the FTDA because of the instantaneous, worldwide nature of Internet communications and because the Supreme Court has held that the "in commerce" clause should be construed liberally.[49] *Intermatic* therefore suggests that a showing that the defendant tried to sell the domain name to the mark holder could constitute a "commercial use in commerce."

The FTDA also requires that the mark in question be famous.[50] Although "Intermatic" hardly seems to be a "famous mark," Magistrate Judge Denlow held that the mark was famous because it "is a strong fanciful federally registered mark, which has been exclusively used by Intermatic for over 50 years."[51]

Based on the holding in *Intermatic v. Toeppen*, therefore, the FTDA may be capable of preventing and punishing cybersquatting. The FTDA must be interpreted rather broadly, however, in order for it to apply to many instances of cybersquatting. Critics argue that *Intermatic* represents an overly broad application of the statute in interpreting the terms "dilution"[52] and "famous."[53] Consequently, not all plaintiffs have prevailed

---

[44] 15 U.S.C. § 1125(c)(1).

[45] *Intermatic*, 947 F. Supp. at 1239.

[46] *Id.* at 1232.

[47] 15 U.S.C. § 1125(c)(1).

[48] *Intermatic*, 947 F. Supp. at 1239.

[49] *Id.* (citing 1 GILSON, TRADEMARK PROTECTION AND PRACTICE § 5.11[2] (1996); Steele v. Bulova Watch Co., 344 U.S. 280 (1952)).

[50] 15 U.S.C. § 1125(c)(1).

[51] *Intermatic*, 947 F. Supp. at 1239.

[52] *See, e.g.*, 4 MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 25:77 (1996) ("In the author's opinion, there is a very poor fit between the actions of a cybersquatter and the federal Anti-dilution Act. The prototypical cybersquatter does not use the reserved domain name as its mark before the public, so there is no traditional dilution by blurring or tarnishment.").

[53] *See, e.g., id.* ("[T]his legal tool only protects 'famous' marks, requiring that the courts expand and devalue the category of 'famous' marks in order to combat cybersquatting.").

against cybersquatters in suits brought under the FTDA. For example, when Avery Dennison Corporation sued a vanity e-mail service that had reserved domain names composed of common surnames and the .net suffix, including Avery.net and Dennison.net,[54] the Ninth Circuit Court of Appeals held that plaintiff Avery Dennison did not satisfy either the famousness[55] or commercial use[56] elements of the statute.[57] Many suits predicated upon the FTDA have suffered similar fates, particularly when it comes to satisfying the FTDA's famousness requirement. Thus, the FTDA seems an imperfect tool for preventing cybersquatting.[58] It poses a dilemma to courts facing incidents of cybersquatting: they must either

---

> Federal dilution law, protecting as it does "famous" trademarks—the ones most attractive to cybersquatters—is a timely development for the owners of those trademarks. One may argue, however, that in their desire to remedy the cybersquatting problem, the courts have violated legislative intent—the statute is intended to protect truly famous marks, a category in which INTERMATIC, and many other marks, may not properly belong.

Robert C. Scheinfeld & Parker H. Bagley, *Long-Arm Jurisdiction: 'Cybersquatting,'* N.Y.L.J., Nov. 27, 1996, at 3

[54] Avery Dennison Corp. v. Sumpton, 189 F.3d 868 (9th Cir. 1999); *see also* Gregory B. Blasbalg, Note, *Masters of Their Domains: Trademark Holders Now Have New Ways to Control Their Marks in Cyberspace*, 5 ROGER WILLIAMS U. L. REV. 563, 582–85 (2000) (suggesting the court might have ruled in favor of plaintiff Avery Dennison had the ACPA been in effect).

[55] *See Avery Dennison*, 189 F.3d at 876–77.

> Applying the famousness factors from the Federal Trademark Dilution Act to the facts of the case at bench, we conclude that Avery Dennison likely establishes acquired distinctiveness in the "Avery" and "Dennison" trademarks, but goes no further. Because the Federal Trademark Dilution Act requires a showing greater than distinctiveness to meet the threshold element of fame, as a matter of law Avery Dennison has failed to fulfill this burden.

Id.

[56] *See id.* at 880.

> All evidence in the record indicates that Appellants register common surnames in domain-name combinations and license e-mail addresses using those surnames, with the consequent intent to capitalize on the surname status of "Avery" and "Dennison." Appellants do not use trademarks qua trademarks as required by the caselaw to establish commercial use. Rather, Appellants use words that happen to be trademarks for their non-trademark value. The district court erred in holding that Appellants' use of avery.net and dennison.net constituted commercial use under the Federal Trademark Dilution Act, and this essential element of the dilution causes of action likewise mandates summary judgment for Appellants.

Id.

[57] *Id.* at 880–81.

[58] *See, e.g.*, Carnival Corp. v. SeaEscape Casino Cruises, Inc., 74 F. Supp. 2d 1261, 1268–71 (S.D. Fla. 1999) (denying plaintiff any relief under the Lanham Act); Hasbro, Inc. v. Clue Computing, Inc., 66 F. Supp. 2d 117, 130–37 (D. Mass. 1999); Washington Speakers Bureau, Inc. v. Leading Authorities, Inc., 33 F. Supp. 2d 488, 502–04 (E.D. Va. 1999) (holding plaintiff's marks were insufficiently famous for plaintiff to maintain a claim under the FTDA, but granting relief under section 1125(a) of the Lanham Act).

over-generalize legal categories or else fail to protect the third-party use of established trademarks as Web site domain names.

Given that previous trademark law did not squarely address cybersquatting, Congress decided that "legislation [was] needed to clarify the rights of trademark owners with respect to bad faith, abusive domain name registration practices, to provide clear deterrence to prevent bad faith and abusive conduct, and to provide adequate remedies for trademark owners in those cases where it does occur."[59] In an allusion to the "use in commerce" requirements of the Lanham Act and the FTDA, the Senate Committee on the Judiciary noted that "many cybersquatters are now careful to no longer offer the domain name for sale in any manner that could implicate liability under existing trademark dilution case law."[60] Commentators also suggested that cybersquatters who avoided offering their domain names for sale were nonetheless successful in getting owners of marks to pay for the right to use their marks on the internet.[61]

Senator Spencer Abraham (R-Mich.) introduced the earliest version of the ACPA[62] on June 21, 1999.[63] Following a Senate Judiciary Committee hearing, committee chairman Senator Orrin Hatch (R-Utah) and ranking Democrat Senator Patrick Leahy (D-Vt.) offered the amended version of the bill,[64] which was ultimately signed into law.[65] This amended version offered trademark holders more protection than that offered by the original draft.[66] Unlike the original version of the bill, the final version states a substantive cause of action,[67] provides for in rem jurisdiction,[68] protects domain name registrants against reverse domain name hijacking,[69] and explicitly states that the ACPA does not preclude mark holders from employing any traditional defenses to a charge of trademark infringement.[70]

---

[59] S. REP. NO. 106-140, at 7–8 (1999).
[60] *Id.* at 7.
[61] *See, e.g.*, Blasbalg, *supra* note 54, at 565.

> Typically, when the company contacts the registrant they are informed by the cybersquatter that the company can purchase the domain name for $10,000, which the cybersquatter is quick to point out, is far less than the cost of litigating the matter. Since their competitors already have successful web sites and the company feels it is losing business by not having an online presence, some companies choose to give in and pay for the right to use their own trademark on the Internet.

Id.

[62] S. 1255, 106th Cong. (1999).
[63] *See* 145 CONG. REC. S7334–36 (daily ed. June 21, 1999).
[64] S. 1461, 106th Cong. (1999).
[65] *See* 145 Cong Rec. S9749–55 (daily ed. July 29, 1999).
[66] *See id.* at S9749 (statement of Sen. Leahy).
[67] 15 U.S.C. § 1125(d)(1)(A) (Supp. V 1999).
[68] *Id.* § 1125(d)(2)(A).
[69] *Id.* § 1114(d)(iv)–(v); *see also infra* text accompanying notes 131–132.
[70] 15 U.S.C. § 1125(c)(3).

The ACPA creates a cause of action distinct from other trademark actions under the Lanham Act and the FTDA, specifically addressing cybersquatting.[71] The mere registration of another's mark as a domain name is actionable under the ACPA, irregardless of the manner in which the domain name is used and whether the registrant actively seeks to sell the domain name.[72] Liability under the ACPA merely requires "bad faith intent to profit,"[73] a broader, easier-to-meet standard than the use in commerce requirements of the Lanham Act and the FTDA.

The ACPA targets the bad-faith registration of trademarks as domain names.[74] Specifically, bad faith intent to profit is a necessary element of the civil cause of action created by the ACPA.[75] The ACPA suggests that courts consider nine factors in determining whether an alleged cybersquatter has acted with the required bad-faith intent.[76] The first four factors suggest instances in which use of a domain name is not motivated by a bad-faith intent to profit, while the next four suggest behavior that signals such bad-faith intent.[77] The final factor directs courts to consider whether the mark at issue would be protected by the FTDA.[78]

---

[71] *Id.* § 1125(d)(1)(A).

> A person shall be liable in a civil action by the owner of a mark, including a personal name which is protected as a mark under this section, if, without regard to the goods or services of the parties, that person—
>
> (i) has a bad faith intent to profit from that mark, including a personal name which is protected as a mark under this section; and
>
> (ii) registers, traffics in, or uses a domain name that—
>
> (I) in the case of a mark that is distinctive at the time of registration of the domain name, is identical or confusingly similar to that mark;
>
> (II) in the case of a famous mark that is famous at the time of registration of the domain name, is identical or confusingly similar to or dilutive of that mark; or
>
> (III) is a trademark, word, or name protected by reason of section 706 of title 18, United States Code, or section 220506 of title 36, United States Code.

Id.

[72] *See id.* § 1125(d)(1)(A)(ii).

[73] *Id.* § 1125(d)(1)(A)(i).

[74] *See* S. REP. No. 106-140, at 4 (1999).

> The purpose of the bill is to protect consumers and American businesses, to promote the growth of online commerce, and to provide clarity in the law for trademark owners by prohibiting the bad-faith and abusive registration of distinctive marks as Internet domain names with the intent to profit from the goodwill associated with such marks—a practice commonly referred to as "cybersquatting."

Id.

[75] 15 U.S.C. § 1125(d)(1)(A)(i).

[76] *Id.* § 1125(d)(1)(B)(i)(I)–(IX).

[77] S. REP. No. 106-140, at 13.

[78] 15 U.S.C. § 1125(d)(1)(B)(i)(IX) (directing courts to consider "the extent to which the mark incorporated in the person's domain name registration is or is not distinctive and famous within the meaning of subsection (c)(1) of this section").

The first of the ACPA's bad-faith factors recommends that courts consider "the trademark or other intellectual property rights of the person, if any, in the domain name."[79] This factor acknowledges that a domain name may bear a relation to two different trademarks.[80] In *Sporty's Farm v. Sportsman's Market*, which "appear[ed] to be the first interpretation of the ACPA at the appellate level,"[81] the Second Circuit Court of Appeals applied this factor narrowly, considering only rights that existed at the time the domain name was registered.[82]

The second factor is "the extent to which the domain name consists of the legal name of the person or a name that is otherwise commonly used to identify that person."[83] Congress intended this factor to protect individuals who registered their names or nicknames as domain names.[84] The statute's drafters specifically disclaimed an intent to permit an individual to avoid liability by adopting a well-known mark as a nickname and then registering that mark as a domain name.[85]

The third factor considers "the person's prior use, if any, of the domain name in connection with the bona fide offering of any goods or services."[86] Congress intended for courts to infer an absence of bad-faith intent to profit from the use of a domain name in the course of legitimate commerce that neither creates a likelihood of confusion regarding the identity of the registrant nor attempts to profit by exploiting the goodwill of a trademark owner's name.[87] One commentator suggests that application of this factor may lead courts to employ faulty reasoning:

> [C]ourts routinely look to a registrant's bad-faith intent to *establish* a finding of likelihood of confusion. Thus, use of this factor may require circular logic: registering a domain name with bad faith indicates a likelihood of confusion . . . which in turn indicates the registrant's bad faith for the purpose of the ACPA.[88]

---

[79] 15 U.S.C. § 1125(d)(1)(B)(i)(I).

[80] S. REP. NO. 106-140, at 13.

[81] Sporty's Farm L.L.C. v. Sportsman's Mkt., Inc., 202 F.3d 489, 496 (2d Cir. 2000).

[82] In an opinion by Judge Calabresi, the court ruled that counter-defendant Sporty's Farm had no intellectual property rights in the domain name at issue—"sportys.com"—because Sporty's Farm had not been formed until nine months after its parent company registered the name. *Id.* at 498.

[83] 15 U.S.C. § 1125(d)(1)(B)(i)(II).

[84] S. REP. NO. 106-140, at 13.

[85] *Id.*

[86] 15 U.S.C. § 1125(d)(1)(B)(i)(III).

[87] S. REP. NO. 106-140, at 13–14.

[88] *See* Neil L. Martin, *The Anticybersquatting Consumer Protection Act: Empowering Trademark Owners, but not the Last Word on Domain Name Disputes*, 25 J. CORP. L. 591, 600–01 (2000) (citation omitted).

However, this logical concern may be addressed by pointing out that the statute does not direct the court to infer bad-faith intent from the existence of a likelihood of confusion; rather, the *absence* of a likelihood of confusion is merely intended to suggest good faith.[89]

The statute's fourth bad-faith factor is the registrant's "bona fide noncommercial or fair use of the mark in a site accessible under the domain name."[90] This factor allows courts to take into account "the interests of those who would make lawful noncommercial or fair uses of others' marks online, such as in comparative advertising, comment, criticism, parody, news reporting, etc."[91]

The fifth bad faith factor—the first of the set that lists instances in which the presence of bad-faith intent to profit should be inferred—urges courts to consider

> the person's intent to divert customers from the mark owner's online location to a site accessible under the domain name that could harm the goodwill represented by the mark, either for commercial gain or with the intent to tarnish or disparage the mark, by creating a likelihood of confusion as to the source, sponsorship, affiliation, or endorsement of the site. . . .[92]

This factor's inclusion in the ACPA indicates Congress's recognition that cybersquatting is typically motivated by a desire to register a well-known mark as a domain name in order to divert internet users to the cybersquatter's own site.[93] Such behavior allows the cybersquatter to

> pass off inferior goods under the name of a well-known markholder, to defraud consumers into providing personally identifiable information, such as credit card numbers, to attract eyeballs to sites that price online advertising according to the number of "hits" the site receives, or even just to harm the value of the mark.[94]

The sixth bad-faith factor is the registrant's

> offer to transfer, sell, or otherwise assign the domain name to the mark owner or any third party for financial gain without having used, or having an intent to use, the domain name in the

---

[89] S. REP. No. 106-140, at 13–14.
[90] 15 U.S.C. § 1125(d)(1)(B)(i)(IV).
[91] S. REP. No. 106-140, at 14.
[92] 15 U.S.C. § 1125(d)(1)(B)(i)(V).
[93] S. REP. No. 106-140, at 14.
[94] *Id.* at 14–15.

> bona fide offering of any goods or services, or the person's prior conduct indicating a pattern of such conduct.[95]

This factor reflects Congress's finding that "in practice, the offer to sell domain names for exorbitant amounts to the rightful mark owner has been one of the most common threads in abusive domain name registrations."[96] The legislative history stresses that this factor, like the other factors,[97] is not conclusive with regard to the domain name registrant's bad faith.[98]

The seventh bad-faith factor asks courts to consider the accused cybersquatter's "provision of material and misleading false contact information when applying for the registration of the domain name, the person's intentional failure to maintain accurate contact information, or the person's prior conduct indicating a pattern of such conduct."[99] This factor is intended to balance Congress's recognition that "[f]alsification of contact information with the intent to evade identification and service of process by trademark owners is also a common thread in cases of cybersquatting,"[100] with its concern "that there may be circumstances in which the provision of false information may be due to other factors, such as mistake or, as some have suggested in the case of political dissidents, for purposes of anonymity."[101]

The eighth factor directs courts to consider the alleged cybersquatter's:

> registration or acquisition of multiple domain names which the person knows are identical or confusingly similar to marks of others that are distinctive at the time of registration of such domain names, or dilutive of famous marks of others that are fa-

---

[95] 15 U.S.C. § 1125(d)(1)(B)(i)(VI).

[96] S. REP. No. 106-140, at 15.

[97] *See id.* at 14 ("[T]he bill gives courts the flexibility to weigh appropriate factors in determining whether the name was registered or used in bad faith, and it recognizes that one such factor may be the use the domain name registrant makes of the mark.").

[98] *Id.* at 15.

> [This factor] does not suggest that a court should consider the mere offer to sell a domain name to a mark owner or the failure to use a name in the bona fide offering of goods and services is sufficient to indicate bad faith. . . . It merely provides a court with the necessary discretion to recognize the evidence of bad-faith when it is present.

Id.

[99] 15 U.S.C. § 1125(d)(1)(B)(i)(VII).

[100] S. REP. No. 106-140, at 15.

[101] *Id.* at 15.

> mous at the time of registration of such domain names, without regard to the goods or services of the parties.[102]

The drafters intended this factor to address a form of cybersquatting known as "warehousing," in which an individual registers multiple domain names that resemble various trademarks.[103] The cybersquatter then waits for one of the trademark owners to offer to buy one of the addresses. Since the cybersquatter does not solicit buyers for the domain names, he avoids liability under the FTDA.[104]

The ninth factor is "the extent to which the mark incorporated in the person's domain name registration is or is not distinctive and famous within the meaning of [the FTDA]."[105] The statute refers to the subsection of the FTDA listing factors courts may consider in determining whether a mark is "distinctive and famous."[106]

The ACPA provides a fair use absolute defense against the bad-faith intent to profit element of the cybersquatting cause of action: "Bad faith intent described under subparagraph (A) shall not be found in any case in which the court determines that the person believed and had reasonable grounds to believe that the use of the domain name was a fair use or otherwise lawful."[107] Thus, if a defendant proves both a subjective and objective belief that his use of the domain name was fair or lawful, he may defeat an ACPA claim.

The ACPA also treats the problems faced by plaintiffs in instituting in rem proceedings under prior law. In the case of *Porsche v. Porsch.com*,[108] the famous automobile maker brought an in rem suit under the FTDA against 128 domain names that were variations on Porsche.com. The variations on Porsche.com included Porsch.com and Porsche.net, as well as domain names related to the company's automobiles, such as Boxter.com.[109] The district court noted:

> "PORSCHE.NET" and "PORSCHECLUB.NET" are registered domain names that Porsche cannot use because NSI already has assigned them away to others. Both domain names are dormant in the sense that they have no connection to other existing web sites, but the mere act of registration creates an immediate injury by preventing Porsche from utilizing those domain names itself in order to channel consumers to its own web site. Cus-

---

[102] 15 U.S.C. § 1125(d)(1)(B)(i)(VIII).

[103] *See* S. REP. No. 106-140, at 5–6.

[104] *Id.* at 15–16.

[105] 15 U.S.C. § 1125(d)(1)(B)(ii)(IX).

[106] *Id.* § 1125(c)(1); *see also supra* text accompanying notes 50–58 (discussing famousness as applied to cybersquatting cases).

[107] *Id.* § 1125(d)(1)(B)(ii).

[108] Porsche Cars N. Am., Inc. v. Porsch.com, 51 F. Supp. 2d 707 (E.D. Va. 1999).

[109] *Id.* at 709–10.

> tomers might try to contact Porsche through "PORSCHE.NET," for example, only to find that they have reached a "dead end" on the Web and then to conclude that the strength of Porsche's brand name is not as great as they first thought.[110]

Though the court expressed sympathy for Porsche's position, it ultimately dismissed the company's action for lack of personal jurisdiction, holding that the FTDA did not permit in rem proceedings.[111] Judge Cacheris reasoned that interpreting the FTDA as allowing in rem claims, in the absence of language in the statute explicitly authorizing plaintiffs to proceed in rem, would call into question the Act's constitutionality.[112]

In response to the problem demonstrated by *Porsche v. Porsch.com*, the ACPA allows mark holders to bring in rem actions for trademark infringement when the owner of a domain name cannot be located.[113] This provision of the ACPA reflects the congressional finding that many cybersquatters register offending domain names using aliases or other false information, thus making them difficult to track down.[114] Congress also articulated a desire to protect those who are online incognito for some legitimate purpose, and intended the ACPA's in rem provision to reduce the need for trademark owners to "root out" those who seek online anonymity as a means of protecting themselves.[115] The ACPA permits mark owners to file in rem civil actions in the judicial district in which the authority that assigned the domain name is located, provided that: (1) the domain name violates rights triggered by registration of the mark with the Patent and Trademark Office, or rights protected by the either the Lanham Act or the FTDA; and (2) the owner either cannot get in personam jurisdiction over a potential defendant under the ACPA or, through due diligence, cannot find a defendant, after both sending and publishing notice of the action.[116] This aspect of the ACPA was challenged and up-

---

[110] *Id.* at 710.
[111] *Id.* at 711–13.
[112] *Id.* at 713.

> Porsche correctly observes that some of the domain names at issue have registrants whose identities and addresses are unknown and against whom *in personam* proceedings might be fruitless. But most of the domain names in this case have registrants whose identities and addresses are known, and who rightly would object to having their interests adjudicated *in absentia*. The Due Process Clause requires at least some appreciation for the differences between these two groups, and Porsche's pursuit of an *in rem* remedy that fails to differentiate between them at all is fatal to its Complaint.

Id.

[113] *See* 15 U.S.C. § 1125(d)(2)(A) (Supp. V 1999).
[114] S. REP. No. 106-140, at 10 (1999).
[115] *Id.* at 11.
[116] 15 U.S.C. § 1125(d)(2)(A)(i)–(ii).

held in federal district court in Virginia. In the case of *Caesars World, Inc. v. Caesars-palace.com*,[117] the court found that the in rem action provided by the ACPA meets constitutional due process standards.

Like the Lanham Act and the FTDA, the ACPA provides for both injunctive relief[118] and actual damages.[119] Because it is often difficult to quantify the losses caused by diverting customers to another Web site,[120] the ACPA allows plaintiffs to recover statutory damages in lieu of actual damages. If the plaintiff so chooses, the court may set damages between $1,000 and $100,000 per domain name as it "considers just."[121]

The drafters of the ACPA intended to encourage domain name registrars[122] "to work with trademark owners to prevent cybersquatting."[123] The statute provides a mechanism by which domain name registrars may insulate themselves from liability for assigning addresses to cybersquatters. The ACPA states that a domain name registrar cannot be held liable for "refusing to register, removing from registration, transferring, temporarily disabling, or permanently canceling a domain name" as long as such action complies with a court order under 15 U.S.C. § 1125(d) or with the registrar's own "reasonable policy" prohibiting registration of domain names identical to, confusingly similar to, or dilutive of another's mark.[124] The registrar can never be liable for damages for registering or maintaining a domain name unless it can be shown that the registrar, in doing so, had a bad faith intent to profit.[125] The only relief that may be awarded against a registrar is injunctive, which may be available if the registrar: (1) failed to provide a court in which an ACPA action has been filed documents regarding the registration and use of the domain name; (2) transferred, suspended or otherwise modified the domain name while the action was pending (i.e., otherwise than by court order); or (3) willfully failed to comply with a court order to transfer, suspend or otherwise modify the domain name.[126] One commentator has expressed concern that the ACPA's lack of a requirement for domain name registries to conduct bad-faith analyses of questionable registrations may encourage "domain name registries to create a policy canceling a domain name registration based only on a showing that the domain name is identical to

---

[117] 112 F. Supp. 2d 502, 502–05 (E.D. Va. 2000).

[118] 15 U.S.C. § 1116(a) (extending injunctive relief already available under the Lanham Act to plaintiffs suing under the ACPA).

[119] 15 U.S.C. § 1117(a) (extending damage remedies already available under the Lanham Act to plaintiffs suing under the ACPA).

[120] *See* Martin, *supra* note 88, at 607.

[121] 15 U.S.C. § 1117(d).

[122] Registrars are companies that "assign domain names on a first-come, first served basis upon payment of a registration fee." Morrison & Foerster LLP v. Brian Wick & Am. Distribution Sys., Inc., 94 F. Supp. 2d 1125, 1126–27 (D. Colo. 2000).

[123] S. REP. No. 106-140, at 11 (1999).

[124] *See* 15 U.S.C. § 1114(1)(D).

[125] *See id.* §1114(1)(D)(iii).

[126] *See id.* § 1114(1)(D)(i)(II).

or confusingly similar to a trademark."[127] If a registry adopts such a policy of presuming cybersquatting, registrants will be unable to employ the "fair use" defense[128] provided by the ACPA.[129] Nothing in the ACPA's legislative history assuages this fear, and in fact Congress seems to have contemplated that domain name registrars' policies not include a bad faith element.[130]

The ACPA also protects individuals from those trademark holders that would use their marks to prevent the registration of unique domain names. This practice, called "reverse domain name hijacking," refers to any attempt to enjoin another's use of a mark as a domain name, even though the domain name uses the mark in a way that is unlikely to confuse consumers.[131] In instances where a person causes a domain name registrar to take unwarranted action against a registrant by misrepresenting that a domain name is identical to, confusingly similar to, or dilutive of a mark, the ACPA provides that the person making the misrepresentation may be liable to the domain name registrant for resulting damages, costs, and attorney's fees.[132]

The ACPA has been criticized for being overly generous to mark owners,[133] for its broad empowerment of domain name registrars, coupled with its encouragement of registrars to comply with the wishes of mark holders,[134] and for not sufficiently respecting the degree of Internet users' sophistication.[135] One commentator believes that, rather than simply extending trademark law into cyberspace, the ACPA creates a whole new

---

[127] Blasbalg, *supra* note 54, at 595.

[128] *See supra* text accompanying note 107.

[129] Blasbalg, *supra* note 54, at 595.

[130] *See* S. REP. No. 106-140, at 17 (1999) (stating only the following about registrar's policies: "The bill anticipates a reasonable policy against cybersquatting will apply only to marks registered on the Principal Registrar of the Patent and Trademark Office in order to promote objective criteria and predictability in the dispute resolution process."). If anything, the allusion to "objective criteria" implies that the registrant's intent ought not be considered.

[131] *See* Martin, *supra* note 88, at 594–95 ("'Reverse domain name hijacking' occurs where senior trademark users (that is, the person first to use a mark) protest the registration of .com domain names upon discovery that a junior user (a later user of the mark) has already registered the senior user's desired name.").

[132] *See* 15 U.S.C. § 1114(1)(D)(iv) (Supp. V 1999).

[133] Blasbalg, *supra* note 54, at 567.

> The attempts by Congress to fit cybersquatting into the existing framework of Federal trademark law creates a situation whereby a trademark use which might be non-infringing in the everyday world becomes actionable under the ACPA when it is used as a domain name. Consequently, the ACPA has granted trademark owners more control over the use of their marks as domain names than they have for any other use.

Id.

[134] *See id.* at 599.

[135] *See* Martin, *supra* note 88, at 609–10.

set of rights for trademark holders and, in doing so, may stifle commercial and personal free speech.[136]

These criticisms disregard the cautious approach that underlies the drafting of the ACPA. Congress has made clear its hesitance to curb the growth of the Internet:

> The Internet remains a relatively new and exciting medium for communication, electronic commerce, education, entertainment, and countless other yet-to-be-determined uses. It is a global medium whose potential is only just beginning to be understood. Abusive conduct, like cybersquatting, threatens the continued growth and vitality of the Internet as a platform for all these uses. But in seeking to curb such abuses, Congress must not cast its net too broadly or impede the growth of technology, and it must be careful to balance the legitimate interests of Internet users with the other interests sought to be protected.[137]

In light of this statement, it should not be surprising that the ACPA will produce outcomes which favor Internet commerce over other legitimate uses of the Internet, particularly those primarily employed by individuals. Congress seems reluctant to take any steps which might have a chilling effect on the growth of the Internet as a business medium.

This reluctance suggests a response to those using *Avery Dennison*[138] as an illustration of how the ACPA may be overly generous to mark owners. One commentator has argued that the plaintiff, who lost the case under the FTDA, would probably have been successful under the ACPA.[139] This outcome, it is claimed, would not be infringement in the "real world," because defendant used plaintiff's marks for their non-trademark value, in what would traditionally be considered a non-commercial manner.[140] This analysis correctly predicts the case's outcome under the ACPA, but evaluates this outcome improperly. Defendant Sumpton, who registered and sold domain names consisting of common last names,[141] most likely did not intend to ransom www.avery.net and www.dennison.net to plaintiff Avery Dennison, and it would therefore be unwarranted for him to receive a stiff punishment for registering his domain names. However, it does not follow that the domain names should not be transferred to the plaintiff, which, being a business with an identifiable brand name, is more likely than the defendant (or one who buys the domain

---

[136] Blasbalg, *supra* note 54, at 600.

[137] S. REP. No. 106-140, at 8 (1999).

[138] Avery Dennison Corp. v. Sumpton, 189 F.3d 868 (9th Cir. 1999); *supra* text accompanying notes 54–57.

[139] *See* Blasbalg, *supra* note 54, at 582–85.

[140] *Id.* at 584–85.

[141] *See Avery Dennison*, 189 F.3d at 872.

names from the defendant) to use the domain names in a way consistent with Congress's vision of the Internet as a global medium for electronic commerce.

Similarly, Congress's cautious, market-favoring approach explains the ACPA's perfunctory treatment of Internet users' degree of sophistication. The ACPA has been criticized for not adopting a clear stance regarding the gap between those who are proficient in their use of the Internet and those who are not.[142] But the ACPA is primarily concerned with protecting the integrity of brand names on the Internet, and Congress correctly noted that all consumers may have trouble distinguishing between genuine and pirate sites.[143] This is likely to be true even of sophisticated Internet users: a polished site available at ibm.com is likely to be taken for a site actually maintained by IBM. In the absence of evidence to the contrary, a sophisticated Internet user has no reason to suspect that a Web site is not what it claims to be.

The ACPA's hands-off approach to domain name registrars can also be explained in terms of Congress's fundamental goals and concerns. Domain name registrars are the media through which domain names are obtained, and therefore are the gatekeepers of electronic commerce. By limiting the liability of registrars, Congress has turned their role into a purely administrative one: they are not to judge for themselves whether a domain name infringes on another's mark. Rather, they are to assume at the time of registration that the domain name does not infringe. Recall that the ACPA shields registrars from liability when they refuse to register, remove from registration, transfer, temporarily disable, or permanently cancel a domain name in compliance with either a court order under the ACPA or the registrars' own "reasonable policy" prohibiting registration of domain names identical to, confusingly similar to, or dilutive of another's mark.[144] The legislation leaves unsaid how registrars are to adopt their reasonable policies. Concurrent to the passage of the ACPA, the Internet Corporation for Assigned Names and Numbers[145] ("ICANN") developed its own Uniform Dispute Resolution Policy[146] ("UDRP"). The UDRP is a model policy that was intended to be, and in fact has been, adopted by the overwhelming majority of domain name registrars.[147] This

---

[142] *See* Martin, *supra* note 88, at 609–10.

[143] *See* S. REP. No. 106-140, at 13 (1999).

[144] *See* 15 U.S.C. § 1114(1)(D)(ii) (Supp. V 1999).

[145] ICANN is a non-profit corporation to which the federal government assigned the task of centralizing the management of the domain name system. *See* About ICANN, *at* http://www.icann.org/general/abouticann.htm (last updated May 1, 2001); Luke A. Walker, *ICANN's Uniform Domain Name Dispute Resolution Policy*, 15 BERKELEY TECH. L.J. 289 (2000).

[146] *See* Uniform Domain-Name Dispute-Resolution Policy, *at* http://www.icann.org/udrp/udrp.htm (last updated June 17, 2000).

[147] *See id.* ("This policy has been adopted by all accredited domain-name registrars for domain names ending in .com, .net, and .org. It has also been adopted by certain managers of country-code top-level domains (e.g., .nu, .tv, .ws).").

means that both domain name registrars and registrants are afforded a great degree of predictability. The role of individual registrars in resolving disputes is minimized because each one follows the mandates set forth by the ACPA and the UDRP. While it is true that registrars are generally immune from liability, this is only so when they follow a particular set of policies with which every domain name registrant should be familiar. This sort of predictability seems to accord with the ACPA's pro-business policies.

The charge that the ACPA may stifle commercial and personal free speech is a serious one.[148] Unlike the aforementioned criticisms of the ACPA, which call into question its consistency and administrability, and which can be assuaged by pointing to the policies latent within the ACPA, this charge questions those very policies. One commentator writes:

> [The ACPA] grants a trademark owner the right to control the use of the textual identity of the mark itself so long as the offending use is the registration of a domain name that is similar to the textual equivalent of the mark itself. No corresponding right exists in any other context under modern trademark law. This can have the effect of stifling both commercial and personal free speech by restricting a competitor or commentator from using a portion of a trademark within a domain name to help the public locate his commentary or commercial page.[149]

If this statement is true, then the ACPA would abridge the rights of parody and comparative advertising that exist in the world outside the Internet. All that can be said in response is that Congress clearly meant to avoid such a narrowing of well-established rights:

> [T]he bill does not extent [sic] to innocent domain name registrations by those who are unaware of another's use of the name, or even to someone who is aware of the trademark status of the name but registers a domain name containing the mark for any reason other than with a bad faith intent to profit from the goodwill associated with that mark.[150]

Presumably, the passage of the ACPA will not prevent anyone from registering, say, microsoftsucks.com, or from otherwise incorporating a trademark into a domain name used for any reason other than generating profit at the expense of an established mark. Case law has not yet firmly

---

[148] *See* Blasbalg, *supra* note 54, at 600.
[149] *Id.*
[150] S. REP. No. 106-140, at 12 (1999).

established a rule on this matter, although it has arisen in two cases which, read together, suggest that as long as the domain name is clearly parodic, its use falls outside categories of cybersquatting proscribed by the ACPA. In *Lucent v. Lucentsucks.com*, the Eastern District of Virginia suggested that, in the absence of a likelihood of confusion and bad faith, Lucentsucks.com could not be enjoined under the ACPA.[151] However, because plaintiff's claim was dismissed for failure to satisfy the elements required of in rem jurisdiction,[152] the portion of the opinion suggesting the legitimacy of parodic use of a mark within a domain name is dictum. In *Morrison & Foerster v. Wick*, the District of Colorado ordered defendant to forfeit his interests in www.morrisonfoerster.com, www.morrisonand foerster.com, www.morrisonforester.com, and www.morrisonandforester. com despite his claims that, because the content of his Web sites was parodic, he was protected by the First Amendment.[153] The court held that because these domain names were inherently confusing, a user would need to actually explore the sites in order to determine that they were parodies. Thus, the defendant could not successfully argue that the domain names themselves were parodies entitled to First Amendment protection.[154]

A lesson drawn from these two cases is that it is important to maintain an analytical distinction between domain names and Web sites. A parodic Web site maintained without bad faith may still use a domain name that the ACPA would otherwise (appropriately) prohibit. A critique or parody of Microsoft available at microsoftsucks.com should be permitted because no reasonable user would think that the site was maintained by Microsoft; such a site available at Microsoft.com should probably not be permitted. Thus far, courts seem to understand and be capable of implementing this distinction.

Remedies available to victims of cybersquatting under the Lanham Act and the Federal Trademark Dilution Act were inconsistent and often at odds with the intent of those acts. Because cybersquatters could, in theory and in practice, avoid the "famousness" and "use in commerce" requirements of the FTDA, a new weapon was needed to combat the use of trademarks as internet Web site domain names. The Anticybersquatting Consumer Protection Act has provided cybersquatting victims with such a weapon. The ACPA does what it sets out to do quite effectively. The essence of the criticisms of the ACPA is that it does not go far enough. Congress, however, has chosen to proceed cautiously, preferring to wait and see how the Internet develops and how case law applies the ACPA, and to encourage registrants and registrars to resolve disputes by

---

[151] 95 F. Supp. 2d 528, 535–36 (E.D. Va. 2000).
[152] *Id.* at 531–34.
[153] 94 F. Supp. 2d 1125, 1134–35 (D. Colo. 2000).
[154] *Id.*

means other than litigation before it regulates with a heavy hand. This seems to be the most prudent approach, and the one that best accords with Congress's express goals of protecting consumers and American businesses, promoting the growth of online commerce, and providing clarity in the law for trademark owners.[155] The ACPA is hardly revolutionary; it merely extends into cyberspace the trademark protections necessary for businesses to develop and profit from their distinctive brand names. The worst that can be said is that it does no more.

—Allon Lifshitz

[155] S. REP. No. 106-140, at 4 (1999).

RACIAL PROFILING

On April 15, 1999,[1] Representative John Conyers (D-Mich.), the ranking Democrat on the House Judiciary Committee, introduced House Bill 1443, the Traffic Stops Statistics Study Act (the "Act"),[2] a bill designed to initiate the gathering of comprehensive data about the racial distribution of police traffic stops. This bill, along with its identical Senate counterpart introduced by then-Senator Frank Lautenberg (D-N.J.),[3] is a comprehensive attempt to address racial profiling via national legislation. The bill failed to obtain passage in the 106th Congress and, given partisan division over the issue as well as uncertainty about the bill's efficacy, the outcome of any future congressional efforts to address racial profiling remains in doubt. Regardless of whether the Act ultimately achieves passage, in the absence of more concrete measures, this legislation will be more symbolic than substantive.

The political impetus behind House Bill 1443 stems from the national publicity surrounding racial profiling. On April 23, 1998, two New Jersey state troopers conducting a traffic stop on the New Jersey Turnpike pulled over a van containing four occupants, three of them African Americans, who were on their way to a basketball camp in North Carolina.[4] A few minutes after the stop, the officers fired shots, wounding three of the motorists.[5] Given the races of the victims and the absence of any evidence of wrongdoing on their part, many suspected that the police officers had been motivated by racial bias.[6] A subsequent study indicated that police stopped African American motorists in New Jersey to a disproportionate degree, corroborating the claim that the state's law enforcement officials engage in racial profiling.[7] The release of a photograph of then-Governor Christine Todd Whitman (R-N.J.) frisking a black motorist while accompanying police officers on a patrol exacerbated public criticism.[8]

---

[1] 145 CONG. REC. E673 (weekly ed. Apr. 15, 1999) (statement of Rep. Conyers).

[2] Traffic Stops Statistic Study Act of 1999, H.R. 1443, 106th Cong. (1999).

[3] Traffic Stops Statistic Study Act of 1999, S. 821, 106th Cong. (1999).

[4] Jeffrey Gold, *State Police Deny Race is a Factor in Pulling Over Motorists*, REC. (BERGEN COUNTY), May 2, 1998, at A4.

[5] *Id.*

[6] *See id.* In February, the victims received $13 million in a settlement with New Jersey officials. Iver Peterson, *New Jersey Agrees to Pay $13 Million in Profiling Suit*, N.Y. TIMES, Feb. 3, 2001, at A1.

[7] PETER VERNIERO & PAUL H. ZOUBEK, INTERIM REPORT OF THE STATE POLICE REVIEW TEAM REGARDING ALLEGATIONS OF RACIAL PROFILING (1999). In his presidential debate with former Senator Bill Bradley (D-N.J), Vice President Gore quipped, "Racial profiling practically began in New Jersey, Senator Bradley." Katharine Q. Seelye, *Gore's Image, Focused and Relentless*, N.Y. TIMES, Oct. 2, 2000, at A23.

[8] David Kocieniewski, *Frisking Photo Puts Whitman On Defensive*, N.Y. TIMES, July 11, 2000, at B1. Deborah Jacobs, executive director of the New Jersey chapter of the American Civil Liberties Union, claimed that the frisk was also illegal because the troopers had already searched the man, finding no contraband, before he was turned over to Whit-

In response to the publicity surrounding the incidents in New Jersey and other states,[9] several state and local governments commissioned studies on racial profiling. Studies conducted in New York City,[10] Maryland,[11] Dallas,[12] and Los Angeles[13] all revealed evidence that law enforcement officers stop African American motorists at a disproportionate rate. In response to these findings, the United States Department of Justice filed a number of suits, resulting in settlements in which various police departments agreed to reform their practices.[14] Twenty-three of the nation's fifty largest cities, including Miami, San Diego, Philadelphia, Seattle, and Houston, have made progress in developing or implementing programs that seek to eliminate the use of racial profiling in road stops.[15] Additionally, several states have enacted legislation to provide for additional studies or reform law enforcement practices.[16] In 1999, President Clinton directed federal agencies to develop a plan to collect data on the race, gender, and ethnicity of persons stopped by agency officers.[17]

In an attempt to provide for more systematic research on the issue, Representative Conyers first introduced the Traffic Stops Statistics Study Act in 1997.[18] The bill passed the House of Representatives the following year,[19] but the Senate declined to vote on it.[20] The new version, intro-

---

man. According to Phil Moran, the lawyer who subpoenaed the Whitman photos, a state police supervisor offered an extra week of paid vacation to the troopers escorting Whitman if they brought back a photo of the Governor frisking a black suspect. *Id.*

[9] *See, e.g.*, Paul Zielbauer, *Blacks in New Haven Cite Racial Profiling by Police in Nearby Towns*, N.Y. TIMES, July 17, 2000, at B1; Tina Kelley, *Call for Calm After Shooting of Policeman by Colleagues*, N.Y. TIMES, Jan. 30, 2000, § 1, at 14.

[10] Benjamin Weiser, *Prosecutors See Profiling by New York Police*, N.Y. TIMES, Oct. 8, 2000, § 4, at 2.

[11] Julian Bond & Wade Henderson, Editorial, *The Bias the Candidates Deplore*, N.Y. TIMES, Oct. 13, 2000, at A33.

[12] Christy Hoppe, *Race Disparity Found in Traffic Stops: Blacks Get More Tickets in Some Counties; DPS Says Study Flawed*, DALLAS MORNING NEWS, Oct. 4, 2000, at 1A.

[13] Letter from Bill Lann Lee, Acting Assistant Attorney General, Civil Rights Division, to James K. Hahn, City Attorney, City of Los Angeles (May 8, 2000), *available at* http://www.usdoj.gov/crt/split/ documents/lapdnoti.htm.

[14] In an agreement with the Justice Department, the New Jersey State Police Department agreed to 97 specific changes in policies and rules. In October 2000, two monitors overseeing implementation of the agreement commended the police force for its first steps toward compliance. Robert Hanley, *Monitors Commend Police on Effort to End Profiling*, N.Y. TIMES, Oct. 7, 2000, at B5. Los Angeles city officials negotiated an agreement with the Justice Department in which the police department will collect data on the race of those stopped by police. Benjamin Weiser, *U.S. Detects Bias in Police Searches*, N.Y. TIMES, Oct. 5, 2000, at A1.

[15] Eric Lipton, *Police Report Rekindles the Mayor's Fiery Side*, N.Y. TIMES, Oct. 7, 2000, at B3.

[16] *See infra* notes 103–117.

[17] Memorandum on Fairness in Law Enforcement, 35 WEEKLY COMP. PRES. DOC. 1067 (June 9, 1999).

[18] *See* 143 CONG. REC. E10 (daily ed. Jan. 7, 1997) (statement of Rep. Conyers); Traffic Stops Statistics Study Act of 1997, H.R. 118, 105th Congress (1997).

[19] 144 CONG. REC. H1389 (daily ed. Mar. 24, 1998).

[20] Lori Litchman, *Philadelphia Statistics on Racial Profiling Buttress Findings of National ACLU*, LEGAL INTELLIGENCER, June 3, 1999, at 3.

duced in 1999,[21] retained most of the original bill's provisions and also required police to record the driver's gender and immigration status.[22] The House Judiciary Committee unanimously passed the new bill March 13, 2000,[23] but thereafter the bill stalled. In the Senate, the bill was referred to the Judiciary Committee on April 15, 1999,[24] where it remained for the rest of the legislative session.

Recognizing the bill's dim chance of passage in the Senate in the fall of 2000, the American Civil Liberties Union ("ACLU") launched a paid advertising campaign in support of the bill, specifically targeting three Republican members of the Senate Judiciary Committee: Senators Spencer Abrams (R-Mich.), John Ashcroft (R-Mo.), and Orrin Hatch (R-Utah).[25] The magazine ads sponsored by the ACLU featured Dr. Martin Luther King, Jr. on the left side and Charles Manson on the right, with a caption above the pictures reading, "The man on the left is 75 times more likely to be stopped by the police while driving than the man on the right."[26] On August 26, 2000, two days before the thirty-seventh anniversary of Dr. King's "I Have a Dream Speech," his son, Martin Luther King III, and the Reverend Al Sharpton led a march on Washington to call on Congress and the White House to address the problem of racial profiling.[27] Though the bill again failed to become law during the 106th Congress, Representative Conyers plans to reintroduce the bill later this year.[28]

House Bill 1443 would require the Attorney General to review the existing data on racial profiling, particularly complaints alleging racial bias in traffic stops.[29] The bill then directs him to collect data from a nationwide sample.[30] The data collected would include the traffic violation prompting the stop; the driver's identifying characteristics, including race, gender, ethnicity, and approximate age; and whether the driver's

---

[21] 145 CONG. REC. E673 (weekly ed. Apr. 15, 1999) (statement of Rep. Conyers).

[22] *See* Traffic Stops Statistics Study Act of 1999, H.R. 1443, 106th Cong. § (2)(a)(3) (1999).

[23] 146 CONG. REC. H930 (daily ed. Mar. 13, 2000).

[24] 145 CONG. REC. S3776 (daily ed. Apr. 15, 1999).

[25] William Raspberry, Editorial, *Racial Profiling Bill Stalls*, DALLAS MORNING NEWS, Sept. 19, 2000, at 15A.

[26] Patricia Winters Lauro, *The A.C.L.U. is Taking a Provocative Madison Avenue Route to Raise Support for its Causes*, N.Y. TIMES, May 30, 2000, at C10. In addition to its magazine ads featuring King and Manson, the ACLU initiated a series of radio and television public service announcements in both English and Spanish. The ads feature reality-based enactments that communicate the humiliation felt when innocent motorists are stopped by the police and offer a toll-free complaint hotline for victims to call. Press Release, American Civil Liberties Union, ACLU Joins Protest Against Racial Bias and Police Brutality at D.C. "Redeem the Dream" Rally (Aug. 25, 2000), *available at* http://aclu.org/news/2000/n082500b.html.

[27] Cindy Loose & Chris L. Jenkins, *Rallying to "Redeem the Dream": Rights Leaders Target Racial Profiling*, WASH. POST, Aug. 27, 2000, at C01.

[28] Telephone Interview with Cynthia Martin, Press Secretary, Office of Representative John Conyers (Feb. 1, 2001).

[29] Traffic Stops Statistics Study Act of 2000, H.R. 1443, 106th Cong. § 2 (2000).

[30] *Id.*

immigration status was questioned.[31] Police would also document the number of people in the vehicle, whether a search resulted from the stop, whether they requested consent to search the vehicle, and whether any alleged criminal activity by the driver justified the search.[32] Furthermore, the bill would require the officer to report whether he gave a warning or citation due to the stop, the justification for any arrest made, and the duration of the stop.[33] All data would be collected by law enforcement officials, and the Attorney General would have the option to provide grant money to facilitate such information gathering.[34]

Supporters of the Act argue that gathering comprehensive data will permit assessment of the extent and scope of racial profiling.[35] They contend that a national effort will determine conclusively whether the problem exists and limit advocates' reliance on anecdotal evidence.[36] A study by the General Accounting Office found that there currently exists "no comprehensive, nationwide source of information that could be used to determine whether race has been a key factor in motorist stops."[37] Given the "active hostility" towards record keeping in the law enforcement community,[38] a mandate may be required to address the shortfall of information on the issue.

Proponents argue that racial profiling merits statistical study for several reasons. On the most basic level, Professor Randall Kennedy of Harvard Law School argues that any distinctions based on race are suspect and should be treated as such. He explains:

> In America, the making of racial distinctions has proven to be more destructive and more popularly distasteful than other lines of social stratification .... When officials discriminate on racial grounds, judges have typically demanded "strict scrutiny"—the most intense level of judicial review .... Strict scrutiny embodies a recognition, born of long and terrible experience, that the presence of a racial factor in decision making should raise anxiety and signal that the government is likely to be doing something wrong.[39]

---

[31] *Id.*

[32] *Id.*

[33] *Id.*

[34] *Id.* § 3.

[35] *See* David A. Harris, *The Stories, the Statistics, and the Law: Why "Driving While Black" Matters* [hereinafter *The Stories, the Statistics, and the Law*], 84 MINN. L. REV. 265, 320 (1999).

[36] *The Stories, the Statistics, and the Law*, *supra* note 35, at 320.

[37] U.S. GEN. ACCT. OFF., RACIAL PROFILING: LIMITED DATA AVAILABLE ON MOTORIST STOPS 1 (2000).

[38] *The Stories, the Statistics, and the Law*, *supra* note 35, at 276.

[39] RANDALL KENNEDY, RACE, CRIME, AND THE LAW 146–47 (1997).

Another concern voiced by proponents of the bill is that the prevailing perception of widespread racial profiling creates tension between African Americans and police officers.[40] Eleanor Holmes Norton (D-D.C.), the District of Columbia's non-voting delegate to the House of Representatives, notes:

> [This tension] is most interesting because the African American community has embraced police because there was such high crime, especially in the 1990's. Crime is down 10 percent now from last year, 34 percent over the last few years; and yet there is this intense hostility based on what is happening particularly to black men but also to black women.[41]

As a result of this hostility, African Americans are less likely to aid the police in their investigations and African American jurors are less likely to believe the testimony of police officers in court.[42] As Professor David A. Harris of the University of Toledo College of Law observes, "One need only think of the split screen television images that followed the acquittal in the O.J. Simpson case[:] stunned, disbelieving whites, juxtaposed with jubilant blacks literally jumping for joy—to understand how deep these divisions are."[43] Because African Americans are also disproportionately likely to be victims of crime, this reluctance to cooperate with authorities creates a perverse effect.[44]

The debate on racial profiling has also been influenced by a Supreme Court decision unanimously holding that traffic stops do not violate the

---

[40] Interview data indicate that African Americans strongly believe that officers subject them to stops and tickets more frequently than they do whites. *The Stories, the Statistics, and the Law*, *supra* note 35, at 267. Christopher Darden, one of the O.J. Simpson prosecutors and an African American, notes that he "learned the rules of the game years before .... Don't move. Don't turn around. Don't give some rookie an excuse to shoot you." *Id.* at 275.

[41] 146 CONG. REC. H7545 (statement of Del. Eleanor Holmes Norton).

[42] *The Stories, the Statistics, and the Law*, *supra* note 35, at 268–69.

[43] *Data Collection: The First Step in Coming to Grips with Racial Profiling: Hearing on S.B. 821 Before the Subcomm. on the Constitution, Federalism, and Property Rights of the Senate Judiciary Comm.*, 106th Cong. (2000) [hereinafter *Data Collection*] (prepared testimony of Prof. David A. Harris), *available at* http://www.senate.gov/~judiciary/330200dah.htm.

[44] *The Stories, the Statistics, and the Law*, *supra* note 35, at 290–91. Some argue that a fear of racial profiling has created opposition to campaigns to provide for safer driving. Katheryn K. Russell, *"Driving While Black": Corollary Phenomena and Collateral Consequences*, 40 B.C. L. REV. 717, 729 (1999). For instance, in January 2000, when the New Jersey State Legislature passed a measure allowing police officers to stop drivers solely for not wearing seat belts, opponents of the bill voiced the concern that the bill would provide greater opportunity for racial profiling. *Bill Would Allow the Police to Stop Drivers Just to Check for Seat Belt Use*, N.Y. TIMES, Jan. 12, 2000, at B5. In the fall of 1998, the National Urban League ("NUL") withdrew its support for President Clinton's "Buckle Up America" campaign, which would have made failure to wear a seatbelt grounds for police officers to pull over and ticket motorists, due to concerns that it would lead to more racial profiling. Russell, *supra* note 44, at 729.

Fourth Amendment protections against search and seizure so long as the stops meet constitutional standards of reasonableness, regardless of the actual motivations of the individual officers involved.[45] The 1996 case of *Whren v. United States* severely impedes the ability of those pulled over on a seemingly pretexual basis to challenge profiling on a case-by-case basis.[46] In *Whren*, the Court held that law enforcement officers may conduct a search if a motorist violates any traffic law.[47] Given that virtually every driver regularly violates one minor traffic law or another, police officers have nearly unlimited discretion to stop motorists.[48] Rejecting racial profiling as grounds for invalidating a search, the Court stated that "the constitutional basis for objecting to intentionally discriminatory application of laws is the Equal Protection Clause, not the Fourth Amendment."[49]

While police departments have, for the most part, opposed legislation addressing racial profiling,[50] supporters of the bill argue that it could actually improve the effectiveness of law enforcement. As Rachel King, legislative counsel for the ACLU, explains, police spend a great deal of time pulling over drivers even when there are no charges to file against them.[51] In Florida, for instance, only ten percent of the thirty-two to thirty-five million traffic stops each year result in tickets.[52] According to King, "you must have a lot of time on your hands when [a large percentage] of the stops you do result in nothing."[53] Furthermore, if the bill alters

---

[45] Whren v. United States, 517 U.S. 806, 813 (1996).

[46] Russell, *supra* note 44, at 718.

[47] *Whren*, 517 U.S. 806, at 813.

[48] David A. Harris, *"Driving While Black" and All Other Traffic Offenses: The Supreme Court and Pretextual Traffic Stops*, 87 J. CRIM. L. & CRIMINOLOGY 544, 545–46 (1997) [hereinafter "Driving While Black"]. Harris notes:

> In Utah, drivers must signal for at least three seconds before changing lanes; a two second signal would violate the law. In many states, a driver must signal for at least one hundred feet before turning right; ninety-five feet would make the driver an offender. And the driver making that right turn may not slow down "suddenly" without signaling. Many states have made it a crime to drive with a malfunctioning taillight, a rear-tag illumination bulb that does not work, or tires without sufficient tread. They also require drivers to display not only license tags, but yearly validation stickers, pollution control stickers, and safety inspection stickers; driving without these items displayed on the vehicles in the proper place violates the law.

Id. at 558–59. However, in the post-*Whren* case of *United States v. Leviner*, 31 F. Supp. 2d 23, 24 (D. Mass 1998), the trial judge departed from the United States Sentencing Commission Guidelines and reduced a repeat offender's sentence based on a conclusion that the defendant's previous convictions for motor vehicle-related crimes were likely the result of racial profiling. *See* Russell, *supra* note 44, at 728.

[49] *Whren*, 517 U.S. at 813.

[50] *See infra*, notes 83–102 and accompanying text.

[51] Tamara Lytle, *Initiatives Would Track Race in Traffic Stops*, ORLANDO SENTINEL, Apr. 25, 1999, at A1.

[52] *Id.*

[53] *Id.*

police practices, relationships between law enforcement officers and minority communities may improve.[54] Data from San Diego and Boston indicate that less confrontational, more community-oriented policing strategies have successfully reduced crime by making local residents less suspicious of the police, leading to more cooperation and information-sharing between residents and law enforcement.[55]

Proponents further argue that the bill will not significantly increase the administrative burden already placed on police.[56] Many police departments require police officers to fill out a report when making stops,[57] and many police cars have computer terminals that would facilitate entering the data.[58] Moreover, some police departments have already installed video cameras in their cars to record stops, which could facilitate research on racial profiling.[59]

Supporters of the Act also note that it will provide statistics on the number of innocent motorists stopped.[60] They argue that current estimates on the prevalence of racial profiling may actually understate the problem, because for the most part, researchers only learn of incidents of racial profiling when innocent parties take legal action, which rarely occurs unless the officers involved behaved violently.[61] A comprehensive study would bring to light not just the egregious cases of racial profiling that are reported as a result of legal action, but also subtle, day-to-day harassment.

Proponents of the bill stress that such racially motivated, harassing stops of innocent people can be more than mere inconveniences. For instance, research indicated that in 1992 most drivers stopped in Volusia County, Florida on Interstate 95 by the Sheriff's highway drug squad were minorities and that deputies seized cash from many of them, even when no charges were filed.[62] Even more seriously, Kennedy argues, racially motivated stops involving innocent motorists may escalate to violence:

> The people stopped will vent their resentment. The officer—and recall that we are here talking about the initially good, nonracist, courteous officer—will respond in some defensive manner, which will in turn provoke further negative responses from

---

[54] *See Data Collection*, *supra* note 43 (prepared testimony of Prof. Harris).
[55] *See id.* (prepared testimony of Prof. Harris).
[56] *See "Driving While Black,"* *supra* note 48, at 581–82.
[57] *See id.* at 581–82.
[58] *See id.* at 581.
[59] *See id.* at 582.
[60] KENNEDY, *supra* note 39, at 155.
[61] *Id.* at 155.
[62] Lytle, *supra* note 51. According to videotape obtained by the *Orlando Sentinel*, deputies seized money almost three times as often as they arrested anyone on drug charges. *"Driving While Black,"* *supra* note 48, at 562.

> those who feel aggrieved. That, in turn, will further aggravate the officer, leading to a deteriorating relationship that will often create bruised feelings, sometimes generate needless arrests, and occasionally spark violence.[63]

Finally, the fear of being stopped often leads minorities to alter their driving patterns, go out less often, and even compute traffic stops into their travel time.[64]

Some supporters of the Act believe it does not go far enough. Because substantial evidence already suggests that racial profiling is a problem, they argue, the study authorized by the bill will incur much time and expense confirming what many studies already show.[65] The passage of the bill and the subsequent long and perhaps expensive study, they fear, could serve to preclude the enactment of stronger, more substantive measures, as lawmakers and police personnel could point to the bill as evidence that they are addressing racial profiling even as the problem persists.[66]

They also fault the bill for requiring police officers to include only a limited amount of information. It does not require that any data be recorded about the police officer making the stop, such as his age or race.[67] The Act also does not call for information about the location of the traffic stop, precluding an analysis of cities and states particularly unfriendly to minority drivers.[68]

Many fear that the information gathered will fail to conclusively prove the existence of racial profiling and will instead lead to prolonged debates over the proper interpretation of that information. For instance, the Texas Department of Public Safety immediately disputed quantitative studies that pointed to a racial profiling problem.[69] The department claimed that the studies erroneously compared the racial composition of drivers pulled over in a locality to the racial composition of that locality,

---

[63] KENNEDY, *supra* note 39, at 157.

[64] *See "Driving While Black," supra* note 48, at 570–71.

[65] Recent surveys illustrate that from 1995 to 1997, 70% of the drivers pulled over by the Maryland State Police on Interstate 95 were African American, despite the fact that blacks constituted only 17.5% of all drivers. Bond & Henderson, *supra* note 11. Until 1999, the New Jersey State Police denied that state troopers used racial profiling. However, as early as 1996, the agency's internal audits revealed widespread racial profiling along the New Jersey Turnpike. Newly released documents indicate that senior commanders declined to take aggressive steps to combat the problem and instead tried to withhold information from federal civil rights prosecutors. David Barstow & David Kocieniewski, *Records Show New Jersey Police Withheld Data on Race Profiling*, N.Y. TIMES, Oct. 12, 2000, at A1.

[66] *See* Barstow & Kocieniewski, *supra* note 65.

[67] *See* Russell, *supra* note 44, at 727. The anonymity of officers, however, may be necessary to ensure honesty in reporting the data. *See "Driving While Black," supra* note 48, at 580.

[68] Russell, *supra* note 44, at 727.

[69] Jim Yardley, *Studies Find Race Disparities in Texas Traffic Stops*, N.Y. TIMES, Oct. 7, 2000, at A12.

thereby ignoring the possibility that many of the drivers might reside elsewhere.[70] Likewise, New York City Mayor Rudolph Giuliani vehemently disputed that a study revealing that blacks and Hispanics were stopped and frisked disproportionately to the arrest rate for these groups proved the existence of racial profiling.[71] He argued that because crime victims more frequently identify African Americans and Latinos as suspects, police are more likely to target these groups for searches.[72]

The bill may also fall short of proponents' hopes regarding its ability to aid in civil rights litigation. In the past, plaintiffs have found it difficult to prevail on equal protection claims based on statistical evidence. While the studies authorized by the bill would provide a plethora of such evidence, the bill does not define what statistical threshold would support a legal discrimination claim, thereby leaving present legal standards unaltered.[73] As a result, even if the data collected appeared to offer proof of racial profiling for political and public relations purposes, it might still fail to meet the high statistical threshold necessary to prevail on an equal protection claim.

Another concern about the bill is that it fails to incorporate proactive proposals for the purpose of addressing racial profiling. In its report on the profiling practices of the New York City Police Department, the United States Civil Rights Commission recommended several remedial steps, including increasing recruitment of minority police officers, attempting to improve relations between police officers and members of the communities they patrol, and abolishing the so-called "forty-eight hour rule," which prevents superiors from interviewing officers about the

---

[70] *Id.*

[71] *See* Lipton, *supra* note 15.

[72] *See id.* The fact that less than one-third of stops occurred because the police were looking for a particular suspect somewhat undermines the Mayor's contention. *Id.* Other studies similarly indicate the difficulties involved in interpreting statistics. In Connecticut, while a six-month analysis of traffic stops by state and local police concluded that "minority drivers do not appear to be systematically treated differently than nonminority drivers" in terms of the rate at which different races were stopped compared to the racial makeup of the locality, it did show significant differences in the number of misdemeanor summonses issued after a stop. The study further noted the obvious point that "we cannot definitively conclude that individual police officers do not practice racial profiling." STEPHEN M. COX, OFF. OF THE CHIEF STATE'S ATT'Y, STATE OF CONN., INTERIM REPORT OF TRAFFIC STOPS STATISTICS FOR THE STATE OF CONNECTICUT ii (2001), *available at* http://web.wtnh.com/Report_Narrative.pdf. The Connecticut study indicates a further problem surrounding the use of statistical surveys. Carolyn Nah, president of the Bridgeport, Connecticut chapter of the NAACP, responded to the study by vowing to recommend that her organization perform its own statistical study, claiming, "I know it's a problem .... It's not just a statewide problem, it's a national problem. I don't care what the study shows." Shaila K. Dewan, *State Study in Connecticut Finds No Pattern of Racial Profiling*, N.Y. TIMES, Jan. 25, 2001, at B5. A six-month study in New Jersey similarly delivered mixed results, indicating that while stop rates for African Americans had increased, arrest rates had declined. Iver Peterson, *Racial Math on Turnpike: More Stops but Fewer Arrests*, N.Y. TIMES, Jan. 14, 2001, § 1, at 36.

[73] David Crump, *Evidence, Race, Intent, and Evil: The Paradox of Purposelessness in the Constitutional Racial Discrimination Cases*, 27 HOFSTRA L. REV. 285, 322 (1998).

use of abusive tactics within two days of the incident.[74] The Commission also advocates increasing the college-education requirement for officers from two to four years, on the theory that more educated officers are less likely to engage in misconduct.[75] Most notably, it calls for enhanced diversity training.[76] Representative Asa Hutchinson (R-Ark.), a member of the House Judiciary Committee, argues that Congress should improve training for law enforcement officers, providing them with race-neutral factors to be used in deciding whether to make an investigatory stop. He also argues that Congress should pass legislation preventing police from using race as a criterion for making stops, similar to the current bar on the use of preemptive strikes against jurors for racial reasons.[77] In addition to changes in training, more direct steps are available. New Jersey, for instance, has instituted a campaign to install video cameras on all patrol cars so that traffic stops can be recorded and monitored.[78] In short, because alternative solutions exist, many of them seemingly more substantive than merely taking more surveys, the Act's continued reliance on the gathering of statistics arguably sells the problem short.

Finally, many argue that "driving while black" is symptomatic of a broader problem of police conduct toward racial minorities, and, by focusing only on driving, the bill ignores racial profiling in a host of other settings.[79] In New York City, the shooting death of Amadou Diallo represents one of the most high-profile examples of police shootings attributed to racial profiling.[80] Since Diallo was not a motorist, however, incidents like his would not be included in the statistics provided for by this bill. Conyers remarked, "It's not just a matter of driving while black. It's a matter of shopping while black, living while black."[81] Similarly, the controversy extends beyond just African Americans. Many American Muslims, for instance, have complained about being targeted by airport security guards, a possible form of racial profiling that does not receive as

---

[74] U.S. COMM'N ON CIVIL RIGHTS, POLICE PRACTICES AND CIVIL RIGHTS IN NEW YORK CITY iv (2000).

[75] *Id.* at 15.

[76] *Id.* at iv.

[77] Asa Hutchinson, Editorial, *What Actions Should Congress Take to Prevent Racial Profiling? Racial Profiling Endangers Justice*, ROLL CALL, Feb. 7, 2000, § 1, at 9.

[78] David Kocieniewski, *After Profiling Scandal, Tough Choices for New Jersey Police Leader*, N.Y. TIMES, Mar. 5, 2000, at 35.

[79] *See generally* Russell, *supra* note 44, at 721–25.

[80] Weiser, *supra* note 14.

[81] Nichole M. Christian, *Reno is Asked to Investigate Death of Black Man at Mall*, N.Y. TIMES, July 7, 2000, at A10. In Michigan, on June 22, 2000, Frederick Finley, an African American man, was choked to death in a shopping mall parking lot during an altercation with security guards trying to seize a $4 bracelet from his stepdaughter, whom they suspected had shoplifted it. Representative Conyers believes that racial profiling played a role and asked then-Attorney General Janet Reno to investigate the incident. *Id.* Federal prosecutors are currently investigating whether New York Police Department officers engage in racially discriminatory stop-and-frisk tactics. William K. Rashbaum, *U.S. Says City Has Failed to Release Data on Frisks*, N.Y. TIMES, Jan. 31, 2001, at B4.

much attention and is not addressed by the bill.[82]Opponents of the bill, on the other hand, believe the legislation goes too far, either because they deny the existence of a racial profiling problem or because they believe that the bill represents unwarranted federal interference into the activities of police departments. Critics who take the first point of view offer several arguments that call into question the ability of statistics to confirm the existence of racial profiling. First, law-enforcement agencies often cite the fact that the crime rate for blacks is higher than that for whites, which they argue justifies statistical disparities in stops.[83] Even Professor Kennedy acknowledges that "blacks, particularly young black men, commit a percentage of the nation's street crime that is strikingly disproportionate to their percentage in the nation's population."[84] Professor David Crump of the University of Houston Law Center describes the use of the higher black crime rate as an explanation for the disparity in traffic stops:

> African-Americans . . . are economically less fortunate as a group than members of other races. This . . . may mean that . . . they are less likely to have the wherewithal to afford up-to-the-minute maintenance of their automotive equipment, such as brakes, taillights and mufflers, and thus they may be subjected to a race-neutrally higher incidence of that percentage of traffic stops that flows from the resulting kinds of infractions. Unless we examined the percentages of stops attributable to this kind of offense in each racial category, we could not definitively answer this argument of opponents, and indeed we could not do so con-

---

[82] Caryle Murphy, *Muslims See New Clouds of Suspicion; Mideast Backlash Cited As Ramadan Fasts Begin*, WASH. POST, Nov. 27, 2000, at B1.

[83] Jeffrey Prescott, Editorial, *New Facts on Racial Profiling*, CHRISTIAN SCI. MONITOR, May 10, 2000, at 8. Of course, this argument begs the question of whether statistics showing a higher than average crime rate amongst African Americans might in fact be the result of racial profiling. Professor Harris explains, for example, that:

> [D]rug offenses are much less likely to be reported, since possessors, buyers, and sellers of narcotics are all willing participants in these crimes . . . . [This means that arrest] data do not measure the extent of drug crimes. Rather, they measure law enforcement activity and the policy choices of . . . the criminal justice system.

Data Collection, *supra* note 43 (prepared testimony of Prof. Harris). In Maryland, a study indicated that drugs were found at an equal rate between blacks and whites when vehicles were stopped and searched. Custom Service statistics similarly reveal that while over 43% of those searched were either black or Hispanic, the rate of drugs found was 6.7% for whites, 6.3% for blacks, and 2.8% for Hispanics. It is also the case that most drug users are white, and that most users buy drugs from those of their own race. *Id.*

[84] KENNEDY, *supra* note 39, at 145. In 1993, the Reverend Jesse Jackson went so far as to admit, "There is nothing more painful to me at this stage in my life than to walk down the street and hear footsteps and start thinking about robbery—then look around and see somebody white and feel relieved." Stuart Taylor, Jr., *Cabbies, Cops, Pizza Deliveries, and Racial Profiling*, 2000 NAT'L. J. 1891, 1892.

> vincingly even then. These hypothetical counterarguments, based as they would be upon cultural and economic generalizations about African-Americans, would sound offensive. They unquestionably would partake of stereoptyping. But they would furnish an alternative inference that opponents of the discrimination hypothesis should be expected to advance, and in reality, the opponents legitimately could argue that such arguments are no more stereotypical than the original inference of racism. From the point of view of the honest state trooper, the inference of discrimination in traffic stops from racial impact may be the aspect of the argument that is most stereotypical and offensive.[85]

Indeed, from the police perspective, commissioning these studies would create different standards for when to pull over blacks versus whites—in essence, a reverse racial profiling.[86]

Some critics question the accuracy of statistics in light of other demographic variables. For instance, San Diego officials, in response to data that indicated that police were more likely to stop blacks and Hispanics than whites and Asian Americans, pointed out that census figures used as a baseline in the study tend to underreport minorities, thereby potentially creating the appearance of disproportionate statistics when, in fact, the total population figures for some groups may be higher than reported.[87]

Law enforcement officials also fear that the surveys mandated by the bill will strain the limited resources of police departments. There are already significant disputes as to the cost of these measures.[88] The Florida Department of Law Enforcement estimates that collecting the required information on a statewide basis would cost $8 million per year.[89] On the other hand, California officials in 1998 estimated it would cost only $1 million a year to gather the data in their much larger state.[90]

Another concern regards the potential for use of these statistics in litigation. In November 2000, the United States Commission on Civil Rights pointed out that access to data on racial profiling would make it easier to bring suits against officers alleged to have engaged in abusive practices.[91] In the context of a criminal trial or civil rights suit, use of this

---

[85] Crump, *supra* note 73, at 323–24.

[86] Robert T. Scully, Editorial, *Police Not Guilty of Racial Profiling*, WASH. TIMES, June 14, 1999, at A17.

[87] Barbara Whitaker, *San Diego Police Found to Stop Black and Latino Drivers Most*, N.Y. TIMES, Oct. 1, 2000, § 1, at 31. Furthermore, the study would not account for the many Mexicans that drive into San Diego. *Id.*

[88] *See* Lytle, *supra* note 51.

[89] *Id.*

[90] *Id.*

[91] *Panel Urges Remedies to Abuses by Police*, N.Y. TIMES, Nov. 4, 2000, at A18.

data would likely be more persuasive than anecdotal evidence.[92] Data collected pursuant to the Act may also provide evidence that would support suits by the Department of Justice against police precincts that use abusive tactics.[93] While the version of the bill introduced in 1998 prevented the use of the data in litigation,[94] the 1999 bill would not have prevented such use.[95] Opponents of the bill fear that this omission would force police departments to defend themselves against a wave of new lawsuits by making the data available to "the cottage industry of lawyers who make their living suing police officers across the country."[96]

Police advocates also fear that the actual collection of the data could have consequences that would make traffic stops at best awkward and at worst even more dangerous. Robert T. Scully, executive director of the National Association of Police Organizations, notes:

> Many individuals would likely consider being questioned on personal characteristics by a law enforcement officer highly offensive. If an officer is uncertain of someone's ethnic background, the officer would often have to ask for this information and can be expected to meet resistance and hostility to such questions .... One of the most vulnerable moments for a law enforcement officer, is when he or she pulls over a car for a traffic violation .... Since the advent of the automobile, approximately 300 law enforcement officers are known to have died during traffic stops, and approximately 80 percent of those were shot to death. The proposed study would make a dangerous situation worse and escalate bad tempers, by bringing race into the discussion. An officer's life may be put further at risk, as well as the passenger's if the officer has to act in self–defense.[97]

Thus, from the law enforcement perspective, the bill could actually increase racial tension, undermining its good intentions.

Critics also argue that existing measures can solve any racial profiling that occurs. For example, they argue that the need for probable cause protects against illegal searches, since illegally seized evidence

---

[92] Andrew D. Leipold, *Objective Tests and Subjective Bias: Some Problems of Discriminatory Intent in the Criminal Law*, 73 CHI.-KENT. L. REV. 559, 590–91 (1998).

[93] *See* Lytle, *supra* note 51 (quoting Rep. Conyers's assertion that the facts gathered by racial profiling "can give the attorney general the right to haul a precinct into federal court").

[94] Traffic Stops Statistics Study Act of 1998, H.R. 118, 105th Cong. § 3 (1998).

[95] *See* Traffic Stops Statistic Study Act of 1999, H.R. 1443, 106th Cong. (1999).

[96] Scully, *supra* note 86.

[97] *Id.* Nonetheless, the first six months of a similar study in Connecticut yielded no complaints about increased racial tensions as a result of the information-gathering process. Dewan, *supra* note 72.

will be suppressed in court.[98] Moreover, they assert that if private citizens believe they have been wrongfully pulled over, they may file complaints, thereby providing appropriate oversight and protection against the profiling problem.[99] Furthermore, the Attorney General can already investigate accusations of police wrongdoing wherever he sees problems.[100]

Finally, critics argue that a stop, while inconvenient, does not violate a driver's rights in and of itself:

> If an individual, whether that person be African-American, Caucasian, Latino—or a member of any other racial or ethnic group—has been pulled over by an officer with probable cause to make that traffic stop and it turns out that individual has done nothing wrong, then that person is free to go. As a society, sometimes law-abiding citizens will be inconvenienced when police aggressively enforce laws and investigate crimes. Just being stopped by the police when they have good reason to do so should not cause those stopped to believe that their rights were violated.[101]

The Supreme Court has echoed this view: "[O]ne's expectation of privacy in an automobile and of freedom in its operation are significantly different from the traditional expectation of privacy and freedom in one's residence."[102]

Although the Act failed to obtain passage, efforts on the state level have been more successful. By mid-1999, legislators in North Carolina[103] and Connecticut[104] had approved state studies on racial profiling. Legislators in Arkansas,[105] Rhode Island,[106] Pennsylvania,[107] Illinois,[108] Virginia,[109] Massachusetts,[110] Ohio,[111] New Jersey,[112] Maryland,[113] South

---

[98] *See* Scully, *supra* note 86.

[99] Vincent DeMaio, president of the New Canaan, Connecticut police union, advocates strengthening departments' internal affairs divisions and making communities aware that the police are adequately policing their own members. David M. Herszenhorn, *Police and Union Chiefs Meet to Address Racial Profiling*, N.Y. TIMES, Oct. 22, 2000, at A1.

[100] Robert T. Scully, Editorial, *Exaggerating the Racial Profiling Problem*, SAN DIEGO UNION-TRIB., Apr. 23, 1999, at B7.

[101] *Id.*

[102] United States v. Martinez-Fuerte, 428 U.S. 543, 561 (1976).

[103] N.C. GEN. STAT. § 114-10(2a) (1999).

[104] CONN. GEN. STAT. § 54-1m (2000).

[105] H.R. 1261, 82d Gen. Assem., Reg. Sess. (Ark. 1999).

[106] H.B. 8430, 1998 Gen. Assem., Jan. Sess. (R.I. 1998).

[107] H.R. 2617, 182d Gen. Assem., Reg. Sess. (Pa. 1998).

[108] H.B. 1503, 91st Gen. Assem., Reg. Sess. (Ill. 1999).

[109] H.R.J. Res. 736, 1999 Gen. Assem., Reg. Sess. (Va. 1999).

[110] S.B. 1180, 181st Gen. Ct., Reg. Sess. (Mass. 1999).

[111] H.B. 363, 123d Gen. Assem., Reg. Sess. (Ohio 1999).

[112] Con. Res. 111, 208th Leg., Reg. Sess. (N.J. 1999).

[113] S.B. 430, 1999 Gen. Assem., Reg. Sess. (Md. 1999).

Carolina,[114] Oklahoma,[115] and Florida,[116] have all introduced variations on the Traffic Stops Statistics Study Act.[117] In addition, the San Diego police department began collecting data on traffic stops on its own initiative.[118] Police in San Jose, Oakland, Houston, and thirty other cities soon followed suit.[119] According to Professor Harris, the Act has "become the catalyst and the template for data collection by local law enforcement agencies all across the country."[120]

Upon reintroduction, the bill's chances of passage are uncertain. Racial profiling was an issue in the 2000 elections, especially the presidential race. During the presidential campaign, the NAACP ran advertisements encouraging voters to call George W. Bush and complain that he had not supported the prohibition of racial profiling strongly enough.[121] The issue was also a major focus of the New Jersey senatorial campaign.[122] Before leaving office, President Clinton also called for a national ban on racial profiling.[123] That being said, there is a limit to how far even Democrats will go in making racial profiling a significant issue.[124] Laura W. Murphy, director of the ACLU's Washington office, warns that the bill is more likely to be defeated due to neglect rather than opposition.[125]

The opposition of most national law enforcement organizations facilitates this stagnation. In 1998, after the original bill passed the House of Representatives with unanimous bipartisan support, the National Association of Police Organizations, which represents over four thousand police groups nationwide, voiced its opposition to the legislation. Subse-

---

[114] S.B. 778, 113th Gen. Assem., Reg. Sess. (S.C. 1999).

[115] S.B. 590, 47th Leg., 1st Sess. (Okla. 1999).

[116] H.B. 769, 1999 Gen. Assem., Reg. Sess. (Fla. 1999).

[117] *The Stories, The Statistics, and the Law*, *supra* note 35, at 322–23. Opponents of these bills have sometimes succeeded in blocking their passage. Most notably, Governor Gray Davis (D-Cal.) vetoed a bill that called for the tracking of racial profiling by police agencies. Evelyn Nieves, *California's Governor Plays Tough on Crime*, N.Y. TIMES, May 23, 2000, at A16.

[118] *The Stories, The Statistics, and the Law*, *supra* note 35, at 323.

[119] *See The Stories, The Statistics, and the Law*, *supra* note 35, at 323.

[120] *Data Collection*, *supra* note 43 (prepared testimony of Professor Harris).

[121] David Firestone, *Big Push Starts to Lift Turnout of Black Vote*, N.Y. TIMES, Oct. 29, 2000, § 1, at 1.

[122] Newly elected Senator Jon Corzine (D-N.J.) made racial profiling one of his key campaign themes, alongside traditional issues such as health care and education. Robert Hanley, *New Jersey Senate Rivals Push Hard Amid Signs Race is Narrowing; Corzine Calls Campaign a "Joyful Experience,"* N.Y. TIMES, Nov. 7, 2000, at B1.

[123] John F. Harris, *Clinton Urges Extension of Civil Rights Policies*, WASH. POST, Jan. 15, 2001, at A3.

[124] One of Bill Bradley's presidential rallies in March is a good illustration of these limitations. After discussing racial profiling with the crowd, Bradley was asked a question about health care. Appearing almost relieved, Bradley responded, "I'm glad you asked that question, since that is the main aim of my campaign." James Dao, *With Their Biggest Day at Hand, The Candidates Paths Diverge*, N.Y. TIMES, Mar. 7, 2000, at A1.

[125] Raspberry, *supra* note 25.

quently, the bill stalled in the Senate.[126] Even amongst the public, attitudes towards racial profiling vary by race.[127] Survey data indicate that African Americans tend to see racism as continuous and pervasive throughout American society, while whites tend to believe that racism stems from individual bigotry that is the exception, not the rule.[128]

Questions persist regarding President Bush's stance on the issue. On the one hand, in the October 2000 presidential debate at Wake Forest University, Bush stated that he does not want to "federalize" the police; instead, he explained, he seeks to make sure that internal affairs divisions at the local level do their job, with consequences at the federal level if these divisions fail.[129] On the other hand, at the very same debate Bush said that he would support a federal ban on racial profiling.[130]

The views of his Attorney General, who would be charged with enforcing such legislation, are unclear. During his confirmation hearings, Attorney General John Ashcroft told the Senate Judiciary Committee, "I pledge to you that if I'm confirmed as attorney general, the Justice Department will meet its special charge. Injustice against individuals will not stand—no ifs, ands or buts. Racial profiling is wrong. I think it's unconstitutional. I will make racial profiling a priority of mine."[131] However, Ashcroft opposed the Traffic Stops Statistics Study Act while he was in the Senate.[132] Overall, the Bush administration's position on the bill is unclear.

In summary, the prominent nature of racial profiling has created pressure for a federal response. Though limited in scope, the Traffic Stops Statistics Study Act has provoked substantial controversy. The divisive nature of the issue and the uncertain effects of the bill place the future of the Act in doubt. On an uncertain road to stopping racial profiling, this bill may be no more than a speedbump.

—*Gregory M. Lipper*

---

[126] *The Stories, The Statistics, and the Law*, *supra* note 35, at 276.

[127] For instance, 51% of white New Yorkers believe that the New York Police Department engages in racial profiling, whereas 79% of blacks and 66% of Hispanics believe racial profiling occurs. C.J. Chivers, *Approval and Wariness in Poll on Police*, N.Y. TIMES, Sept. 15, 2000, at B9.

[128] Crump, *supra* note 73, at 315.

[129] *Second Presidential Debate Between Gov. Bush and Vice President Gore*, N.Y. TIMES, Oct. 12, 2000, at A22.

[130] *Id.*

[131] David S. Broder, Editorial, *Who Is Ashcroft?*, WASH. POST, Jan. 23, 2001, at A17. This and other statements led one Senate Democratic staff member to remark, "If John Ashcroft is telling the truth, George Bush has given us another David Souter." *Id.* Senator Robert G. Torricelli (D-N.J.) said he would vote to confirm Ashcroft only if he promised to continue federal monitoring of New Jersey state troopers' racial profiling. Helen Dewar, *Boxer to Vote Against Ashcroft Nomination*, WASH. POST, Jan. 11, 2001, at A4.

[132] William Raspberry, Editorial, *No!*, WASH. POST, Dec. 29, 2000, at A33.

TRADING WITH CHINA

On October 10, 2000,[1] President Clinton signed into law "[a]n Act to authorize the extension of nondiscriminatory treatment (normal trade relations treatment) to the People's Republic of China and to establish a framework for relations between the United States and the People's Republic of China."[2] The Permanent Normal Trade Relations for China Act (the "PNTR Law") renders Title IV of the Trade Act of 1974[3] ("Title IV"), which required the President to decide whether to renew normal trading relations on an annual basis,[4] inapplicable to China.[5] Previous extensions of permanent normal trading status to transitional economies (in particular in Eastern Europe) involved changing the treatment of that country's goods from that afforded to communist states to that afforded to market economies. On the other hand, the PNTR Law creates a series of special mechanisms applied only to China that represent a departure from previous U.S. trade legislation and reflect the unique and often difficult nature of the relationship between the United States and China.

Precisely because the legislation mirrors the problematic U.S.-China relationship, the PNTR Law contains measures unique in U.S. trade law and international trade agreements to which the United States is a party. These mechanisms, specifically the product-specific safeguard, appear at odds with the spirit of the principle of most-favored nation status that underlies the World Trade Organization framework. Further, while these mechanisms secured enough support to ensure the bill's passage, they do not address issues related to workers in industries that will be adversely affected by trade with China. While the merits of including labor concerns in international trade treaties remains hotly debated, the PNTR Law, as a piece of U.S. legislation, provided an opportunity to address the needs of those at the losing end of the trade liberalization equation—workers in import-competing industries. The Law's failure to do so in any way beyond the product-specific safeguard represents a lost opportunity.

---

[1] Remarks on Signing Legislation on Permanent Normal Trade Relations with China, WEEKLY COMP. PRES. DOC. 2417 (Oct. 10, 2000).

[2] Permanent Normal Trade Relations for China Act, Pub. L. No. 106-286, 114 Stat. 880 (2000) (to be codified at 19 U.S.C. § 2431). In June 1999 the term "normal trade relations" replaced the term "most-favored nation" in United States law. However, "most-favored nation" continues to be used in WTO agreements and other trade agreements. *See* GOV'T ACCOUNTING OFFICE, WORLD TRADE ORGANIZATION: CHINA'S MEMBERSHIP STATUS AND NORMAL TRADE RELATIONS ISSUES (1994) [hereinafter CHINA'S MEMBERSHIP STATUS].

[3] Trade Act of 1974, Pub. L. No. 93-618, 88 Stat. 1978 (1975) (codified as amended at 19 U.S.C. ch. 12) (1994 & Supp. III 1997)).

[4] 19 U.S.C. § 2432.

[5] Permanent Normal Trade Relations for China Act § 101.

From 1974 to the present, U.S. trade relations with China have been governed primarily by two documents: Title IV[6] and the Agreement on Trade Relations between the United States of America and the People's Republic of China of July 7, 1979 (the "1979 Agreement").[7] At the time of the enactment of the Trade Act of 1974, China, the U.S.S.R., the Soviet Bloc nations, and other communist countries did not enjoy permanent normal trading relations with the United States. Title IV of that Act authorized the President to grant normal trade relations concessions to these economies[8] subject to the provisions of the "Jackson-Vanik Amendment,"[9] which presumptively precludes the grant of normal trade relations to any country that places restrictions on its citizens' ability to emigrate.[10] However, Title IV grants the President the authority to issue an annual waiver to any country, even those otherwise covered by the Jackson-Vanik Amendment, thereby establishing normal trading relations.[11] If Congress does not override the President's waiver by a concurrent resolution of a majority of both houses,[12] it remains in effect for twelve months.[13] China has received waivers of this kind every year since 1980,[14] though not without annual congressional hearings and considerable public attention.[15]

In addition to granting normal trade relations in a unilateral manner, Title IV empowers the President to negotiate bilateral commercial treaties extending normal trading relations when he believes that such agreements "would promote the purposes of the [Trade Act of 1974] and are in the national interest."[16] Title IV, however, significantly limits the

---

[6] *See* Trade Act of 1974 §§ 401–21.

[7] Agreement on Trade Relations, July 7, 1979, U.S.-P.R.C., 31 U.S.T. 4651. The 1979 Agreement is one of 19 bilateral commercial agreements negotiated pursuant to section 405 of the Trade Act of 1974 with countries covered by the so-called "Jackson-Vanik Amendment." WHITE HOUSE CHINA TRADE RELATIONS WORKING GROUP, THE 1979 AGREEMENT (2000) [hereinafter CHINA TRADE RELATIONS WORKING GROUP] (on file with author).

[8] Trade Act of 1974 § 401.

[9] *Id.* § 402.

[10] *Id.* The amendment applies to a country that:

> (1) denies its citizens the right or opportunity to emigrate; (2) imposes more than a nominal tax on emigration or on the visas or other documents required for emigration, for any purpose or cause whatsoever; or (3) imposes more than a nominal tax, levy, fine fee or other charge on any citizen as a consequence of the desire of such citizen to emigrate to the country of his choice.

Id.

[11] *Id.* § 402(c).

[12] *Id.* § 402(d).

[13] *Id.*

[14] CHINA'S MEMBERSHIP STATUS, *supra* note 2, at 4.

[15] *See, e.g., id.* at 4–5; Jim Abrams, *US Must Court China, Clinton Says; Sees Danger in Political Isolation*, CHI. SUN TIMES, June 22, 1998, at 3; Mary Curtius, *Clinton Backs China Trade*, BOSTON GLOBE, May 27, 1994, at 1.

[16] *See* Trade Act of 1974 § 405. The stated purposes of the Act are:

scope of these bilateral agreements.[17] The 1979 Agreement represents one of these conditional trade pacts.[18] This bilateral treaty between the United States and China provides for limited non-discriminatory treatment of goods. Each side gains access to commercial infrastructure—wholesale and retail outlets, trucking and other forms of transport, and maintenance and service shops—on the same terms as other countries but does not gain the right to own or operate distribution systems.[19] Further, while the agreement rhetorically commits both sides to liberalize trade in services, it does not provide for an independent obligation to permit American firms to provide services in China or vice versa.[20] Since 1979, the United States and China have negotiated a number of additional market access agreements.[21]

The annual Title IV-based process of reviewing trade relations with China and granting normal trade relations on an annual basis will no longer be operable once China joins the World Trade Organization ("WTO"). China began the process of accession to the WTO in 1986.[22] The process consists of four phases: "(1) 'fact finding,' (2) negotiation, (3) WTO decision, and (4) implementation."[23] China has completed the fact-finding stage and can complete phase two by negotiating a series of bilateral agreements with current WTO members. Any member country may directly negotiate with an acceding member, but the majority of members only vote to ratify or reject whatever agreement a group of

---

> (1) to foster the economic growth of and full employment in the United States and to strengthen economic relations between the United States and foreign countries through open and nondiscriminatory world trade; (2) to harmonize, reduce, and eliminate barriers to trade on a basis which assures substantially equivalent competitive opportunities for the commerce of the United States; (3) to establish fairness and equity in international trading relations, including reforms of the General Agreement on Tariffs and Trade; (4) to provide adequate procedures to safeguard American industry and labor against unfair or injurious import competition, and to assist industries, firms, workers, and communities to adjust to changes in international trade flows; (5) to open up market opportunities for United States commerce in nonmarket economies; and (6) to provide fair and reasonable access to products of less developed countries in the United States market.

Id. § 2.

[17] *See id.* § 405. Such agreements are terminable at any time for national security reasons, include safeguard arrangements that provide for prompt consultations with the other party in the event of import competition—a rapid increase in imports causing material injury, or threat thereof, to a domestic industry—and provide for the imposition of duties and other countermeasures in the event of a market disruption. The agreements remain valid for only three years, although they can be renewed if the balance of trade is "satisfactory" and the other party to the agreement reciprocates American concessions by lowering its trade barriers. *See id.*

[18] CHINA TRADE RELATIONS WORKING GROUP, *supra* note 7.

[19] *See id.*

[20] *See id.*

[21] CHINA'S MEMBERSHIP STATUS, *supra* note 2, at 13.

[22] *Id.* at 3.

[23] *Id.* at 6.

WTO members including the largest trading economies (referred to as the "working group") negotiates.[24] The agreements will then be combined into a final global agreement that represents the best "deal" any negotiating party secured.[25] For example, if the United States extracts the largest reduction in tariffs on industrial goods and the Europeans extract the most significant concessions on the entrance of foreign financial firms into the Chinese banking sector, these two provisions will be incorporated into the final accession agreement that will operate between China and all other WTO members.

In general terms, the current state of negotiations between China and the WTO includes eight major areas of agreement: tariffs, non-tariff barriers, services, trade framework, intellectual property rights, standards and regulatory practices, agriculture, and monitoring and compliance mechanisms.[26] Chinese negotiators, however, have expressed discontent with final requests made by the United States and the European Union.[27] These requests include the submission of detailed compliance plans and in some cases drafts of the regulations that will implement China's WTO commitments.[28] Chinese officials claim they are being asked to do more than other nations to explain how their country will live up to its compliance promises.[29] China also continues to negotiate with working group members over copyright and trademark laws, rules on issuing business licenses, and the creation of independent judicial agencies that will hear trade disputes and guarantee foreign companies the ability to choose their own Chinese partners and distribute their own goods.[30]

Once all of the bilateral agreements between China and interested WTO members are concluded, those members that have not negotiated with China can take one of three actions: disapprove the accession, approve the accession, or approve the accession but invoke nonapplicability.[31] This last option allows WTO member countries to refuse to apply most favored nation status and other WTO obligations to one another without giving any justification for this decision.[32] Thus, when and if China joins the WTO, the United States, as a WTO member, will be obliged to accord permanent normal trading relations to China unless it invokes a "nonapplication" clause.[33]

---

[24] *Id.* at 8.
[25] *Id.*
[26] *Id.* at 5.
[27] *See* John Pomfret & Phillip P. Pan, *Chinese Talks on WTO Stall; Leaders Resist Reform, Negotiators Say*, WASH. POST, Oct. 10, 2000, at E1.
[28] *Id.*
[29] *See id.*
[30] *Id.*
[31] CHINA'S MEMBERSHIP STATUS, *supra* note 2, at 16.
[32] *See id.* at 16–17.
[33] *Id.* at 5.

Once China becomes a member of the WTO, the Title IV system of granting annual normal trade relations to China will be in violation of WTO regulations. As a first step in addressing this situation, on July 15, 1999 China and the United States signed a new treaty, the Agreement on China's Accession to the World Trade Organization of July 15, 1999 ("1999 Agreement").[34] The PNTR Law incorporates this agreement and amends the Trade Act of 1974 accordingly. It therefore serves as an umbrella under which U.S.-China trade relations may be conducted.

The PNTR Law is functionally divided into two sections. Division A establishes permanent normal trade relations with China by amending the Trade Act of 1974 to make the Jackson-Vanik Amendment inapplicable to China.[35] However, this change in status only becomes operative if China's Protocol of Accession to the WTO is at least "equivalent" to the terms of the 1999 Agreement.[36] Since China's accession negotiations continue and the final text of the accession treaty remains in flux, the concept of equivalence remains to be tested. Division B contains several independent titles that are designed to govern U.S.-China relations on both economic and non-economic matters.[37] This section of the PNTR Law preserves some of the critical functions Congress exercised by way of the annual waiver debate (such as emigration rights) and addresses some of the human rights concerns expressed by those who opposed granting permanent normal trade relations to China.[38]

The Trade Act of 1974 and the General Agreement on Tariffs and Trade ("GATT") contain provisions that regulate the use of safeguards.[39] The term "safeguards" refers to a category of retaliatory measures, such as tariffs and quotas, that countries impose in reaction to increased import competition in a particular category of goods.[40] Unlike tariffs and quotas used in retaliation to dumping, safeguards are invoked in response to lawful instances of import competition. Countries that impose safeguards to protect domestic industries must compensate WTO members adversely affected by the safeguards.[41]

---

[34] Agreement on China's Accession to the World Trade Organization, July 15, 1999, U.S.-P.R.C. [hereinafter 1999 Agreement], *available at* http://www.uschina.org.

[35] Permanent Normal Trade Relations for China Act, Pub. L. No. 106-286, § 101(a), 114 Stat. 880 (2000).

[36] *Id.*

[37] *See id.* §§ 201–701.

[38] *See id.* §§ 301–302, 501–502, 511–513.

[39] Trade Act of 1974, Pub. L. No. 93-618, § 203, 88 Stat. 1978 (1975); WORLD TRADE ORG., THE AGREEMENTS: ANTI-DUMPING, SUBSIDIES, SAFEGUARDS: CONTIGENCIES, ETC. (2000) [hereinafter THE AGREEMENTS] (on file with author).

[40] *See* THE AGREEMENTS, *supra* note 39.

[41] WORLD TRADE ORG., SUMMARY OF THE FINAL ACT OF THE URUGUAY ROUND (1995), *available at* http://www.wto.org/ english/docs_e/legal_e/ursum_e.htm.

> The agreement envisages consultations on compensation for safeguard measures. Where consultations are not successful, the affected members may withdraw equivalent concessions or other obligations under GATT 1994. However, such ac-

The PNTR Law complements this system by creating a new "product-specific safeguard"[42] that is easier to trigger and may be imposed for longer periods than the safeguards contained either in section 201 of the Trade Act of 1974 ("Section 201")[43] or in the WTO Agreement on Safeguards, which limit the use of safeguards to temporary injuries to domestic industry caused by imports.[44] The product-specific safeguard outlined in the PNTR Law differs in several ways from the safeguards authorized by Section 201, which apply to all countries to which the United States has granted permanent normal trade relations. First, the PNTR Law establishes a lower threshold for imposing safeguards than does Section 201. The President may impose Section 201 safeguards against imports that are a "substantial cause" of "serious injury," or present a threat of serious injury, to U.S. firms producing similar articles.[45] Section 201 defines a substantial cause as "a cause which is important and not less than any other cause."[46] In contrast, the PNTR Law permits the United States to utilize product-specific safeguards when a "market disruption" occurs.[47] The statute defines a market disruption as an increase in Chinese imports constituting a "significant cause" of "material injury" to an import-competing domestic industry.[48] The PNTR Law does not define "significant cause" as equivalent to "substantial cause." In the absence of statutory guidance, the "significant cause" standard seems to permit the International Trade Commission ("ITC") to use safeguards in a wider range of circumstances. In addition, the PNTR Law merely requires that the injury be "material,"[49] while Section 201 offers relief only for "serious" injury.[50]

Second, the PNTR Law uses less rigorous, more inclusive factors to determine whether a safeguard should be applied. The PNTR Law considers the volume of imports, the effect of imports on prices for similar U.S. articles, and the effect of imports on the domestic industry producing like or directly competitive products.[51] In contrast, Section 201 offers more specific and targeted factors for safeguard imposition. These differ depending on whether the United States has suffered an injury or faces

---

tion is not allowed for the first three years of the safeguard measure if it conforms to the provision of the agreement, and is taken as a result of an absolute increase in imports.

Id.

[42] Permanent Normal Trade Relations for China Act § 103.
[43] *See* Trade Act of 1974 § 201.
[44] *See* THE AGREEMENTS, *supra* note 39.
[45] Trade Act of 1974 § 202.
[46] *Id.* § 201.
[47] Permanent Normal Trade Relations for China Act § 103.
[48] *Id.*
[49] *Id.*
[50] Trade Act of 1974 § 201(B)(2)(b).
[51] Permanent Normal Trade Relations For China Act § 103.

the threat of injury. If substantial injury is alleged, the ITC must consider "the idling of productive facilities in the industry, the inability of significant numbers of firms to operate at a . . . profit," and underemployment in the industry.[52] If the serious injury is only threatened, the ITC looks at "a decline in sales, a higher and growing inventory, and a downward trend in profits, production, wages or unemployment."[53] The much broader standards in the PNTR Law presumably will make safeguards easier to justify and apply than the tight wording of the 1974 Act.

Third, in contrast to Section 201, the PNTR Law targets Chinese goods for the application of safeguards.[54] Under section 203 of the Trade Act of 1974, the President must globally apply counter-measures against a rise in imports of a certain product rather than apply them against an increase in imports from a specific country.[55] The PNTR Law singles out China as the only country whose products the President may specifically target.[56]

Fourth, the PNTR Law places greater emphasis on executive action than does the Trade Act of 1974. While Title II of the Trade Act of 1974, which contains Section 201, outlines procedures by which the ITC should initiate and conduct investigations,[57] the PNTR Law directly instructs the President to impose import duties or quantitative restrictions in response to market disruptions.[58] This suggests that executive action should be swift in the case of a sudden increase in Chinese imports.[59]

Finally, the PNTR Law contemplates the possibility of trade diversion caused by the use of safeguards by other countries, whereby there is diversion of Chinese imports to the United States.[60] If a third party requests consultations with China under the product-specific safeguard,[61] the Act requires the U.S. Trade Representative to inform the Customs Service, which shall then monitor imports into the United States of those products that are the subject of the consultation request.[62]

The PNTR Law enumerates a host of factors that should be used to determine whether trade diversion exists.[63] It defines trade diversion as a

---

[52] Trade Act of 1974 § 201.

[53] *Id.*

[54] Permanent Normal Trade Relations for China Act § 103.

[55] Trade Act of 1974 § 203.

[56] Permanent Normal Trade Relations for China Act § 103.

[57] Trade Act of 1974 § 201.

[58] Permanent Normal Trade Relations for China Act § 103.

[59] *Id.*

[60] *See id.*

[61] Under the most-favored-nation principle, WTO member nations will be able to apply the PNTR Law to China once it accedes to the WTO. *See* CHINA'S WTO MEMBERSHIP, *supra* note 2, at 8.

[62] *Id.*

[63] Permanent Normal Trade Relations for China Act § 103. These include the actual or imminent increase in the Chinese share of the U.S. domestic market, the absolute level of Chinese exports to the United States or the country seeking consultations, the actual or imminent changes in exports due to the application of safeguards by a third party against

situation in which another WTO member has imposed safeguards against some Chinese imports (thus lowering the overall level of importation of that good into the country invoking the safeguard), causing those imports to be diverted to the U.S. market.[64] If the diversion is substantial enough, it can cause market disruption.[65] Like the provisions on market disruptions directly caused by China, the provision on trade diversion is broad and expansive.[66] The PNTR Law directs the ITC to hold public hearings to determine if a Chinese action has caused or threatens to cause a significant diversion of trade into the domestic market of the United States.[67] In contrast, Section 201 does not have a special procedure for addressing trade diversion, stating only that trade diversion should be considered as a factor in determining if import competition exists, which in turn would trigger the safeguard.[68]

In addition to its differences from Title IV, the product-specific safeguard also differs from the WTO Agreement on Safeguards.[69] In general, the product-specific safeguard appears to offer greater protection against import competition than the Agreement on Safeguards contemplates. Under the Agreement on Safeguards, the President can only impose safeguards in response to "serious injury,"[70] the same standard used in Section 201. The PNTR Law only requires that imports cause a "material injury" before the President can implement a product-specific safeguard.[71] In addition, the Agreement on Safeguards prohibits any ratifying WTO member from imposing safeguard measures against imports from a single country.[72] This is precisely what the PNTR Law allows the United States, and would allow other countries, to do with regard to Chinese goods.[73] Further, under the Agreement on Safeguards, a safeguard can only be applied for a maximum of four years with the possibility of ex-

---

China, and the cyclical or seasonal changes in imports and exports in the United States. *Id.*

[64] *Id.*

[65] *Id.*

[66] *Id.*

[67] *Id.* The PNTR Law states that an action may be caused:

> (1) by the People's Republic of China to prevent or remedy market disruption in a WTO member other than the United States; (2) by a WTO member other than the United States to withdraw concessions under the WTO Agreement or otherwise to limit imports to prevent or remedy market disruption; (3) by a WTO member other than the United States to apply a provisional safeguard within the meaning of the product-specific safeguard provision of the Protocol of Accession of the People's Republic of China to the WTO; or (4) [by] any combination of actions described in paragraphs (1) through (3).

Id.

[68] Trade Act of 1974, Pub. L. No. 93-618, § 201, 88 Stat. 1978 (1975).

[69] *See* THE AGREEMENTS, *supra* note 39.

[70] *Id.*

[71] Permanent Normal Trade Relations for China Act § 103.

[72] *See* THE AGREEMENTS, *supra* note 39.

[73] Permanent Normal Trade Relations for China Act § 103.

tension to eight years under certain circumstances.[74] The PNTR Law, by virtue of the incorporated 1999 Agreement, allows the use of safeguards against China for twelve years following its formal admittance to the WTO.[75] Finally, the Agreement on Safeguards prevents WTO members from imposing safeguards on developing countries if less than three percent of the imports of a particular good come from that nation.[76] It is unclear whether this provision applies to China.

It is similarly unclear under the PNTR Law what methodology the United States will apply to establish the existence of dumping, that is, the export of a good into the American market at a price below that for which the good is sold in the exporter's home market or at a price below its production cost. This definition of dumping only applies to market economies; under the 1999 Agreement the ITC may treat China as a non-market economy for fifteen years for the purpose of establishing whether exporters are dumping goods.[77] As a result, China should face a less strict application of anti-dumping laws at least for the foreseeable future.

WTO agreements also establish anti-dumping rules. Specifically, the WTO Anti-Dumping Agreement and Article VI of GATT set out three factors for countries to use in determining whether an exporter is dumping a product: (1) the price charged in the exporter's domestic market; (2) the price charged by the exporter in another country; or (3) a combination of the exporter's production costs, other expenses, and normal profit margins.[78] The latter two standards roughly correspond to the standard used to determine whether an exporter from a non-market economy is dumping its product.[79] The WTO Agreement on Anti-Dumping is silent on how and when the various methodologies should be employed so there does not appear to be a conflict with using a non-market methodology for fifteen years after China's accession.

In addition to establishing powerful tools, such as the product-specific safeguard and the possibility of treating China as a non-market economy, the PNTR Law goes to unprecedented lengths to provide mechanisms for monitoring China's compliance with the terms of its accession to the WTO and with the rules of the organization. Section 401 of the PNTR Law makes it U.S. policy to obtain, as part of the final Protocol of Accession of the People's Republic of China to the WTO, an annual review compiled by the WTO Secretariat evaluating China's compliance with the terms of its accession.[80] Since the WTO does not require

---

[74] THE AGREEMENTS, *supra* note 39.

[75] CHINA'S MEMBERSHIP STATUS, *supra* note 2, at 14.

[76] THE AGREEMENTS, *supra* note 39.

[77] *See* 1999 Agreement, *supra* note 34, at 2.

[78] THE AGREEMENTS, *supra* note 39.

[79] *See* 1999 Agreement, *supra* note 34, at 2.

[80] Permanent Normal Trade Relations for China Act, Pub. L. No. 106-286, § 401, 114 Stat. 880 (2000).

such a report for any other nation, this arguably violates the principle of most-favored-nation treatment. The PNTR Law also authorizes appropriations to U.S. federal agencies to initiate programs overseeing China's compliance with WTO rules.[81] Finally, the U.S. Trade Representative is required to submit an annual report to Congress on China's compliance with its terms of accession.[82]

Beyond these specific monitoring institutions, the PNTR Law calls for the creation of several institutions and programs to monitor non-trade related aspects of China's behavior. The policy statements contained in the PNTR Law, which embody the concerns expressed by a number of members of Congress, are designed to preserve some of Congress's power to scrutinize and criticize Chinese conduct in these areas.[83] The PNTR Law declares the following to be the policies of the United States: increasing labor and environmental standards in China; encouraging the Chinese to afford worker's internationally recognized rights and to respect human rights; promoting ratification of the International Covenant on Civil and Political Rights; encouraging free movement of peoples within the territory of China; and affording criminal defendants a number of rights found in liberal societies.[84]

To implement these broad policy statements, the PNTR Law creates the "Congressional-Executive Commission on the People's Republic of China" (the "Commission"),[85] the "Task Force on Prohibition of Importation of Products of Forced or Prison Labor from the People's Republic of China" (the "Task Force"),[86] and a multi-agency program to "conduct rule of law training and technical assistance related to commercial activities."[87] The Commission will monitor China's respect for its citizens'

---

[81] The Act provides for additional funding to the Commerce Department for monitoring China's WTO compliance, monitoring import surges (market disruptions), expediting investigations, enforcing unfair trade laws involving Chinese goods and practices (principally the anti-dumping statutes), and establishing a Trade Law Technical Assistance Center. Funds are also authorized for the United States Trade Representative to negotiate further agreements with the Chinese and to monitor China's existing agreements with the United States. The Act further includes authorizations for the Department of Agriculture to provide additional resources to bolster legal and technical assistance, especially food safety and biotechnology. Finally, the PNTR Law authorizes the creation of a joint State/Commerce Department "Overseas Compliance Program" that will include hiring new trade experts to be dispatched to American embassies. *Id.* § 413.

[82] *Id.* § 103.

[83] *Id.* § 203. For a discussion of the chief opponents of the PNTR Law and their primary concerns, see E.J. Dione Jr., *China: Profit and Principle*, WASH. POST, Sept. 15, 2000, at A27. Dione notes that various constellations of opposition have formed around issues such as religious liberty (Senators Paul Wellstone (D-Minn.) and Jesse Helms (R-N.C.) joined forces on this point), national security, environmental and labor standards, and a host of other issues. *See id.*

[84] Permanent Normal Trade Relations Act § 203.

[85] *Id.* § 301.

[86] *Id.* § 501.

[87] *Id.* § 511.

human rights,[88] compile "victim lists" of persons believed to be imprisoned, tortured, or otherwise persecuted by the Chinese government due to political agitation concerning human rights, and monitor the development of the rule of law.[89] The PNTR Law also empowers the Commission to encourage both governmental and non-governmental cultural exchange programs.[90] The provision concerning the Commission also strengthens the institutional role of the Coordinator for Tibet[91] within the State Department and instructs the Commission to work with the Coordinator, thus giving it an independent legislative mandate for overall involvement in U.S.-China relations.[92] Finally, the PNTR Law requires the Commission to compile an annual report on these issues and discuss its findings at congressional hearings—thus recreating the annual congressional debate on China under the Jackson-Vanik Amendment.[93] The membership of the Commission includes members of the House and Senate; representatives of the State, Commerce, and Labor Departments; and two at-large representatives of the executive branch.[94]

The Task Force will focus on China's compliance with legislation concerning prison labor.[95] The Task Force, which is chaired by the Secretary of Treasury, is composed of the Secretaries of Commerce, Labor, and State; the Commissioner of Customs; and the heads of other agencies, as designated by the President.[96] The PNTR Law charges the Task Force with enforcing section 307 of the Tariff Act of 1930,[97] which prohibits entry into the United States of goods mined, produced, or manufactured wholly or in part by convict labor, forced labor, or indentured labor.[98] As with the Commission, the PNTR Law requires the Task Force to submit annual reports to Congress on any attempted violations of section 307. However, the PNTR Law does not require the Task Force to present its findings at a Congressional hearing.

The multi-agency technical assistance and rule of law programs offer incentives to China to improve public administration in several key areas.[99] The PNTR Law provides for Commerce, Labor, and Legal Sys-

---

[88] *Id.* § 302.
[89] *Id.* § 302(b)–(c).
[90] *Id.* § 302(d)–(e).
[91] In the Clinton administration, the Coordinator for Tibet was involved in a broad range of subjects relating to the U.S.-China relationship, including trade. Under the current Bush administration, however, it is unclear what role the Coordinator will play given ongoing State Department reorganization and position cuts.
[92] *Id.* § 302(f).
[93] *Id.* § 302(g)–(i).
[94] *Id.* § 303(a).
[95] *Id.* §§ 501–514.
[96] *Id.* § 511.
[97] 19 U.S.C. § 1307 (1994).
[98] Permanent Normal Trade Relations for China Act § 502.
[99] *Id.*

tem and Civil Society Rule of Law Programs.[100] The statutory language broadly mandates that these programs conduct training and technical assistance related to commercial activities, internationally recognized worker rights, and the development of the legal system, respectively.[101] The statute provides authorizations to cover the cost of the programs' operations,[102] but prohibits these programs from disbursing funds to any organization for which credible evidence exists that it engages in human rights abuses.[103]

Trade agreements with other communist or non-market countries provide a useful comparison for the PNTR Law and the associated 1999 Agreement. Since the enactment of the Trade Act of 1974, the United States has maintained normal trading relations and signed bilateral commercial treaties with the former Soviet Union and many of the former Eastern Block countries despite the Jackson-Vanik Amendment.[104] The trade agreements with these countries contain "escape clauses" that would suspend the agreement if "[e]ither Contracting Party does not have domestic authority to carry out its obligation under [the] Agreement."[105] Title IV mandated the inclusion of this clause and thus prohibited the establishment of permanent normal trade relations.[106] The issuance of these proclamations also required a presidential determination that these agreements were in the national interest.[107] However, the annual proclamations tended to feature little substance.[108]

Starting in 1992, Congress began extinguishing the applicability of Title IV to former Soviet bloc countries that adopted a democratic government and a market economy.[109] The permanent trade agreements signed subsequently created none of the programs or agencies designed to monitor human rights policy featured in the PNTR Law.[110] Furthermore, in the late 1990s the U.S. government began extinguishing Title IV applicability to former communist countries that had applied for WTO membership.[111] While these actions have thus far dealt with fairly small

---

[100] *Id.* § 511.

[101] *See id.*

[102] *Id.* § 512.

[103] *Id.* § 513.

[104] *See, e.g.*, Proclamation No. 6175, 55 Fed. Reg. 37,643 (Sept. 6, 1990) (Czechoslovakia); Proclamation No. 6307, 56 Fed. Reg. 29,787 (June 28, 1991) (Bulgaria).

[105] Trade Act of 1974, Pub. L. No. 93-618, § 405, 88 Stat. 1978 (1975).

[106] *See id.*

[107] *See id.* § 402.

[108] *See, e.g.*, Proclamation No. 9143, 56 Fed. Reg. 31,037 (June 24, 1991); Proclamation No. 9144, 56 Fed. Reg. 31,039 (June 24, 1991); Proclamation No. 9147, 56 Fed. Reg. 40,741 (Aug. 15, 1991).

[109] *E.g.*, Proclamation No. 6419, 57 Fed. Reg. 12,865 (Apr. 10, 1992); Proclamation No. 6922, 61 Fed. Reg. 51,205 (Sept. 27, 1996); Proclamation No. 6951, 61 Fed. Reg. 58,129 (Nov. 7, 1999).

[110] *See supra* note 109.

[111] *See* Proclamation No. 7326, 65 Fed. Reg. 41,547 (June 29, 2000).

economies—Albania and Kyrgyzstan, for example—even in those instances conditions similar to those applied to China were not included.[112]

The circumstances under which the United States and the Soviet Union negotiated their agreement in 1991 would seem to parallel the current U.S.-China relationship.[113] However, that agreement contained none of the compliance or behavior scrutinizing provisions of the PNTR Law and associated 1999 Agreement.[114] The agreement contained language that refused to recognize the forcible incorporation of Lithuania, Latvia, and Estonia into the U.S.S.R., a long-standing U.S. policy, and general language discussing the support the agreement would lend to the development of a market-based economy in the Soviet Union.[115] It also included policy language recognizing that the agreement would promote respect for the internationally recognized rights of working people and reaffirming the parties' commitment to the provisions of the Helsinki process—the Conference on Security and Cooperation in Europe.[116] Unlike the PNTR Law, however, the U.S.-U.S.S.R. agreement did not establish any institutions or programs to implement these broad policy statements or to monitor the Soviet Union's progress in achieving these goals.

The PNTR law therefore differs significantly from previous trade normalization legislation. While these additional features enabled the passage of the PNTR Law in the United States, they create problems in completing negotiations for China's accession to the WTO. Several Chinese negotiators have stated publicly that they are being unfairly required to comply with conditions that other new entrants to the WTO, especially the former Eastern bloc states, were spared.[117] In addition, one wonders if the Chinese bureaucracy and political leadership will have the technical and institutional capacity to carry out the extensive concessions under the PNTR Law and the 1999 Agreement in the agreed upon time period.

In terms of promoting liberalized trade, the PNTR Law represents a significant advancement. Overall, Chinese tariffs on American goods are expected to drop from 25% to 9% by 2005.[118] However, the significant differences between the PNTR Law and existing U.S. trade law, international treaties to which the United States is a party, and prior trade

---

[112] *See id.*

[113] *See* Proclamation No. 6320, 56 Fed. Reg. 37,407 (Aug. 2, 1991).

[114] The agreement included 12 articles and several procedural matters: establishment of normal trade relations, general obligations with respect to market access for products and services, expansion and promotion of trade, government commercial offices, business facilitation, transparency, financial provisions relating to trade in products and services, protection of intellectual property, transit, market disruption and safeguards, dispute settlement, national security, and several procedural matters. *Id.*

[115] *See id.*

[116] *See id.*

[117] Since each new entrant negotiates its own accession, it is debatable whether or not this complaint is legitimate. *See* Pomfret & Pan, *supra* note 27.

[118] Deb Riechmann, *Clinton to Sign China Trade Bill*, AUGUSTA CHRON. (Ga.), Oct. 11, 2000, at B05.

agreements between the United States and transitional non-market economies, are also responsible for the legislation's flaws. Internationally, the use of retaliatory measures specifically designed for China—the product-specific safeguard being the prime example—appears at odds with a global trading system predicated on most-favored-nation status enjoyed by all WTO members. This dissonance undermines the United States' position as a leader in the WTO and instead paints it as parochial and self-interested at best and quasi-protectionist at worst.

Domestically, the legislation fails to fully address the needs of those workers whose industries will be negatively impacted by increased trade with China. While the product-specific safeguard promises to offer them temporary relief, no resources are provided for job re-training or other transitional programs. Further, based on their mandates and anticipated resources, it is unclear that the supervisory institutions the legislation creates will contribute to facilitating domestic political change in China in a meaningful way. Overall, the PNTR Law, while delivering liberalized trade with China, fails to address the concerns of those workers likely to be impacted by increased Chinese competition or to address the concerns of those who feel the legislation does not go far enough to facilitate democratic political change in China.

—Leo Wise