HARVARD JOURNAL
on
LEGISLATION

VOLUME 25, NUMBER 2 SUMMER 1988

ARTICLES

NOTE

POLICY ESSAY

SYMPOSIUM

HARVARD JOURNAL ON LEGISLATION
Harvard Law School
Cambridge, MA 02138

Phone: (617) 495-4400

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

Publication Policy

The JOURNAL specializes in the analysis of legislation and the legislative process. It focuses on legislative reform and on organizational and procedural factors that affect the efficiency and effectiveness of legislative decisionmaking, The JOURNAL is especially interested in publishing articles that examine a public policy problem of nationwide significance and propose legislation to resolve it.

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

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

Subscriptions

The JOURNAL's subscription rate for the 1988 volume (Volume 25) is $20.00 domestic rate; $23.00 foreign surface rate; $35.00 foreign air mail. All correspondence concerning subscriptions should be addressed to the Subscriptions Manager, Publications Center, Harvard Law School, Cambridge, MA 02138.

Unless otherwise requested by the subscriber, the JOURNAL is mailed at bulk rate. For some overseas subscribers, delivery via air mail may be more reliable than standard surface rate service. Claims for missing issues of Volume 25 will be honored only until January 1989.

How to Obtain Copies of Individual Issues

Copies of individual issues from the current volume can be obtained directly from the JOURNAL for $10.00 each (plus $1.50 per copy for foreign mailing).

Copies of issues from all past volumes of the JOURNAL are available from Fred B. Rothman & Co., 10368 West Centennial Road, Littleton, CO 80123. Single isues are $8.50 per copy. Contact Fred B. Rothman & Co. for information on prices for bound volumes.

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

ARTICLE

LENDER/OWNERS AND CERCLA: TITLE AND LIABILITY

ANN M. BURKHART*

The Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) represents Congress' response to the problem of cleaning up hazardous waste sites. The Act and its related regulations authorize the Environmental Protection Agency (EPA) either to order the responsible parties to contain the hazardous waste on the site or to clean the site and charge the responsible parties for EPA's response costs. An unresolved issue is whether these provisions contemplate holding a lender/owner liable for response costs.

In this Article, Professor Burkhart rebuts challenges to lender/owner liability. She begins by scrutinizing the language and legislative history of the liability provisions and their exceptions and reviewing the relevant environmental case law. She then considers constitutional challenges to lender/owner liability. Next, she reviews common law bases of liability. Professor Burkhart concludes that lender/owners should be held liable for response costs under CERCLA.

I. INTRODUCTION

The hazardous waste disposal problem has reached a disastrous level in America. Only recently have studies demonstrated the magnitude of the environmental problem, not only in terms of the large number of dump sites that are polluting our air and water, but also in terms of the effects of these toxics on human beings and their habitats. Evidence has established a causal link between toxic chemical exposure and such health problems as cancer, birth defects, and personality disorders.[1] Dramatic large-

* Associate Professor of Law, University of Minnesota Law School. B.S., Purdue University, 1973; M.S., Purdue University, 1974; J.D., University of Illinois, 1977.

[1] C. BOWSHER, COMPTROLLER GENERAL, REPORT TO THE CONGRESS, CLEANING UP HAZARDOUS WASTES: AN OVERVIEW OF SUPERFUND REAUTHORIZATION ISSUES, GAO/RCED-85-69 5–7 (1985) [hereinafter COMPTROLLER GENERAL REPORT]; 126 CONG. REC. 26,208 (1980) (statement of Rep. Ratchford (D-Conn.)); J. Moorman, Land and Natural Resources Division, Dep't of Justice, The Superfund Concept: Report of the Interagency Task Force on Compensation and Liability for Releases of Hazardous Substances 5–9, (June 1979) [hereinafter The Superfund Concept] (on file at the Harvard Journal on Legislation). *See also* Note, *Allocating the Costs of Hazardous Waste Disposal*, 94 HARV. L. REV. 584, 584 n.1 (1981) [hereinafter Note, *Allocating Costs*]:

> In one case, chemicals allegedly migrating from a dump site included dioxin, which produces cancer, birth defects, and mutations; tetrachloroethylene, a carcinogen that has adverse effects on the central nervous system; and chloroform, which causes narcosis of the central nervous system, destruction of liver cells, kidney damage, harmful alterations of blood chemistry, and cardiac problems.

scale environmental disasters, such as those that occurred at Love Canal and at Times Beach,[2] have impressed upon the American public the potentially catastrophic proportions of the problem. Federal, state, and local governments also have felt the impact, not only in human terms, but also in economic terms; the Love Canal cleanup alone has cost the government more than $30 million, whereas proper disposal practices might have amounted to only $3 to $4 million at the time of disposal.[3]

A University of California public health physician has estimated that 6% of all cancer deaths in California are caused by toxic chemical exposure. 131 CONG. REC. H11,111 (daily ed. Dec. 5, 1985) (statement of Rep. Fazio (D-Cal.)). Representative Fazio also noted toxic chemicals' substantial negative impact on the environment. For example, in San Francisco Bay, the reduction in the striped bass population, which is at least partially attributable to toxic pollutants, costs California's fishing industry several billion dollars per year. *Id.* Some experts, however, believe that the health risks from toxic chemical exposure are minimal. *See, e.g.*, Ames, Magaw & Gold, *Ranking Possible Carcinogenic Hazards*, 236 SCIENCE 271 (1987).

[2] On May 21, 1980, President Carter declared a Federal emergency at the Love Canal in Niagara Falls, New York, the site of chemical dumping by the Hooker Chemical and Plastics Corporation from 1942 to 1953. Nearly 1000 families were evacuated, and the State and Federal Governments provided funds to buy contaminated properties. Molotsky, *President Orders Emergency Help for Love Canal*, N.Y. Times, May 22, 1980, at A1, col. 2; *Actions at the Love Canal Site*, N.Y. Times, Feb. 23, 1980, at A19, col. 6.

On February 23, 1983, the Federal Government announced its intention to spend $33 million to buy back all homes in Times Beach, Missouri, a town contaminated by dioxin sprayed on its streets some ten years earlier. Reinhold, *U.S. Offers to Buy all Homes Tainted by Dioxin*, N.Y. Times, Feb. 23, 1983, at A1, col. 6.

[3] 126 CONG. REC. 26,338 (1980) (statement of Rep. Florio (D-N.J.)). In calculating its fiscal 1988 budget request, the Environmental Protection Agency (EPA) estimated an average cleanup cost of $10–$12 million per site. *Justice Official Tells BNA Conference that PRPs Deserve Access to Superfund Sites*, [Current Developments] 17 Env't Rep. (BNA) No. 49, at 2049, 2050 (Apr. 3, 1987). The Director of the Division of Solid and Hazardous Waste for the New York State Dep't of Environmental Conservation estimated an average cleanup cost of more than $15 million per site. *Id.* One federal agency has stated that cost estimates for cleaning a site range from $1–$30 million. OFFICE OF TECHNOLOGY ASSESSMENT, SUPERFUND STRATEGY 61 (1985). *See also* H.R. REP. NO. 253, 99th Cong., 2d Sess. 256, *reprinted in* 1986 U.S. CODE CONG. & ADMIN. NEWS 2835, 2930–31 (statement of Rep. Scheuer (D-N.Y.)):

> Chevron spent $10 to $12 million to resolve the liability for cleaning up a 30,000-gallon gasoline leak. In 1978, Exxon spent between $5 and $10 million as a result of leak [sic] in East Meadow, New York. Estimates for liability resulting from such underground gasoline leaks range as high as $25 million.

Id. at 2930–31; 126 CONG. REC. 25,100 (1980) (statement of Rep. Eckhardt (D-Tex.)) (In one year "the Justice Department has filed a total of 40 suits for remedial work with an estimated cost of between $330 million and $590 million. While these statistics hint at the size of the problem, they are merely the tip of the iceberg."); The Superfund Concept, *supra* note 1, at 11 (estimated per site cleanup cost of $25.9 million). The cost estimates for rehabilitating the most dangerous sites that have been discovered to date are enormous.

> EPA has estimated that the total price tag for cleaning up the nation's worst abandoned hazardous waste sites could run as high as $46 billion. . . . [The] GAO has estimated that the federal share of cleanup could run as high as $39 billion, with private parties paying roughly equivalent amounts to complete cleanup. The Office of Technology Assessment estimates the total cleanup

The pervasive use of hazardous chemicals in commerce and the virtual absence of effective methods for permanent disposal of hazardous wastes guarantee that the problem will be of continuing importance.[4]

---

price tag at roughly $100 billion, counting both public and private sector contributions to such costs.

H.R. REP. NO. 253, 99th Cong., 2d Sess. 278, *reprinted in* 1986 U.S. CODE CONG. & ADMIN. NEWS 2835, 2953. *See also id.* at 257, 1986 U.S. CODE CONG. & ADMIN. NEWS at 2931–32; COMPTROLLER GENERAL REPORT, *supra* note 1, at 17.

[4] As of January 23, 1987, EPA has included 703 sites on the National Priorities List (NPL) and has proposed 248 additional sites for listing. *EPA Seeks Comments on 64 Proposed Sites to be Added to the National Priorities List*, 17 Env't Rep. (BNA) No. 41, at 1725 (Feb. 6, 1987). Once listed on the NPL, a site is eligible for Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) cleanup funds, more commonly known as "Superfund." Comprehensive Environmental Response, Compensation, and Liability Act, Pub. L. No. 96-510, 94 Stat. 2781 (1980) (codified at 42 U.S.C. §§ 9601–9657 (1982 & Supp. II 1984)), *amended by* Superfund Amendments and Reauthorization Act of 1986, Pub. L. No. 99-499, 1986 U.S. CODE CONG. & ADMIN. NEWS (100 Stat.) 1760–74 [hereinafter all citations to CERCLA, or "the Act," will be to the U.S. Code and Supplement]. The NPL includes only those discovered sites that pose the greatest threat to human health and to the environment. The magnitude of the total problem is better demonstrated by EPA's hazardous waste site inventory, which includes approximately 20,000 sites that may pose a threat to human health and to the environment. *CERCLA Enforcement Figures Called Low for Fiscal 1986*, [Decisions] 25 Env't Rep. (BNA) No. 4, at *front cover, inside front cover* (Nov. 28, 1986). Senator Stafford (R-Vt.) has stated that the actual number of such sites could be 378,000:

> [T]he General Accounting Office [GAO] has reported that the potential universe of Superfund sites in fact could be much larger then [sic] previously acknowledged, and could include some 378,000 facilities. The GAO report concluded that relatively little emphasis has been given to site discovery. Aside from the initial effort in 1982 which uncovered most of the sites on the current inventory, the Federal Government has relied primarily on local governments and the public to discover new sites. It has not conducted any other systematic discovery effort. According to the GAO report, the Environmental Protection Agency has acknowleged [sic] that a targeted, systematic discovery effort combined with a change in program emphasis toward cleaning up sites that have not yet received sufficient attention, could increase the number of sites well beyond the 25,000 figure. For example, the Environmental Protection Agency acknowledged in the report that there are some 34,000 to 52,000 municipal landfills and some 9,770 to 63,770 mining waste sites not yet been listed [sic] or evaluated under the Superfund Program.

132 CONG. REC. S14,896 (daily ed. Oct. 3, 1986) (statement of Sen. Stafford).

Despite general knowledge of the injuries and costs caused by improper disposal, the problem is worsening. EPA estimated that of some 57 million metric tons of hazardous waste produced in 1980, as much as 90% was disposed of in an environmentally unsound manner. 126 CONG. REC. 26,339 (1980) (statement of Rep. Staggers (D-W. Va.)). For example, Representative Ambro (D-N.Y.) stated in floor debates that, during 1978, Hooker Chemicals & Plastics Corp. dumped more than 1,600,000 pounds of industrial wastes in a landfill. "Despite Hooker's knowledge that these wastes were largely hazardous and constitute a threat to human health as they leach into the ground water supply, there was no surveillance by the corporation of the site." 126 CONG. REC. 26,351 (1980) (statement of Rep. Ambro). EPA has now banned land disposal of some of the most toxic wastes. *See U.S. Industry in Midst of Profound Change in Management of Hazardous Waste, Florio Says*, [Current Developments] 17 Env't Rep. (BNA) No. 47 at 1919 (Mar. 20, 1987). *See also Developments—Toxic Waste Litigation*, 99 HARV. L. REV. 1458, 1462 (1986) [hereinafter *Toxic Waste Note*].

The severity of the problem also results in part from the pervasive use of hazardous

In response to the problem of improper hazardous waste disposal, Congress in 1980 enacted the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA),[5] often referred to as "Superfund,"[6] and in 1986, the Superfund Amendments and Reauthorization Act (SARA).[7] Among other provisions, CERCLA authorizes the government not only to rehabilitate hazardous waste sites, but also to recover its costs and other specified damages from the entities specified in subsection 107(a).[8]

---

substances. According to Representative Breaux (D-La.), EPA has identified more than 4,000 types of businesses and industries that have contributed waste to now abandoned hazardous waste sites. "EPA's list of potential responsible parties who have actually caused the problem run from automobiles, banking, electronics and electrical manufacturing, furniture, aircraft and aerospace, optical products, computers, food, beverage and grocery manufacturers, paper and packaging product companies, airlines, rubber products, communications, textiles, and utilities." 131 CONG. REC. H11,080 (daily ed. Dec. 5, 1985) (statement of Rep. Breaux). Representative Moore (R-La.) has stated that EPA attributes only 13% of toxic waste sites to chemical companies' disposal practices. The remaining sites are created by a spectrum of users that normally are not considered to be in the business of generating hazardous wastes. According to Representative Moore, an EPA investigation of a hazardous waste site in Zionsville, Indiana revealed waste contributors that included Eli Lilly, Fred's Frozen Food, Coca-Cola, University of Minnesota, and the Indianapolis Department of Public Works. 131 CONG. REC. H11,106 (daily ed. Dec. 5, 1985) (statement of Rep. Moore). *See also* 132 CONG. REC. S14,908 (daily ed. Oct. 3, 1986) (statement of Sen. Bentsen (D-Tex.)).

[5] 42 U.S.C. §§ 9601–9657 (1982 & Supp. II 1984).

[6] Although the term "Superfund" is popularly used to refer to CERCLA in its entirety, the term more accurately applies to the Hazardous Substance Response Trust Fund that Congress established through CERCLA for the payment of governmental response costs. 42 U.S.C. §§ 9631–9633 (1982).

[7] To provide EPA with the funds necessary for its investigations, remedial actions, and law suits against responsible parties, CERCLA provided for taxation of certain industries that generate hazardous substances. This funding mechanism expired on September 30, 1985. 42 U.S.C. § 9653 (1982). Unable to reach agreement on the terms of a reauthorization bill, Congress subsequently enacted two stopgap appropriation bills in 1985: Hazardous Substance Response Trust Fund, Repayable Advance, Pub. L. No. 99-270, 100 Stat. 80 (1986); Superfund Extension, Pub. L. No. 99-411, 100 Stat. 931 (1986). In October 1986, it enacted the Superfund Amendments and Reauthorization Act (SARA), which, among other provisions, extended funding for an additional five years. Superfund Amendments and Reauthorization Act of 1986, Pub. L. No. 99-499, 100 Stat. 1760–74.

[8] Section 107(a) provides:

Notwithstanding any other provision or rule of law, and subject only to the defenses set forth in subsection (b) of this section—

(1) the owner and operator of a vessel (otherwise subject to the jurisdiction of the United States) or a facility,

(2) any person who at the time of disposal of any hazardous substance owned or operated any facility at which such hazardous substances were disposed of,

(3) any person who by contract, agreement, or otherwise arranged for disposal or treatment, or arranged with a transporter for transport for disposal or treatment, of hazardous substances owned by such person, by any other party or entity, at any facility owned or operated by another party or entity and containing such hazardous substances, and

(4) any person who accepts or accepted any hazardous substances for transport to

The magnitude of damage caused by improperly disposed hazardous wastes and the enormous cost of cleaning up contaminated waste sites prompted Congress to cast a broad net of liability in subsection 107(a).[9] Among those caught in this net are "innocent" lender/owners—secured lenders who have acquired encumbered property without having participated in the dumping activities and who have not continued them. Potential liability of innocent lender/owners most often arises when a lender forecloses on property in which it holds a security interest and purchases the property at the foreclosure sale or when a lender accepts a deed to the property in settlement of the secured debt.

CERCLA has proven to be an unexpected source of liability for lenders, because this is the first time that a government agency has pursued lender/owners in court for conditions on the property. Although other federal[10] and

---

disposal or treatment facilities or sites selected by such person, from which there is a release, or a threatened release which causes the incurrence of response costs, of a hazardous substance, shall be liable for—

(A) all costs of removal or remedial action incurred by the United States Government or a State not inconsistent with the national contingency plan;

(B) any other necessary costs of response incurred by any other person consistent with the national contingency plan;

(C) damages for injury to, destruction of, or loss of natural resources, including the reasonable costs of assessing such injury, destruction, or loss resulting from such a release; and

(D) the costs of any health assessment or health effects study carried out under section 104(i).

The amounts recoverable in an action under this section shall include interest on the amounts recoverable under subparagraphs (A) through (D).

42 U.S.C. § 9607(a) (1982), *amended by* 42 U.S.C.A. § 9607 (West Supp. 1987) (addition of subparagraph D).

[9] *See infra* note 19 and accompanying text.

[10] *See, e.g.*, Federal Water Pollution Control Act, 33 U.S.C. § 1321(a)(6) (1982):

"[O]wner or operator" means (A) in the case of a vessel, any person owning, operating, or chartering by demise, such vessel, and (B) in the case of an onshore facility, and an offshore facility, any person owning or operating such onshore facility or offshore facility, and (C) in the case of any abandoned offshore facility, the person who owned or operated such facility immediately prior to such abandonment.

Deepwater Ports Act, 33 U.S.C. § 1517(m)(4) (1982) ("'[O]wner or operator' means any person owning, operating, or chartering by demise, a vessel."); Clean Air Act, 42 U.S.C. § 7411(a)(5) (1982) ("The term 'owner or operator' means any person who owns, leases, operates, controls, or supervises a stationary source."); Outer Continental Shelf Resource Management Act, 43 U.S.C. § 1811(19) (1982):

"[O]wner" means any person holding title to, or in the absence of title, any other indicia of ownership of, a vessel or offshore facility, whether by lease, permit, contract, license, or other form of agreement, or with respect to any offshore facility abandoned without prior approval of the Secretary of the Interior, the person who owned such offshore facility immediately prior to such abandonment, except that such term does not include a person who, without

state[11] environmental acts and regulations include a list of potentially liable parties that is similar to that contained in CER-

---

participating in the management or operation of a vessel or offshore facility, holds indicia of ownership primarily to protect his security interest in the vessel or offshore facility.

Financial Responsibility for Oil Pollution—Alaska Pipeline, 33 C.F.R. § 131.2(j) (1986) ("'Owner' or 'vessel owner' means any person holding legal or equitable title to a vessel; *Provided, however,* That a person holding legal or equitable title to a vessel solely as security is not an owner." (emphasis in original)); Hazardous Waste Management System: General, 40 C.F.R. § 260.10 (1986) ("'Owner' means the person who owns a facility or part of a facility.").

[11] *See, e.g.*, Massachusetts Oil and Hazardous Material Release Prevention and Response Act, MASS. GEN. L. ch. 21E, § 2 (1986):

"Owner" or "Operator", (1) in the case of a vessel, any person owning, operating or chartering by demise such vessel, (2) in the case of a site, any person owning or operating such site, and (3) in the case of an abandoned site, any person who owned, operated, or otherwise controlled activities at such site immediately prior to such abandonment. The term shall not include a person, who, without participating in the management of a vessel or site holds indicia of ownership primarily to protect his security interest in said vessel or site.

Minnesota Environmental Response and Liability Act, MINN. STAT. § 115B.02(11) (1987):

"Owner of real property" means a person who is in possession of, has the right of control, or controls the use of real property, . . . provided that:

. . . (3) Any person holding a remainder or other nonpossessory interest or estate in real property is an owner of the real property beginning when that person's interest or estate in the real property vests in possession or that person obtains the unconditioned right to possession, or to control the use of, the real property.

New Hampshire Hazardous Waste Cleanup Fund, N.H. REV. STAT. ANN. § 147-B:2(VI) (Equity Supp. 1987) ("'Generator' means any person who owns or operates a facility where hazardous waste is generated."); New Jersey Major Hazardous Waste Facilities Siting Act, N.J. STAT. ANN. § 13:1E-51(n) (West Supp. 1987) ("[I]n addition to the usual meanings thereof, every owner of record of any interest in land whereon a major hazardous waste facility is or has been located, and any person or corporation which owns a majority interest in any other corporation which is the owner or operator. . . ."); New Jersey Spill Compensation and Control Act, N.J. STAT. ANN. § 58:10-23.11b(n) (West Supp. 1987):

"Owner" or "operator" means with respect to a vessel, any person owning, operating or chartering by demise such vessel; with respect to any major facility, any person owning such facility, or operating it by lease, contract or other form of agreement; with respect to abandoned or derelict major facilities, the person who owned or operated such facility immediately prior to such abandonment, or the owner at the time of discharge.

North Carolina Oil Pollution and Hazardous Substances Control Act, N.C. GEN. STAT. § 143-215.77(12) (1983) ("'Operator' shall mean any person owning or operating an oil terminal facility or pipeline, whether by lease, contract, or any other form of agreement.").

Interestingly, the Massachusetts and Minnesota Acts impose liability on the "owner *or* operator." MASS. GEN. L. ch. 21E, § 5(a) (1987) (emphasis added), and MINN. STAT. § 115B.03(1) (1986) (emphasis added). Conversely, CERCLA imposes liability on the "owner *and* operator." 42 U.S.C. § 9607(a)(1) (1982) (emphasis added). *See infra* notes 37–61 and accompanying text.

Many other states have enacted legislation imposing liability for hazardous waste cleanup costs. Although only the state acts listed *supra* actually define "owner," several other state acts impose liability on the property owner. *See, e.g.*, KY. REV. STAT. ANN. § 224.877(3), (4), (6)(a) (Michie/Bobbs-Merrill Supp. 1986); ME. REV. STAT. ANN. tit.

CLA, no reported decision exists in which the Environmental Protection Agency (EPA) or another authorized plaintiff has pursued a lender inside or outside the courtroom. Consequently, when two CERCLA cost recovery actions were filed against lenders that had held security interests in properties that EPA cleaned of hazardous wastes,[12] shock waves reverberated through the lending industry. Business journal articles and continuing legal education programs have since warned lenders of this unexpected source of liability and have counseled methods for attempting to avoid it.[13]

Lenders have been particularly concerned about the prospect of CERCLA liability. First, the cost of cleaning a hazardous waste site, especially when combined with the other elements of CERCLA damages, often and substantially exceeds the amount the lender agreed to invest against the security of the property. Indeed, such damage amounts often exceed the fair market value of the land even after it is cleaned of hazardous wastes. In such a case, a lender faces the possibility of a much greater economic burden than was anticipated when it made a loan secured by the contaminated property.

Second, many insurance companies have taken the position that commonly used comprehensive general liability policies do not include liability for hazardous waste related injuries and

---

38, § 1319-J (West Supp. 1986); MICH. COMP. LAWS ANN. § 299.541 (West 1984); N.Y. ENVTL. CONSERV. LAW § 27-1313(3)(a) (Consol. Supp. 1987); OHIO REV. CODE ANN. § 3734.22 (Anderson Supp. 1987); S.C. CODE ANN. § 44-56-60(b)(1)–(c)(2) (Law. Co-op. Supp. 1987); Texas Solid Waste Disposal Act, TEX. REV. CIV. STAT. ANN. art. 4477-7, § 8(g)(2)(A) (Vernon Supp. 1988); Texas Hazardous Substances Spill Prevention and Control Act, TEX. WATER CODE ANN. § 26.263(6) (Vernon Supp. 1988); WIS. STAT. ANN. § 144.76(3), (7) (West Supp. 1987). For a discussion of the impact of state hazardous waste laws on real estate transactions, see Angelo & Bergeson, *The Expanding Scope of Liability for Environmental Damage and its Impact on Business Transactions*, 8 CORP. L. REV. 101, 106–08 (1985); Dean, *How State Hazardous Waste Statutes Influence Real Estate Transactions*, [Current Developments] 18 Env't Rep. (BNA) No. 14, at 933 (July 31, 1987).

[12] United States v. Maryland Bank & Trust Co., 632 F. Supp. 573 (D. Md. 1986); United States v. Mirabile, 15 Envtl. L. Rep. (Envtl. L. Inst.) 20,994 (E.D. Pa. 1985).

[13] *See, e.g.*, Angelo & Bergeson, *supra* note 11, at 113; Apple & Guthrie, *Caveat Emptor When it Comes to Super Fund Liability*, MINN. L.J. 11 (1987); Bleicher & Stonelake, *Caveat Emptor: The Impact of Superfund and Related Laws on Real Estate Transactions*, 14 Envtl. L. Rep. (Envtl. L. Inst.) No. 1, at 10,017 (Jan. 1984); Burcat, *Environmental Liability of Creditors: Open Season on Banks, Creditors, and Other Deep Pockets*, 103 BANKING L.J. 509 (1986); Richman & Stukane, *Avoiding the Environmental Risks in Mortgage Transactions*, 2 REAL EST. FIN. J. 13 (Winter 1987); Shea, *Protecting Lenders Aqainst Environmental Risks*, PRAC. REAL EST. LAW. 11 (May 1987); Berz & Sexton, *Superfund Collides with Lenders' Concerns*, Legal Times, Dec. 23–30, 1985, at 13, col. 1.

cleanup costs. As a result, a lender that believed it was comprehensively covered against liability confronts the possibility of being a self-insurer for a large unanticipated liability.[14]

Surprisingly in light of the effects on lenders, Congress apparently did not consider the issue of lender/owner liability when it enacted CERCLA. Despite hundreds of pages of legislative history,[15] not one reference exists to the lender/owner's potential liability.[16] More surprisingly, SARA's legislative history also does not mention the issue although the opinions in earlier

---

[14] For a discussion of the insurance issues that have arisen with respect to environmental liability, see Adler & Broiles, *The Pollution Exclusion: Implementing the Social Policy of Preventing Pollution Through the Insurance Policy*, 19 LOY. L.A.L. REV. 1251 (1986); Angelo & Bergeson, *supra* note 13, at 114; Bauer & Lakind, *Toward Resolution of Insurance Coverage Questions in Toxic Tort Litigation*, 38 RUTGERS L. REV. 677 (1986); *Toxic Waste Note*, *supra* note 4, at 1573–85; Note, *The Pollution Exclusion in the Comprehensive General Liability Insurance Policy*, 1986 U. ILL. L. REV. 897; Annotation, *Liability Insurance Coverage for Violations of Antipollution Laws*, 88 A.L.R.3d 182 (1978).

[15] It is somewhat inaccurate to refer to CERCLA's legislative history.

> Although Congress had worked on "Superfund" toxic and hazardous waste cleanup bills and on parallel oil spill bills for over three years, the actual bill which became law had virtually no legislative history at all. The bill which became law was hurriedly put together by a bipartisan leadership group of senators (with some assistance from their House counterparts), introduced, and passed by the Senate in lieu of all other pending measures on the subject. It was then placed before the House, in the form of a Senate amendment of the earlier House bill. It was considered on December 3, 1980, in the closing days of the lame duck session of an outgoing Congress. It was considered and passed, after very limited debate, under a suspension of the rules, in a situation which allowed for no amendments. Faced with a complicated bill on a take it-or-leave it basis, the House took it, groaning all the way.

Grad, *A Legislative History of the Comprehensive Environmental Response, Compensation and Liability ("Superfund") Act of 1980*, 8 COLUM. J. ENVTL. L. 1, 1 (1982) (footnote omitted). *See also* Eckhardt, *The Unfinished Business of Hazardous Waste Control*, 33 BAYLOR L. REV. 253 (1981). Because of the limited time Congress had to consider CERCLA, no committee reports exist for the law as enacted. The floor debates constitute the only directly related legislative history. *See* Bulk Distrib. Centers v. Monsanto Co., 589 F. Supp. 1437, 1441 (S.D. Fla. 1984) ("CERCLA's legislative history is riddled with uncertainty because lawmakers hastily drafted the bill and because last minute compromises forced changes that went largely unexplained.") (dicta). For this reason, one court has advised that "the Committee Reports must be read with some caution." United States v. Reilly Tar & Chem. Corp., 546 F. Supp. 1100, 1111 (D. Minn. 1982).

[16] One Representative focused on the problem of lenders refusing loans to companies that may be subject to CERCLA liability, but he did not address the problem of the lenders' potential liability.

> Financial institutions are extremely wary of lending capital for operations when the borrower may or may not be subject to huge liabilities created by the legal disposal of hazardous waste. The impact of this ripples through the economy as small business finds itself unable to borrow needed capital for expansion and investment due to the contingent liabilities generated under the CERCLA liability system.

131 CONG. REC. H11,091 (daily ed. Dec. 5, 1985) (statement of Rep. Brown (R-Colo.)).

CERCLA suits involving lender/owners were published before SARA became law.[17]

The references in the legislative histories concerning the scope of liability provide little insight into whether or to what extent Congress intended lender/owners to be subject to CERCLA liability. Although some statements in the legislative history indicate that Congress may have intended to impose liability only on an entity that generated, transported, or permitted dumping of hazardous materials,[18] other relevant policy statements indicate that Congress did not intend to limit liability in this way.[19] The legislative history includes statements by several

---

[17] According to one commentator, the lenders' lobby did attempt to have Congress exempt lender/owners from liability.

> Although bank lobbies attempted to persuade members of Congress to include in section 101(35) of the 1986 Superfund amendments a provision exempting mortgagees from liability when they acquire possession of land by foreclosure, members of the Senate were so hostile to the idea that it was never even formally considered in committee.

Telephone interview with Robert Norris, Legislative Assistant to Congressman Barney Frank (D-Mass.) (Nov. 14, 1986). Comment, *The Impact of the 1986 Superfund Amendments and Reauthorization Act on the Commercial Lending Industry: A Critical Assessment*, 41 U. MIAMI L. REV. 879, 904 n.139 (1987). While Congress considered the 1986 amendments, the Federal Home Loan Corporation, a major purchaser in the secondary mortgage market, proposed an alternative definition of "contractual relationship" that would insulate a lender/owner from CERCLA liability if it did not have notice of the hazardous wastes when it made the loan secured by the polluted land. *Id.* at 907–08.

[18] *See, e.g.*, 132 CONG. REC. S14,934 (daily ed. Oct. 3, 1986) (statement of Sen. Durenberger (R-Minn.)) ("[Superfund] imposed strict, joint, and several liability on those who manufacture, handle, and dispose of hazardous substances."); S. REP. No. 848, 96th Cong., 2d Sess. (1980); Letter from Douglas Costle to Sen. Randolph (D-W. Va.), (Sept. 25, 1979):

> The supposition of the Administration's proposal is that society should not bear the costs of protecting the public from hazards produced in the past by a generator, transporter, consumer, or dumpsite owner or operator who has profited or otherwise benefited from commerce involving these substances and now wishes to be insulated from any continuing responsibilities for the present hazards to society that have been created;

126 CONG. REC. 26,358 (1980) (statement of Rep. Findley (R-Ill.)) ("[H.R. 7020 would] . . . establish strict lines of liability for those who engage in the waste disposal business."); 126 CONG. REC. 26,339 (statement of Rep. Staggers (D-W. Va.)) ("[H.R. 7020] provides that defendants who caused or contributed to hazardous waste situations necessitating response action by the Administrator shall be strictly, jointly and severally liable for the costs of such action."); 126 CONG. REC. 30,932 (1980) (statement of Sen. Randolph) ("The purposes of [S. 1480] were: First, to make those who release hazardous substances strictly liable for cleanup costs, mitigation, and third-party damages. . . ."). The last two quotations can be interpreted as including a property owner who did not operate the dump and did not authorize anyone else to do so but, having failed to clean the site of hazardous wastes thereby necessitating an EPA cleanup, has "contributed to" and permitted the "release" of hazardous substances.

[19] *See, e.g.*, Senator Lautenberg's (R-N.J.) statements with respect to the CERCLA reauthorization bill:

> Cleaning up these sites will be expensive. It will require assigning liability to the parties who contributed to the creation of these sites. And, it will require

members of Congress to the effect that "polluters should pay"[20] or that the parties "responsible" for a hazardous waste release should be liable.[21]

These statements are consistent with the statutory imposition of liability on a property owner that acquired a dump site but did not operate the dump, such as a lender/owner. Imposition of liability is, moreover, consistent with traditional precepts of property law. Historically, a property owner has been legally responsible for a hazardous condition existing on its land even if it did not create the condition. In this sense, the owner is deemed to be a tort-feasor. Although it did not create the hazardous condition, the owner is permitting the condition on its land to harm others.[22]

In the absence of a clear congressional expression of intent concerning a lender/owner's liability for the cost of cleaning up a hazardous waste site, the resolution of the issue turns on the statutory language and that language's relationship with the policies underlying CERCLA. This Article will analyze CERCLA's liability provisions and will demonstrate that Congress intended that a lender/owner may be held liable for response costs. This Article then will establish that this potential imposition of liability is constitutional and is consistent with a landowner's—

---

a financial contribution from a range of parties, some of whom may not have contributed directly to our toxic waste problem, but all of whom have benefited from the products produced by the chemical and petroleum industries.

132 CONG. REC. S14,911 (daily ed. Oct. 3, 1986) (statement of Sen. Lautenberg).

[20] *See, e.g.*, 132 CONG. REC. S14,932 (daily ed. Oct. 3, 1986) (statement of Sen. Wallop (R-Wyo.)) ("The cost of cleaning environmental problems should be based on the principle that the polluter should pay."); *id.* at S14,923 (statement of Sen. Chafee (R-R.I.)) ("[The funding proposal] is consistent with the principle of Superfund: the principle that the polluter pays."); 131 CONG. REC. H11,118 (daily ed. Dec. 5, 1985) (statement of Rep. Traficant (D-Ohio)) ("Polluters should pay the entire cost of cleaning the mess they created and the grave hazards they caused.").

[21] *See, e.g.*, 132 CONG. REC. S14,934 (daily ed. Oct. 3, 1986) (statement of Sen. Durenberger) ("When it was adopted in 1980, Superfund was designed to assure that those who are responsible for the release of hazardous substances into the environment would also bear the responsibility of responding to the threats that those substances pose. That was the theory of Superfund."); 131 CONG. REC. H11,117 (daily ed. Dec. 5, 1985) (Letter to Speaker of the House and Minority Leader from selected Members of Congress (Dec. 4, 1985) ("Liability is the most effective tool that EPA has for bringing the responsible parties to the bargaining table to negotiate a cleanup agreement.")); 126 CONG. REC. 26,338 (1980) (statement of Rep. Florio) ("[The liability provision] assures that the costs of chemical poison releases are borne by those responsible for the releases. It creates a strong incentive both for prevention of releases and voluntary cleanup of releases by responsible parties.").

[22] *See infra* notes 190–237 and accompanying text. The monumental problems of proof and formidable procedural obstacles, such as statutes of limitations, however, can prevent recovery by hazardous waste victims. *See* Farber, *Toxic Causation*, 71 MINN. L. REV. 1219 (1987); *Toxic Waste Note, supra* note 4, at 1604–16.

including a lender/owner's—common law liability for the condition of its land. Finally, this Article will propose a relaxation of the usual common law rules concerning apportionment of damages in CERCLA-related contribution actions to apportion liability more equitably among the statutorily responsible parties.

II. CERCLA

Congress enacted CERCLA to deal with hazardous waste dump sites that are polluting neighboring lands or water sources. The Act and its related regulations authorize EPA either to order the responsible parties to contain the hazardous waste on the site[23] or to clean the site and charge the responsible parties for EPA's response costs.[24] The statutorily defined response costs include the cost of cleaning the site, any damages for injuries to natural resources, and the costs of any health assessment or health effects studies conducted pursuant to statutory authority.[25] If EPA cleans the site, it may recover its response costs from any one or more of the persons enumerated in subsections 107(a)(1)–(4) of the Act, including: (1) the current "owner and operator" of the waste site;[26] (2) any person who owned or

---

[23] 42 U.S.C. § 9606(a) (1982).

[24] *Id.*

[25] The recoverable response costs are enumerated in *id.* § 9607(a). *See supra* note 8.

[26] 42 U.S.C. § 9607(a)(1) (1982). Subsection 107(a)(1) does not specify when a person must have been the "owner or operator" for liability to attach. It merely provides that "the owner or operator of a vessel . . . or a facility" is liable for CERCLA response costs. This language can be construed to impose liability on the person who owned (1) after dumping stopped but before the cleanup, (2) during the cleanup, or (3) when the cost recovery action is filed. Those persons who owned during any dumping activities are liable pursuant to § 107(a)(2). *Id.* § 9607(a)(2). EPA correctly interprets § 107(a)(1) as imposing liability only on the person who was the owner during the cleanup. S. Leifer, EPA Deputy Associate Enforcement Counsel for Waste, Office of Enforcement & Compliance Monitoring, Paper presented at Environmental Risks for Lenders Conference entitled Lender Liability Under CERCLA 2 (June 25, 1987) [hereinafter Leifer Paper] (on file at the Harvard Journal on Legislation). Courts also have construed § 107(a)(1) in the same way. United States v. Cauffman, 21 Env't Rep. Cas. (BNA) 2167 (C.D. Cal. 1984) (owner at time of cleanup liable); New York v. Shore Realty Corp., 759 F.2d 1032, 1044–45 (2d Cir. 1985) (owner at time of cleanup liable although did not own at time of disposal). This interpretation is consistent with the economics of imposing liability on an owner who did not participate in or permit dumping. The owner during the cleanup benefited by the increase in fair market value attributable to the cleanup, whereas a person who subsequently purchases the property presumably has paid a price that reflects the uncontaminated value of the property. If the owner during the cleanup purchased without notice of the waste on the property and paid the uncontaminated value for the land, it will escape liability pursuant to the third-party defense if it exercised reasonable care after discovering the waste. 42 U.S.C. § 9607(b)(3) (1982). Furthermore,

operated the waste site when any hazardous dumping occurred; (3) the hazardous waste generators; and (4) the hazardous waste transporters.[27] Liability for response costs is strict, joint, and several with limited exception.[28]

The terms "owner and operator" in subsection 107(a)(1) are generic and unmodified.[29] Therefore, a lender that acquires title to a hazardous waste site, whether by foreclosure or by other means, potentially is liable for CERCLA response costs even though the lender did not participate in operation of the site or in generating or transporting the waste at the site. Liability follows title as it does under the common law. Under this theory, lenders who acquired title to encumbered property by foreclosure have been sued as "owners" in two response cost recovery actions.[30] Because lender/owners will usually be easy to locate

---

the owner during the cleanup, if liable for response costs, could have avoided that liability by cleaning the site rather than waiting for EPA or another responsible party to do so. Subsection 9604(a)(1) authorizes the government to clean a hazardous waste site "unless the President determines that such removal and remedial action will be done properly by the owner or operator of the vessel or facility from which the release or threat of release emanates. . . ." *Id.* § 9604(a)(1). Note that this alternative is unavailable to a person who acquires the site after a cleanup. Imposing liability on the owner during the cleanup also is consistent with the common law tort liability of a landowner, whereas such liability does not attach to a person who acquires property after the hazardous condition has been eliminated. *See infra* notes 190–237 and accompanying text. No reported decision involves a suit against a person who acquired a waste site after an EPA cleanup.

The third class of owners, those who did not own during dumping and sold before the cleanup, were not subject to liability under CERCLA as originally enacted, although they may be liable pursuant to the CERCLA amendments. In Cadillac Fairview/California, Inc. v. Dow Chem. Co., 14 Envtl. L. Rep. (Envtl. L. Inst.) 20,376, 20,378 (C.D. Cal. 1984), the court stated the pre-SARA law:

> This defendant argues on this motion that the scope of liability of the Act is not so broad as to encompass a party who merely owned the site at a previous point in time, who neither deposited nor allowed others to deposit hazardous wastes on the site. This appears to be correct under any but a very strained reading of the Act.

That former innocent owners were not liable under CERCLA as originally enacted is further demonstrated by the SARA amendment that provides one circumstance under which such an owner will be liable for response costs. New § 101(35)(c) provides that such an owner will be liable if it had actual knowledge of a release or threatened release from the property but did not disclose that condition to the person who acquired title from it. 42 U.S.C.A. § 9601(35)(c) (West Supp. 1987). Such an amendment would have been unnecessary if this class of owners already were liable under § 107(a)(1).

[27] *See supra* note 8.

[28] *See infra* notes 239–74 and accompanying text. Although courts and commentators frequently state that CERCLA imposes strict liability, liability is less than strict. An otherwise responsible party will not be liable for CERCLA response costs if it can establish that the hazardous condition was caused solely by an act of God, an act of war, or an unrelated third party. 42 U.S.C. § 9607(b) (1982 & Supp. II 1984). The third-party defense is discussed at *infra* notes 115–43 and accompanying text.

[29] 42 U.S.C. § 9607(a)(1) (1982).

[30] *See infra* notes 89–113 and accompanying text.

and to join in a response cost recovery action and will generally have sufficiently deep pockets to pay a judgment against them, lender/owners will continue to be attractive defendants and more suits are likely.

If the lender participated in operation of the waste site, whether before or after acquiring the site, it may be liable as an "operator." In *United States v. Mirabile*, for example, a federal district court stated that a secured lender could incur CERCLA liability as an "operator" if it had been actively engaged in its borrowers' business operations.[31] A lender/owner also bears potential tort liability for injuries resulting from the condition of the land and for any failure to disclose that condition when it sells the land.[32]

If the lender/owner is not potentially liable as an operator, but only as an owner, however, it has at least four plausible arguments to attempt to avoid CERCLA liability: (1) subsection 107(a)(1) imposes liability only on owners that also are operators;[33] (2) pursuant to an exception in the Act's definition of "owner or operator," the lender is not liable because it acquired title primarily to protect its security interest;[34] (3) the lender qualifies for the subsection 107(b)(3) third-party defense;[35] and (4) imposition of liability on a lender/owner is unconstitutional.[36] As will be demonstrated, the third-party defense is the only viable argument to avoid liability, and the exception it offers is narrow.

A. *Subsection 107(a)(1)—"The Owner and Operator"*

Subsection 107(a)(1) provides that "the owner and operator" of a waste site is liable for CERCLA response costs.[37] The Act broadly and circularly defines "the owner or operator" as a person who owns, operates, charters, or otherwise controls a

---

[31] 15 Envtl. L. Rep. (Envtl. L. Inst.) 20,994, 20,997 (E.D. Pa. 1985).

[32] 42 U.S.C.A. § 9601(35)(C) (West Supp. 1987). *See infra* note 140 and accompanying text.

[33] *See infra* notes 37–61 and accompanying text.

[34] *See infra* notes 63–113 and accompanying text.

[35] *See infra* notes 115–43 and accompanying text.

[36] *See infra* notes 146–89 and accompanying text.

[37] 42 U.S.C. § 9607(a)(1) (1982).

vessel or facility.[38] In *United States v. Maryland Bank & Trust Co.*,[39] a CERCLA response cost recovery action against, among others, a lender/owner, the lender/owner argued that subsection 107(a)(1) imposes liability on a property owner only if that person also is an operator. The defendant argued that the conjunctive "and," together with use of the article "the" before the word "owner" but not before the word "operator," provided that an owner is liable only if it also is an operator.[40]

The court was unreceptive to the defendant's grammar-based argument, noting that "by no means does Congress always follow the rules of grammar when enacting the laws of this nation."[41] An authority on statutory construction, as well as a grammarian, provide support for the court's finding of potential liability for non-operating owners in the statute's language.[42] With respect to the use of "and" rather than "or," a leading authority on statutory construction states: "There has been . . . [such] great laxity in the use of these terms that courts have generally said that the words are interchangeable and that one may be substituted for the other, if consistent with the legislative intent."[43] With respect to the absence of the second "the" in the liability provision, H.W. Fowler has stated in his classic work, *A Dictionary of Modern English Usage*, that insistence on "the" being placed before each noun in a series is an example of "needless rigidity."[44] This type of statutory flyspecking is par-

---

[38] Section 9601(a)(20) provides:

> "[0]wner or operator" means (i) in the case of a vessel, any person owning, operating, or chartering by demise, such vessel, (ii) in the case of an onshore facility or an offshore facility, any person owning or operating such facility, and (iii) in the case of any abandoned facility, any person who owned, operated, or otherwise controlled activities at such facility immediately prior to such abandonment. Such term does not include a person, who, without participating in the management of a vessel or facility, holds indicia of ownership primarily to protect his security interest in the vessel or facility.

Id. § 9601(a)(20). The security interest exception is discussed later in this Article. *See infra* notes 63–111 and accompanying text.

[39] 632 F. Supp. 573, 577–78 (D. Md. 1986).

[40] *See id.* at 577.

[41] *See id.* at 578.

[42] *See, e.g.*, 1A SUTHERLAND, STATUTORY CONSTRUCTION § 21.14, at 127 (4th ed. 1985); H.W. FOWLER, A DICTIONARY OF MODERN ENGLISH USAGE 630–31 (2d ed. 1965).

[43] 1A SUTHERLAND, *supra* note 42, § 21.14, at 127.

[44] Fowler writes:

> What may fairly be expected of us is to realize that among expressions of several adjectives or nouns introduced by *the* some obviously cannot have *the* repeated with each item without changing the sense (*the black and white penguins*), and some can logically claim the repetition (*the red and the yellow tomatoes*). A careful writer will have the distinction in mind, but he will not

ticularly rigid in light of CERCLA's generally acknowledged drafting deficiencies.[45]

Placing the language used in subsection 107(a)(1) in context suggests that its adoption was the result of a drafting deficiency. First, construing subsection 107(a)(1) as imposing liability on an "innocent" owner is necessary to make section 107 internally consistent. For example, subsections 107(c)(1) and 107(c)(2), which provide the amount of chargeable response costs, both refer to the liability of an "owner or operator."[46] These provisions contemplate that an owner need not be an operator to be liable for response costs.

Second, Congress used the subsection 107(a)(1) phrase "the owner and operator" in other subsections that clearly contemplate the possibility of two different persons filling these roles. Subsection 107(k)(2), for example, provides that if "the owner or operator" of a facility notifies EPA that specified conditions for transferring liability have been satisfied, such transfer will be effective unless EPA otherwise notifies "the owner and operator," in which case "the owner and operator" continue to be liable.[47]

---

necessarily be a slave to logic; "the red and yellow tomatoes" may be preferred for better reasons than ignorance or indolence. For other attempts to impose a needless rigidity, see ONLY and NOT I.

H.W. FOWLER, *supra* note 42, 630–31 (emphasis in original).

[45] Several courts and commentators have noted the Act's drafting deficiencies. *See, e.g.*, 632 F. Supp. at 578 (D. Md. 1986); Bulk Distrib. Centers, Inc. v. Monsanto Co., 589 F. Supp. 1437, 1441 & n.10 (S.D. Fla. 1984); United States v. Northeastern Pharm. & Chem. Co., 579 F. Supp. 823, 838 n.15 (W.D. Mo. 1984), *rev'd in part on other grounds*, 810 F. 2d 726 (8th Cir. 1986), *cert. denied*, 108 S. Ct. 146 (1987) [hereinafter NEPACCO]; Eckhardt, *supra* note 15, at 253. These drafting deficiencies are explained, in part, by the Act's legislative history. CERCLA was a compromise package of other bills that coalesced in last minute negotiations. Grad, *supra* note 15, at 1–2. *See infra* note 54.

[46] Subsection 9607(c)(1)–(2) provides:

(1) Except as provided in paragraph (2) of this subsection, the liability under this section of an owner or operator or other responsible person for each release of a hazardous substance or incident involving release of a hazardous substance. . . .

(2) Notwithstanding the limitations in paragraph (1) of this subsection, the liability of an *owner or operator* or other responsible person under this section shall be the full and total costs of response and damages, if. . . .

42 U.S.C. § 9607(c)(1)–(2) (1982 & Supp. II 1984) (emphasis added). *See also id.* §§ 9607(e)(1)–(2), (h), (k)(1).

[47] Subsection 9607(k)(2) provides:

Such transfer of liability shall be effective ninety days after *the owner or operator* of such facility notifies the Administrator of the Environmental Protection Agency . . . that the conditions imposed by this subsection have been satisfied. If within such ninety-day period the Administrator . . . determines that any such facility has not complied with all the conditions imposed by this subsection or that insufficient information has been provided to demonstrate

Finally, because no question exists that a waste site operator is liable for CERCLA response costs, a grammarian would be hard pressed to explain why Congress imposed liability on owners and operators rather than on operators alone, unless it intended to impose liability on an owner who was not also an operator. In fact, interpreting subsection 107(a)(1) as imposing liability only on an owner who also is an operator would cause that provision to be mere surplusage, because subsection 107(a)(2) imposes liability on "any person who . . . owned or operated" a facility when any dumping occurred.[48] Thus, if a court employs the rule of construction that statutory language should be construed in a manner that does not render a clause superfluous,[49] the court should interpret subsection 107(a)(1) as imposing liability on the current owner even if that person did not operate the facility.

An examination of a statute's legislative history may help resolve ambiguities in the text.[50] Although the legislative history for CERCLA as enacted is sparse,[51] the floor debates on the Act indicate that Congress intended to impose liability on an owner regardless of whether that person also is an operator. For example, Congressman Broyhill (R-N.C.) stated: "[U]nder the language of section 107 the owner *or* operator of a vessel or a facility can be held strictly liable for various types of costs and damages entirely on the basis of having been found to be an owner *or* operator of any facility or vessel."[52] This conclusion is firmly buttressed by the legislative history accompanying

---

such compliance, the Administrator . . . shall so notify *the owner and operator* of such facility . . ., and *the owner and operator* of such facility shall continue to be liable. . . .

Id. § 9607(k)(2) (emphasis added).

[48] *Id.* § 9607(a)(2).

[49] 2A SUTHERLAND, STATUTORY CONSTRUCTION § 46.06, at 104 (4th ed. 1984). *See also* United States v. Bear Marine Servs., 509 F. Supp. 710 (E.D. La. 1980) (construing a provision of the Federal Water Pollution Control Act, 33 U.S.C. §§ 1251–1376 (1982)). The court noted that (like CERCLA's legislative history) the Water Pollution Control Act's legislative history was unclear and reflected the "'give and take' of the legislative process, and the imprecision of group-undertaken drafting." *Id.* at 713. The court further stated: "Unfortunately, there is a dearth of legislative history as to Subpart (2). And, yet, as this Court has already observed, it 'has got to mean something.' It is the obligation of this Court to give meaning to each and every portion of a legislative act." *Id.* at 716.

[50] 2A SUTHERLAND, *supra* note 49, at 278.

[51] *See supra* note 15.

[52] 126 CONG. REC. 31,969 (1980) (statement of Rep. Broyhill (R-N.C.)) (emphasis added). *See also supra* notes 17–19.

CERCLA's reauthorization[53] and by the language of the bills that coalesced to form CERCLA.[54]

---

[53] *See, e.g.*, Explanation of Purpose and Intent of the Judiciary Committee Report, 131 CONG. REC. H11,083, H11,086 (daily ed. Dec. 5, 1985):

> [The specified criteria] do not easily apply to a landowner who is liable as an "owner" of a "facility" under section 107(a), but who otherwise may be minimally related to the hazardous substance problem at the facility. Therefore, new subsection 122(g)(1)(B) provides that landowners may qualify for these expedited settlements. The criteria for this type of de minimis settlement require that the potentially responsible [sic] party: (1) own the real property on or in which the facility is located; (2) not have conducted or allowed the generation, transportation, storage, treatment, or disposal of any hazardous substance at the facility; and (3) not have contributed to the release or threatened release through any act or omission.

See also 131 CONG. REC. H11,082 (daily ed. Dec. 5, 1985) (statement of Rep. Rodino (D-N.J.)) ("[T]he committee amendments encourage EPA to settle with—and not to sue— . . . individuals who became owners of land without any knowledge or responsibility for the fact that it contained a hazardous waste site.").

[54] Three bills were the foundation of CERCLA: H.R. 85, 96th Cong., 2d Sess., 126 CONG. REC. 26,369–92 (1980); S. 1480, 96th Cong., 1st Sess., 125 CONG. REC. 17,989 (1979), and H.R. 7020, 96th Cong., 2d Sess., 126 CONG. REC. 26, 757 (1980). For a legislative history of H.R. 85, see Grad, *supra* note 15, at 3–4. For a legislative history of S. 1480, see *id.* at 6–14, 29–35; ENVIRONMENT AND NATURAL RESOURCES POLICY DIVISION, SENATE COMM. ON ENV'T AND PUBLIC WORKS, 97TH CONG., 2D SESS., A LEGISLATIVE HISTORY OF THE COMPREHENSIVE ENVTL. RESPONSE, COMPENSATION, AND LIABILITY ACT OF 1980 (SUPERFUND), PUB. L. 96-510, vol. 1 at 462 (Comm. Print 1983) (October 1, 1980 draft); 126 CONG. REC. 30,916 (1980). For a legislative history of H.R. 7020, see Grad, *supra* note 15, at 4–6, 14–18, 29–35; ENVIRONMENT AND NATURAL RESOURCES POLICY DIVISION, SENATE COMM. ON ENV'T AND PUBLIC WORKS, 97TH CONG., 2D SESS., A LEGISLATIVE HISTORY OF THE COMPREHENSIVE ENVTL. RESPONSE, COMPENSATION, AND LIABILITY ACT OF 1980 (SUPERFUND), PUB. L. 96-510, vol. 2 at 3 (Comm. Print 1983) (as introduced on April 2, 1980); *id.* at 138 (as reported); *id.* at 391 (as passed by the House). For merger of three bills to form CERCLA, see 126 CONG. REC. 30,113 (1980) (statement of Sen. Stafford) ("This compromise [CERCLA] incorporates those parts of S. 1480, H.R. 7020 and H.R. 85 on which there is broad consensus.").

Like CERCLA, H.R. 85 imposed liability for "response costs" on the owner and the operator of a vessel or facility. The bill separately defined "owner" and "operator." H.R. 85, 96th Cong., 2d Sess. § 101(x)–(y), 126 CONG. REC. 26,369–392 (1980). The definition of "operator" clearly indicates that the "owner" and "operator" will be different entities. H.R. 85 defines an "operator" as:

> (1) [I]n the case of a vessel, a charterer by demise or any other person, except the owner who is responsible for the operation, manning, victualing, and supplying of the vessel, or
>
> (2) in the case of a facility, any person, *except the owner*, responsible for the operation of the facility by agreement with the owner. . . .

Id. (emphasis added). A Committee Report accompanying H.R. 85 reiterates this distinction between an "owner" and an "operator": "In the case of a facility, an 'operator' is defined to be a person who is carrying out operational functions for the owner of the facility pursuant to an appropriate agreement." HOUSE COMMITTEE ON MERCHANT MARINE AND FISHERIES, H. R. REP. NO. 172, PART 1, 96th Cong., 1st Sess., *reprinted in* 1980 U.S. CODE CONG. & ADMIN. NEWS, 96th Cong., 2d Sess. 6160, 6182 [hereinafter COMMITTEE REPORT ON H.R. 85, PART 1]. *See also* 1980 U.S. CODE CONG. & ADMIN. NEWS, 96th Cong., 2d Sess. 6174 ("Under the scheme developed in the bill, the owner *or* operator of the vessel or facility . . . is in the best position to prevent and control a polluting discharge." (emphasis added)); 125 CONG. REC. 384–386 (1979) (statement of Rep. Biaggi (D-N.Y.)) ("[T]he bill imposes . . . liability on the owner *or* operator. . . ." (emphasis added)). Furthermore, the Committee Report refers to owners forming "one

Moreover, interpreting subsection 107(a)(1) to include lender/owners not only is consistent with CERCLA's legislative history but also promotes CERCLA's goals. In light of the enormous costs of cleaning hazardous waste sites and of the remedial actions necessary to stem the pollution that has seeped from them, Congress necessarily spread liability widely. Before CERCLA, participants in the hazardous waste industry had avoided internalizing all the costs associated with their operations. Waste site owners and operators did not create leakproof sites because of an ignorance of the effects of improperly disposed waste, a virtual dearth of safe disposal technology, and, importantly, an absence of effective legal sanctions. For similar reasons, waste generators and transporters often did not safely package and ship hazardous materials.

By imposing liability on every person who benefited from the dumping activities, CERCLA forces each to internalize the costs caused by unsafe disposal practices and to use safer disposal methods to avoid future liability. Thus, section 107 imposes liability not only on those persons who benefited directly from

---

of the three major classes (the others being operators and guarantors) subject to liability under the Act." COMMITTEE REPORT ON H.R. 85, PART I, *supra*, at 6181. *Cf.* statement by Representative Harsha (R-Ohio):

> I would like to discuss two provisions for the purpose of establishing legislative history. The first is the definition of "owner" contained in title I of H.R. 85. During consideration of this measure by the Public Works Committee I offered an amendment to clarify the definition, as reported by the Committee on Merchant Marine and Fisheries. This change was necessary because the original definition inadvertently subjected those who hold title to a vessel or facility, but do not participate in the management or operation and are not otherwise affiliated with the person leasing or operating the vessel or facility, to the liability provisions of the bill.
>
> While the Merchant Marine Committee report indicated this situation was not intended the statutory language is unclear. Therefore, I offered clarifying language to truly exempt those who hold title but do not participate in the operation or management activities. My amendment also requires that those that hold title cannot be affilated [sic] in any way with those who lease or charter the vessel or facility. This was done to prevent the establishment of "dummy" corporations, with few assets, which would be the responsible party for the purpose of the act.

126 CONG. REC. 26,210–12 (statement of Rep. Harsha) (1980). In contrast to H.R. 85, S. 1480 simply incorporated by reference the Clean Water Act's definition of "owner or operator," which is reproduced at *supra* note 10. That particular definition is virtually identical to CERCLA's definition, though it does not include CERCLA's security interest exception.

H.R. 7020 does not provide a definition of "owner" or of "operator." Instead, it defined a "responsible party" for purposes of liability as including a person who "owned or operated such site at the time during which it was utilized for the treatment, storage, or disposal of any hazardous waste. . . ." H.R. 7020, § 3041(b)(1), 96th Cong., 2d. Sess., 126 CONG. REC. 26,757 (1980). The bill's use of the disjunctive denotes that different persons may be the owner and the operator and that both classes of persons may be liable under the terms of the bill.

the hazardous dump—the waste generators, transporters, and waste site operators—but also on those persons who benefited indirectly by operation of the waste site, including a site owner who does not directly operate it.[55] Courts have imposed section 107 liability on a lessor whose tenant operated a waste site although the lessor did not participate in its operation.[56] In one section 107 case, the court even held a sublessor liable as an "owner" for damages caused by its sublessee's operations even though a sublessor does not own the land.[57] Although a landlord does not benefit directly from its tenant's waste site operations, it does benefit from rental payments generated by those operations. Holding a landlord liable also removes the economic incentive for a property owner to avoid CERCLA liability by leasing the waste site to a subsidiary or other controlled entity.

The same theories apply to lenders that lend to hazardous waste site owners. If the loan was made to finance the acquisition or operation of the site, the lender's situation is analogous to the landlord's. Just as the landlord benefits from rental payments generated by the hazardous waste operations, so the lender benefits from principal and interest payments generated by operation of the site. Even if the loan was for a purpose unrelated to operation of the site, the operation is aided because the loan frees the operator's capital for use at the site. Furthermore, as with the landlord, holding a lender/owner liable for response costs eliminates the economic incentive for a waste

---

[55] 42 U.S.C. § 9607 (1982 & Supp. II 1984).

[56] Caldwell v. Gurley Refining Co., 755 F.2d 645 (8th Cir. 1985); Sand Springs Home v. Interplastic Corp., 25 Env't Rep. Cas. (BNA) 2127 (N.D. Okla. 1987); United States v. Argent Corp., 14 Envtl. L. Rep. (Envtl. L. Inst.) 20,616 (D.N.M. 1984); NEPACCO, 579 F. Supp. 823, 845 n.26 (W.D. Mo. 1984), *aff'd in part*, 810 F.2d 726 (8th Cir. 1986), *cert. denied*, 108 S. Ct. 146 (1987). The landlord cases have involved § 107(a)(2), 42 U.S.C. § 9607(a)(2) (1982), rather than § 107(a)(1), 42 U.S.C. § 9607(a)(1) (1982). Subsection 107(a)(2) imposes liability on "any person who at the time of disposal of any hazardous substance owned or operated any facility at which such hazardous substances were disposed of. . . ." 42 U.S.C. § 9607(a)(2) (1982).

Courts have been equally expansive in their interpretations of parties who are liable as "operators." Despite customary legal deference to corporate form, courts have held that corporate officers and shareholders can be personally liable for § 107 damages if they were personally involved in the activities creating CERCLA liability. *See, e.g.*, New York v. Shore Realty Corp., 759 F.2d 1032, 1052 (2d Cir. 1985); NEPACCO, 579 F. Supp. at 849; United States v. Carolawn Co., 14 Envtl. L. Rep. (Envtl. L. Inst.) 20,699, 20,700 (D.S.C. 1984); United States v. Wade, 577 F. Supp. 1326, 1341 (E.D. Pa. 1983). Similarly, one court has held that an institutional lender may be liable as an "operator" because of its participation in management of a hazardous waste site. United States v. Mirabile, 15 Envtl. L. Rep. (Envtl. L. Inst.) 20,994, 20,997 (E.D. Pa. 1985).

[57] United States v. South Carolina Recycling & Disposal, Inc., 653 F. Supp. 984, 1003 (D.S.C. 1984).

site operator to acquire title to the site in the name of a shell corporation, to "lend" the shell corporation the necessary operating capital, and then to acquire title to the site in settlement of the loan after dumping activity at the site has terminated.

The lender's CERCLA liability is not triggered by an indirect benefit alone, however. Like the landlord who owns the property, the lender's liability is premised on its ownership of the site. Under the statutory scheme, the lender will be liable only if it enjoys the direct benefit of a cleanup that occurs while it owns the site, if it acquired the site with notice of the waste, or if it attempted to avoid liability by selling the property to an innocent purchaser for a price that presumably represented the fair market value of uncontaminated property.[58] Although the amount of response costs normally will not be fully reflected in the increase in the land's fair market value, its value will undeniably be enhanced. Moreover, such liability is not absolute. As described below, a lender/owner can avoid all liability for response costs if it can establish the third-party defense of subsection 107(b)(3),[59] and a contribution action is available if it cannot avoid liability.[60]

Imposing CERCLA liability on lender/owners also furthers CERCLA's goals by creating substantial incentives for lenders to act as hazardous waste watchdogs. If a lender determines that a parcel of land contains hazardous wastes, a knowledgeable lender should refuse to accept the land as collateral unless the potential borrower proves that the waste is properly contained, even if the lender intends to sell the loan on the secondary mortgage market.[61] The security otherwise is valueless. In

---

[58] Subsection 101(35)(C) provides that a former property owner is liable for response costs if it knew of a release or threatened release before selling the property but transferred ownership "without disclosing such knowledge. . . ." 42 U.S.C.A. § 9601(35)(C) (West Supp. 1987). *See infra* note 140 and accompanying text.

[59] *See infra* notes 115–43 and accompanying text.

[60] *See infra* notes 248–74 and accompanying text.

[61] Like an originating lender, a secondary market buyer must be concerned with potential CERCLA liability. Although secondary market mortgage purchasers are a step removed from the loan transaction and may be unfamiliar with the secured land, they are treated the same as any other lender/owner. Thus, under SARA's "contractual relationship" definition, a secondary market mortgage purchaser that forecloses and acquires title to the encumbered property will be unable to assert the third-party defense if it had actual or constructive notice of the hazardous wastes on the site when it foreclosed. 42 U.S.C.A. § 9601(35)(c) (West Supp. 1987). *See infra* notes 115–35 and accompanying text. During Congress' consideration of SARA, Federal Home Loan Corporation, one of the largest buyers in the secondary market, proposed a definition of "contractual relationship" that would protect a lender/owner if it was unaware of the hazardous wastes on the encumbered land when the loan was made. *See supra* note 17.

the absence of this proof, the lender will refuse to make the loan, will charge a higher interest rate, or will require additional collateral. If the loan is for operating expenses or for acquisition of a waste site, the lender may condition the loan on the borrower's agreement to dispose of hazardous wastes in a safe manner. These results will have the beneficial effects of putting some unsafe operators out of business or, by charging unsafe owners higher interest rates, of forcing them to internalize some of the costs generated by their operations.

The results of lender/owner liability have a darker side, however. For example, if a dump site owner or operator is unable to borrow funds, it may have insufficient capital to clean the site or to upgrade its operations by investing in new disposal technologies. That result is diametrically opposed to CERCLA's aims. Additionally, to satisfy the lender's inquiry concerning the land, the site owner or operator will have an incentive to falsify documents concerning the wastes it has accepted and its disposal methods, a practice that will hinder future cleanup efforts. Furthermore, waste site operators may refuse to accept especially toxic wastes or may charge higher prices for them, thereby increasing the incidence of illicit, "midnight" dumping of the worst forms of hazardous waste. Based on the current state of knowledge concerning CERCLA's impact on lenders' decisionmaking, however, these potential problems provide insufficient reason to exempt lender/owners from CERCLA's liability provisions. To determine whether imposing liability on lender/owners will further CERCLA's goals in the long term would require a detailed economic analysis, which is beyond the scope of this Article.

By its terms, subsection 107(a)(1) imposes liability for response costs on a property owner. Any argument that an owner is liable only if it operated the leaking waste site is negated by reference to the well-established rules of statutory construction, the legislative history of CERCLA, and the economics of hazardous waste site cleanups. Subsection 107(a)(1) affords no relief to an owner merely because it acquired the property in connection with a loan relationship. As described below, however, the Act's definition of an "owner or operator" protects lenders who acquire title to land solely by reason of the loan relationship.

The effect of the potential liability of secondary mortgage market purchasers will be to reduce the flow of capital from the secondary market to the loan originators.

B. *Subsection 101(20)(A): Security Interest Exception*

The second argument lender/owners have used to attempt to avoid CERCLA liability is based on the Act's definition of "owner or operator." Subsection 101(20)(A) broadly defines the terms "owner or operator" to include any person owning or operating a vessel or facility.[62] The definition expressly excludes "a person, who, without participating in the management of a vessel or facility, holds indicia of ownership primarily to protect his security interest in the vessel or facility."[63] A person who satisfies the terms of this security interest exception is not liable as an owner or operator pursuant to subsection 107(a)(1).[64] Not surprisingly, lender/owners sued for response costs have argued that they acquired the contaminated property to protect their security interests and, therefore, are not liable.[65] This argument is flawed in two fundamental ways.

First, the security interest exception applies only while a person holds mere "indicia of ownership." When a lender acquires the waste site, it owns a qualitatively different interest in the land than mere indicia of ownership. Examination of the purpose for the security interest exception and of the mortgage law context in which it exists reveals that the exception is directed to this qualitative difference between ownership of land and of a mortgage. The purpose of the exception is to ensure that mortgage holders are treated similarly under the Act despite differing state law treatments of the interests created by a mortgage.

In the majority of states, those denominated lien theory states, state law characterizes a mortgage as conveying no title to the land.[66] Rather, the mortgage in these states creates only the right to sell the encumbered land in case of default.[67] A mortgage

---

[62] 42 U.S.C.A. § 9601(20)(A) (West Supp. 1987). The subsection is reproduced at *supra* note 38.

[63] *Id.* The Act does not similarly except from the definition of owner a trustee of a land trust. In some states, property often is held in a land trust, with the trustee holding legal title. Angelo & Bergeson, *supra* note 11, at 107 n.13a. Although CERCLA does not expressly except the trustee from liability, a trustee has a strong argument that it does not "own" the property. A trustee with active involvement in the property's management, however, may be subject to liability.

[64] 42 U.S.C. § 9607(a)(1) (1982).

[65] United States v. Maryland Bank & Trust Co., 632 F. Supp. 573, 578–80 (D. Md. 1986); United States v. Mirabile, 15 Envtl. L. Rep. (Envtl. L. Inst.) 20,994, 20,996 (E.D. Pa. 1985).

[66] Burkhart, *Freeing Mortgages of Merger*, 40 VAND. L. REV. 283, 327 (1987).

[67] *Id.* at 321–27.

holder in these states would not be an owner under subsection 107(a)(1).[68] In a few other states, denominated title theory states, however, state law characterizes a mortgage as actually conveying title to the encumbered land to the lender.[69] Normally, a lender in a title theory state will not take possession of the property or otherwise hold itself out as being the owner of the property, but the mortgage, by its terms or by operation of common law, nevertheless invests the lender with indicia of ownership. If CERCLA equated title with ownership, similarly situated mortgage holders would be liable as owners in some states but not in others. The security interest exception eliminates this potential for unequal treatment. As long as a mortgage holder has no greater interest in the encumbered land than the mortgage, it will not be liable as an owner.

Although the distinction between title and indicia of ownership is a technical one, examination of the language of subsection 101(20)(A) as a whole reveals that Congress intended to draw this distinction. The first sentence of the "owner or operator" definition provides that an owner is a person "owning" the affected property or who "owned" the property in the case of abandoned property, whereas the security interest exception provided in the next sentence of the definition applies to a person who "holds indicia of ownership."[70] A reference to holding only indicia of ownership indicates that, although the holder possesses an instrument indicating a conveyance of title to it, such ownership is not ownership in the usual sense but is given to serve another purpose—in this case, to secure a loan. The logical conclusion is that the security interest exception applies only to holders of security interests and not to persons who own property.

The legislative histories of the three bills that formed the foundation of CERCLA further support the conclusion that Congress intended to distinguish between ownership for all purposes and title as mere indicia of ownership. H.R. 85, the Comprehensive Oil Pollution Liability and Compensation Act,[71] for example, includes a definition of owner that is similar to that contained in CERCLA but with an important difference. The references in CERCLA to ownership and to holding indicia of

[68] 42 U.S.C. § 9607(a)(1) (1982).
[69] Burkhart, *supra* note 66, at 328.
[70] 42 U.S.C. § 9601(20)(A) (1982).
[71] H.R. 85, 96th Cong., 1st Sess., 126 CONG. REC. 26,369–92 (1980).

ownership are contained in separate sentences, thereby creating a possible inference that Congress' reference to indicia of ownership in the second sentence was a drafting error. The definition in H.R. 85, however, demonstrates that Congress was cognizant of the difference between the terms. "Owner" is defined in H.R. 85 as:

> [A]ny person holding title to, *or, in the absence of title, any other indicia of ownership* of, a vessel or facility, but does not include a person who, without participating in the management or operation of a vessel or facility, *holds indicia of ownership* primarily to protect his security interest in the vessel or facility.[72]

The Committee Reports concerning H.R. 85 echo this distinction between title and indicia of ownership in their discussions of the "owner" definition.[73] Similarly, the two other bills that eventually joined to form CERCLA each contained a definition of owner that reflects the distinction between ownership and possession of mere indicia of ownership.[74] The distinction undercuts the lender/owner's first argument that it is protected from CERCLA liability based on the security interest exception. Once the lender/owner acquires the property at a foreclosure sale or by voluntary settlement, it no longer holds mere indicia of ownership; instead, it owns the property and is liable as an owner.

Notwithstanding the legislative history of the "owner or operator" definition, the phrase "primarily to protect his security interest" provides lender/owners sufficient room for a second argument against CERCLA liability based on subsection 101(20)(A).[75] They may argue that the statutory exception is broad enough to include a lender that acquires the land to ensure repayment of the loan. When a lender forecloses its security interest and purchases the property at the foreclosure sale or when it otherwise acquires title in satisfaction of the secured debt, it usually does so "primarily to protect" its interest in the property.[76] Because most lenders are not in the business of

---

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

[73] HOUSE COMMITTEE ON PUBLIC WORKS & TRANSPORTATION, H.R. REP. NO. 172, PART 2, 96th Cong., 2d Sess., *reprinted in* 1980 U.S. CODE CONG. & ADMIN. NEWS 6212, 6213; COMMITTEE REPORT ON H.R. 85, PART 1, *supra* note 54, at 6181.

[74] H.R. 7020, 96th Cong., 2d Sess., 126 CONG. REC. 26,757–99 (1980); S. 1480, 96th Cong., 1st Sess. (1979), 126 CONG. REC. 30,898–987 (1980).

[75] 42 U.S.C. § 9601(20)(A) (1982).

[76] Burkhart, *supra* note 66, at 331–32.

acquiring land and managing it for its investment potential, lenders normally resort to foreclosure or accept encumbered land in settlement of the debt only when the debtor is unable to repay the loan or when state law prohibits a lender from pursuing personal liability until it has attempted to satisfy the debt from the collateral securing it. Particularly when the defaulting debtor has abandoned the property or is mismanaging it, a lender may conclude that unless it intervenes to put the property ownership in new hands, the value of the lender's security interest and the lender's chances of repayment in full from a sale of the land will also deteriorate.

While the argument may be consistent with a lay person's first reading of the statute, it is contradicted by the statute's wording and legislative history,[77] which demonstrate that Congress did not use the phrase "primarily to protect his security interest" to refer to the lender's reasons for acquiring title to the encumbered land. Instead, the phrase refers to the lender's reasons for accepting a mortgage in those jurisdictions that treat a mortgage as conveying indicia of ownership to the encumbered land. The placement of the phrase in the statute strongly supports this conclusion: "'[O]wner or operator'. . . does not include a person, who . . . holds indicia of ownership primarily to protect his security interest. . . ."[78] As discussed above, the term "indicia of ownership" refers in this context to a mortgage interest in the title theory states. As can be seen, the "primarily" phrase specifies when a mortgage holder will be free of liability as an owner. As an exception to a statutory imposition of liability, the security interest exception must be strictly construed.[79] Therefore, once a lender holds more than mere indicia of ownership, as it does when it buys the property, it no longer qualifies for the exception.

The legislative history of subsection 101(20)(A) further demonstrates that the particular placement of the "primarily" clause was not the result of imprecise drafting; a Committee Report concerning H.R. 85, which included a definition of owner similar to CERCLA's, states that the term "owner" does not include "certain persons possessing indicia of ownership (such as a financial institution) who, without participating in the manage-

---

[77] *See supra* note 54 (sources of CERCLA).
[78] 42 U.S.C. § 9601(20)(A) (1982).
[79] *See* 2A SUTHERLAND, *supra* note 49, § 47.11, at 144.

ment or operation of a vessel or facility, hold title either in order to secure a loan or in connection with a lease financing arrangement under the appropriate banking laws, rules, or regulations."[80] Furthermore, in describing the security interest exception, the Report states: "[A] financial institution which held title primarily to secure a loan but also received tax benefits as the result of holding title would not be an 'owner' as long as it did not participate in the management or operation of the vessel or facility."[81] Therefore, the measure of the security interest exception is the lender's purpose for taking indicia of ownership as security for the loan, rather than the lender's reasons for subsequently acquiring the land.

Finally, the lender/owner's argument that it acquired the encumbered property primarily to protect its security interest is inconsistent with the general concepts traditionally and currently employed in real estate finance law. When a mortgagee acquires the encumbered property, it does not do so to protect its security interest. That interest is protected by recording it in the public property records in accordance with the terms of the state's recording act. Instead, the lender acquires the property to recover the outstanding loan amount. If the lender acquired the property by a voluntary conveyance, usually by a deed in lieu of foreclosure, the lender has agreed to accept the land in full or partial satisfaction of the debt. The lender's ownership of the land has no preservative effect on the security interest. On the contrary, by acquiring the property title, the lender risks losing its security interest through operation of the doctrine of merger.[82]

The lender/owner's argument is attenuated even further if the lender acquired the property by foreclosing on a mortgage and buying at the sale, rather than by voluntary conveyance. Because a foreclosure sale extinguishes the lender's security interest, a lender/owner simply will not possess a security interest at the time of the cleanup. The security interest exception, however, is worded in the present tense: "Such term ['owner or operator'] does not include a person, who . . . *holds* indicia of ownership primarily to protect his security interest. . . ." Thus, to qualify for the security interest exception, that interest must

---

[80] COMMITTEE REPORT ON H.R. 85, PART 1, *supra* note 54, at 6181.
[81] *Id.*
[82] Burkhart, *supra* note 66, at 331–52.

exist when the cleanup occurs.[83] When the lender has foreclosed, the security interest is extinguished and so is the protection of the exception. In sum, the lender's acquisition of the property protects the lender's source of payment, not its security interest.

Although the above arguments may appear to be unduly technical in light of CERCLA's drafting deficiencies, SARA's amendments corroborate the limits of the exception. SARA amended the definition of "owner or operator" to provide that if the facility is acquired by a unit of state or local government due to a "bankruptcy, foreclosure, tax delinquency, abandonment, or similar means," the term "owner or operator" means the person who owned, operated, or otherwise controlled the facility immediately before the foreclosure.[84] Congress did not provide a similar exception for lender/owners, although the business journals had highlighted this issue before SARA's enactment.[85]

Final confirmation that a lender/owner is not included within the security interest exception is supplied by use of the term "owner or operator" throughout the Act in contexts that would cause such a limitation of liability to violate CERCLA's statutory goals. A determination that a lender/owner is protected from section 107[86] liability because of the security interest exception would also except the lender/owner from all other CERCLA provisions that otherwise apply to an owner or operator and that should apply to a lender/owner. For example, subsection 104(e) provides that a government agent who takes samples of hazardous substances from a property must "give to the owner, operator, or person in charge a receipt describing the sample obtained and if requested a portion of each such sample equal in volume of weight to the portion retained."[87] Certainly, a lender/owner, no less than any other variety of owner, should have the advantage of this statutory right to a receipt and to a portion of any sample taken. Similarly, subsection 111(g) requires an owner or operator to provide notice to potential injured parties of a hazardous substance that has been released

---

[83] United States v. Maryland Bank & Trust Co., 632 F. Supp. 573, 579 (D. Md. 1986).
[84] 42 U.S.C.A. § 9601(20)(A)(iii) (West Supp. 1987).
[85] *See, e.g.*, Angelo & Bergeson, *supra* note 11, at 111–12; Burcat, *supra* note 13, at 513–15.
[86] 42 U.S.C. § 9607 (1982 & Supp. II 1984).
[87] 42 U.S.C. § 9604(e)(1)(B) (1982).

from the vessel or facility.[88] Imposition of this responsibility on a lender/owner is necessary because, as the person in possession of the vessel or facility, the lender/owner is in the best position to discover such discharges, to analyze them, and to notify potentially affected parties. If a lender/owner were excepted from this requirement by virtue of the security interest exception, the notice requirement of subsection 111(g) would be nullified because it does not impose this duty on any other person. For public safety reasons, the lender/owner must be subject to this statutory responsibility. The lender/owner, therefore, must be excluded from the security interest exception to preserve the Act's balance.

The security interest exception was the primary focus of both reported decisions involving a response cost recovery action against a lender/owner.[89] Predictably, the lender/owner in each case argued that it was not liable because of the security interest exception. One federal district court agreed with the defendant's argument,[90] and the other federal district court indicated that on a different set of facts it also would hold that the security interest exception insulates a lender/owner from liability.[91]

In the first of these cases decided, *United States v. Mirabile*, American Bank and Trust Company (ABT) had financed a paint manufacturing business. The loan was secured, in part, by a mortgage on the manufacturing site. When the loan went into default, ABT foreclosed and was the highest bidder at the sheriff's sale. Rather than accept a sheriff's deed to the property, however, ABT searched for a purchaser for its interest and, four months after the sale, assigned its interest in the site to the Mirabiles.[92]

---

[88] *Id.* § 9611(g).

[89] An earlier case, In re T.P. Long Chem., Inc., 45 Bankr. 278 (N.D. Ohio 1985) has been described as stating in dictum that a lender is not an owner subject to CERCLA liability if it acquires the property in which it has a security interest. *See, e.g., Fear of Foreclosure: United States v. Maryland Bank & Trust Co.*, 16 Envtl. L. Rep. (Envtl. L. Inst.) No. 7, at 10,165, 10,165 n.5 (July 1986). *In re T.P. Long*, however, does not support that proposition. Instead, the court stated in dictum that if the holder of a security interest in personal property repossessed the property for sale pursuant to its security agreement without acquiring title, the security interest holder would not be an owner as defined by CERCLA. This dictum is a correct interpretation of CERCLA. If the lender does not acquire title to the contaminated property, whether real or personal, it is not an owner.

[90] United States v. Mirabile, 15 Envtl. L. Rep. (Envtl. L. Inst.) 20,994 (E.D. Pa. 1985).

[91] United States v. Maryland Bank & Trust Co., 632 F. Supp. 573 (D. Md. 1986).

[92] *Mirabile*, 15 Envtl. L. Rep. (Envtl. L. Inst.) at 20,996.

Approximately fourteen months after the Mirabiles acquired the property, an EPA agent inspected it and discovered some 550 drums of waste from the paint manufacturing operation, many of which were in a deteriorated condition. When the Mirabiles failed to respond to an EPA cleanup notice, EPA cleaned the site. EPA then sued the Mirabiles to recover its response costs, and the Mirabiles joined ABT and others as third-party defendants. ABT moved for summary judgment on the grounds that: (1) it was never an owner because it acquired only equitable title to the property at the foreclosure sale and never acquired legal title by accepting a sheriff's deed; and (2) it foreclosed and took steps to secure the property after the foreclosure sale solely to protect its security interest.[93]

ABT's argument that it was never an owner of the property within the meaning of section 107[94] was weak. ABT conceded that it had acquired equitable title to the property, pursuant to which it took possession of the property. Moreover, no other person could properly be denominated the owner of the property. Under the relevant state foreclosure law, neither the former owner nor any junior lienor retained a right to redeem the property title after the sale.[95] Furthermore, state law required the sheriff who conducted the foreclosure sale to perfect title in the foreclosure sale purchaser by issuing a deed within ten days after the sale.[96] ABT successfully delayed issuance of the deed by advising the sheriff and the local tax department that "it intended to take title to the property."[97] Considering ABT's knowledge of the waste on the property and of the cost of disposing of it,[98] ABT may have been attempting to avoid CERCLA liability by avoiding legal title to the property.

Although the court recognized that the property title ABT acquired at the foreclosure sale could be sufficient to bring it within the scope of CERCLA's liability provisions, the court determined that the original security interest controlled regardless of the subsequent acquisition of title:

---

[93] *Id.*

[94] 42 U.S.C. § 9607 (1982).

[95] *See* In re Rouse, 48 Bankr. 236, 239 (E.D. Pa. 1985) (". . . [A]ccording to Pennsylvania law, the debtor's right to redeem the mortgage terminated when the sheriff's sale took place."); G. NELSON & D. WHITMAN, REAL ESTATE FINANCE LAW § 8.4, at 616 n.1 (2d ed. 1985).

[96] 42 PA. CONS. STAT. ANN. § 3135 (Purdon 1987). The only statutory exception is if a "petition has been filed to set aside the sale."

[97] *Mirabile*, 15 Envtl. L. Rep. (Envtl. L. Inst.) at 20,996.

[98] *Id.*

> I need not resolve the issue of whether, under Pennsylvania law, ABT's successful bid at the sheriff's sale technically vested ABT with ownership as defined by the statute. Regardless of the nature of the title received by ABT, its actions with respect to the foreclosure were plainly undertaken in an effort to protect its security interest in the property. ABT made no effort to continue [the paint manufacturing] operations on the property, and indeed foreclosed some eight months after all operations had ceased.
>
> . . . [I]n enacting CERCLA Congress manifested its intent to impose liability upon those who were responsible for and profited from improper disposal practices. Thus, it would appear that before a secured creditor such as ABT may be held liable, it must, at a minimum, participate in the day-to-day operational aspects of the site. In the instant case, ABT merely foreclosed on the property after all operations had ceased and thereafter took prudent and routine steps to secure the property against further depreciation.[99]

In effect, the court's reasoning eliminates the statutory imposition of liability on the person who owns the property when the cleanup occurs. The court's reasoning is flawed in three ways. First, under the statutory scheme, the current owner is liable not because it is "responsible for and profited from improper disposal practices," but because it benefits from the cleanup through the increase in the property's value.[100] Second, nothing in the legislative history justifies a conclusion that Congress intended to treat lender/owners differently than other innocent purchasers. In fact, a lender such as ABT that provided financing for the waste generating activity is a less sympathetic defendant than a purchaser who did not contribute in any way to the operation and, indeed, may have acquired the land without knowledge of the hazardous wastes on the property. Finally, as developed below,[101] a landowner has a common law responsibility to correct hazardous conditions on its property regardless of whether it caused or contributed to the hazard. The security interest exception to liability properly does not protect lender/owners.

Ironically, the court in *Mirabile* had an obvious and incontrovertible ground for granting ABT's motion for summary judgment: ABT did not own the site during the EPA cleanup. ABT

---

[99] *Id.*

[100] *See supra* note 26.

[101] *See infra* notes 190–237 and accompanying text.

assigned its property interest to the Mirabiles on December 15, 1981, and an EPA official first visited the property in February 1983.[102] No court or commentator has previously identified this particular flaw in the court's reasoning. In fact, *United States v. Maryland Bank & Trust Company*[103] indicates that the seed planted by the *Mirabile* court may mature into a noxious weed.

In *Maryland Bank & Trust*, Maryland Bank & Trust (MB&T) lent money to the owner and operator of a trash and garbage dump site. The borrower ran the business on property owned jointly with his spouse. According to the district court's factual summary, MB&T knew the nature of the borrower's business; but the record does not reveal whether MB&T obtained this information before or after making the loan.[104] When the owners sold the property to their son, MB&T made him a $335,000 purchase money loan. The son defaulted on the loan, and MB&T foreclosed on the property and purchased it at the sale. Eleven months later, the son reported the existence of hazardous wastes on the site to the county health department, which then notified EPA. Some seventeen months after MB&T had acquired the property for $381,500, EPA cleaned it at a cost of approximately $551,713.50 and demanded payment from MB&T. When MB&T failed to pay, EPA brought suit.[105]

Among other defenses, MB&T argued that it was not liable pursuant to the security interest exception. The court rejected this defense for three reasons. First, the court recognized that the foreclosure sale extinguished MB&T's security interest and that, therefore, MB&T did not possess "indicia of ownership primarily to protect [its] security interest" when the EPA cleanup occurred.[106] Second, the court cited legislative history to demonstrate that Congress created the security interest exception for the limited purpose of insuring that similarly situated lenders are treated equally regardless of differing state law rules

---

[102] *Mirabile*, 15 Envtl. L. Rep. (Envtl. L. Inst.) at 20,993. Although EPA did not inspect the property until 1983, the Pennsylvania Department of Environmental Resources informed Mr. Mirabile "in the winter of 1981–1982" that drums on the property contained hazardous waste and requested that he remove them. Nothing in the opinion indicates, however, that EPA or the state agency ever contacted ABT about the property, either before or after ABT assigned its interest in the land to the Mirabiles.

[103] 632 F. Supp. 573 (D. Md. 1986).

[104] *Id.* at 575.

[105] *Id.* at 575–76.

[106] *Id.* at 579.

of mortgage and title.[107] Finally, the court recognized that MB&T would be unjustly enriched if it were not held liable for EPA's response costs:

> Under the scenario put forward by the bank, the federal government alone would shoulder the cost of cleaning up the site, while the former mortgagee-turned-owner, would benefit from the clean-up by the increased value of the now unpolluted land. At the foreclosure sale, the mortgagee could acquire the property cheaply. All other prospective purchasers would be faced with potential CERCLA liability, and would shy away from the sale. Yet once the property has been cleared at the taxpayers' expense and becomes marketable, the mortgagee-turned-owner would be in a position to sell the site at a profit.
>
> In essence, [MB&T's] position would convert CERCLA into an insurance scheme for financial institutions, protecting them against possible losses due to the security of loans with polluted properties.[108]

Unfortunately, the court did not stop its analysis there but continued by discussing *Mirabile*. Despite the *Maryland Bank & Trust* court's strong statements of the policy reasons for holding lender/owners liable under CERCLA, the court indicated that it would not hold a lender/owner liable if it had owned the contaminated property for only a few months before the cleanup. Like the *Mirabile* court, the *Maryland Bank & Trust* court failed to recognize that the lender in *Mirabile* was not liable for response costs simply because it was not the owner during the cleanup. Instead, the court characterized the *Mirabile* holding as turning on the lender/owner's sale of the property four months after the foreclosure sale:

> The [*Mirabile*] court found that the mortgagee's purchase of the land at the foreclosure was plainly undertaken in an effort to protect its security interest in the property. That holding pertained to a situation in which the mortgagee-turned-owner promptly assigned the property. *To the extent to which that opinion suggests a rule of broader application, this Court respectfully disagrees.*[109]

This statement implies that the *Maryland Bank & Trust* court agreed with the *Mirabile* holding on its facts. The conclusion is

---

[107] *Id.* at 580.
[108] *Id.*
[109] *Id.* (emphasis added).

reinforced by the court's further statement: "The [security interest] exclusion does not apply to former mortgagees currently holding title after purchasing the property at a foreclosure sale, *at least when, as here, the former mortgagee has held title for nearly four years, and a full year before the EPA cleanup.*"[110] Again, this statement implies that the court might hold that the security interest exception protects a lender if it purchased the property less than a year before the EPA cleanup.

The *Maryland Bank & Trust* court, however, expressly reserved the issue of whether a lender can qualify for the security interest exception after acquisition of contaminated property.[111] Its discussion in dictum of the CERCLA liability of a lender/owner that owns for less than a year, however, is troubling. Although the court clearly recognized the reasons for denying lender/owners the protection of the security interest exception, it indicated that it would follow the *Mirabile* court's lead when deciding a case in which the lender/owner held title for several months. Such a decision would interfere with CERCLA's legislative scheme by creating a class of landowners who could avoid liability for response costs and other obligations CERCLA imposes on landowners for the public safety.

Regardless of how long a lender/owner owns land before a hazardous waste cleanup, the lender does not qualify for the security interest exception in subsection 101(20)(A).[112] Once the lender acquires the land, it no longer holds mere indicia of ownership. Furthermore, a lender's argument that it acquired the property primarily to protect its security interest is unavailing. That statutory phrase refers not to the lender's acquisition of the land but to its reasons for accepting a mortgage. Thus, the security interest exception serves the very limited purpose of ensuring that CERCLA treats secured lenders equally despite differences in state law mortgage characterizations.[113] Although the security interest exception provides no relief to a lender/owner, the third-party defense of subsection 107(b)(3)[114] offers a safe harbor, albeit a narrow one.

---

[110] *Id.* at 579 (emphasis added).
[111] *Id.* at 579 n.5.
[112] 42 U.S.C. § 9601(20)(A)(iii) (1982).
[113] *See supra* notes 66–69 and accompanying text.
[114] 42 U.S.C. § 9607(b)(3) (1982).

C. *Subsection 107(b)(3): Third-Party Defense*

Although a lender/owner is potentially liable for response costs as an owner under subsection 107(a)(1), it still may avoid liability even if the cleanup occurs during its ownership. Congress provided three affirmative defenses to liability. Subsection 107(b)(1)–(3) provides that an otherwise responsible party will avoid liability if it can prove that the hazardous waste release or threat of release and the resulting damages were caused solely by (1) an act of God, (2) an act of war, or (3) an act or omission of a third party.[115] Not surprisingly, litigation concerning the subsection 107(b) defenses has centered on the third-party defense.

Originally, to establish the third-party defense, an owner had to prove only three facts. First, the owner had to establish that the third party's actions were the sole cause of the hazardous condition.[116] The CERCLA legislative scheme requires that anyone that contributed to a hazardous waste problem share in the cost of correcting it. Second, the owner had to establish that it exercised due care with respect to the hazardous substances on its land.[117] Apparently, Congress did not intend to relieve an owner qualifying for the third-party defense of any duty to exercise due care with respect to the hazardous waste.[118] Third, the owner had to establish that the third party was not the

---

[115] Subsection 9607(b) provides:

> There shall be no liability under subsection (a) of this section for a person otherwise liable who can establish by a preponderance of the evidence that the release or threat of release of a hazardous substance and the damages resulting therefrom were caused solely by—
>
> (1) an act of God;
>
> (2) an act of war;
>
> (3) an act or omission of a third party other than an employee or agent of the defendant, or than one whose act or omission occurs in connection with a contractual relationship, existing directly or indirectly, with the defendant (except where the sole contractual arrangement arises from a published tariff and acceptance for carriage by a common carrier by rail), if the defendant establishes by a preponderance of the evidence that (a) he exercised due care with respect to the hazardous substance concerned, taking into consideration the characteristics of such hazardous substance, in light of all relevant facts and circumstances, and (b) he took precautions against foreseeable acts or omissions of any such third party and the consequences that could foreseeably result from such acts or omissions; or
>
> (4) any combination of the foregoing paragraphs.

Id. § 9607(b) (1982 & Supp. II 1984).

[116] *Id.*

[117] *Id.*

[118] INTERSTATE AND FOREIGN COMMERCE COMMITTEE, H.R. REP. 1016, PART 1, 96th Cong., 2d. Sess., *reprinted in* 1980 U.S. CODE CONG. & ADMIN. NEWS, 6119, 6137.

owner's employee, agent, or "one whose act or omission occurs in connection with a contractual relationship, existing directly or indirectly, with the [owner] (except where the sole contractual arrangement arises from a published tariff and acceptance for carriage by a common carrier by rail). . . ."[119] By providing, in effect, an exception to the exception, Congress prevented an owner from escaping CERCLA liability by the simple expedient of contracting with a third party.

Before SARA, CERCLA did not include a definition of the type of contractual relationship that would defeat the third-party defense. According to authorities on statutory construction, the proper construction of the term in this context is that an owner cannot avoid liability if the third party was an independent contractor, just as the owner cannot avoid liability if the third party was the owner's agent or employee.[120] This interpretation of the contractual relationship provision is supported by the parenthetical provision following it concerning common carrier contracts. The term "contractual relationship" also properly applies to lease agreements. A landlord cannot avoid liability for its tenant's dumping activities based on the third-party defense.[121] Interpreting the term "contractual relationship" to apply to agreements with independent contractors or with tenants is consistent with the common-law concept of a landowner's nondelegable duties.[122]

Although CERCLA originally did not include a definition of contractual relationship for the third-party defense, SARA provides a definition of that term in a lengthy new definitional subsection 101(35).[123] Consistent with Congress' determination to spread response costs over the widest possible base and to encourage property owners to clean their waste sites, SARA's contractual relationship definition restricts the availability of the

---

[119] 42 U.S.C. § 9607(b)(3) (1982).

[120] *See, e.g.*, 2A SUTHERLAND, *supra* note 49, § 47.17, at 166 (defining *ejusdem generis*: "Where general words follow specific words in a statutory enumeration, the general words are construed to embrace only objects similar in nature to those enumerated by the specific words.").

[121] New York v. Shore Realty Corp., 759 F.2d 1032, 1048–49 (2d Cir. 1985) (lessor of a warehouse where lessee is releasing hazardous waste is a potentially liable party under CERCLA); United States v. Argent Corp., 14 Envtl. L. Rep. (Envtl. L. Inst.) 20,616, 20,616 (D.N.M. 1984) (defendant may not assert the third-party affirmative defense in connection with activities done by tenants of its property).

[122] RESTATEMENT (SECOND) OF TORTS § 416 (1979); W.L. PROSSER & W.P. KEETON, TORTS § 71, at 511–15 (5th ed. 1984).

[123] 42 U.S.C.A. § 9601(35) (West Supp. 1987).

third-party defense to the most innocent owners and creates new incentives for private cleanups. Before amendment, CERCLA did not impose liability on an innocent owner who acquired the contaminated property after dumping stopped and who sold the property before the cleanup. New subsection 101(35) virtually has eliminated that safe harbor for an innocent lender/owner by providing circumstances under which a former property owner is liable for response costs if it had actual or constructive knowledge of the hazardous condition—even if it did not contribute to the contamination by action or by inaction.

Subsection 101(35)(A) defines a contractual relationship to include "land contracts, deeds or other instruments transferring title or possession."[124] The third-party defense therefore now applies only if the release or threatened release and resulting damage are caused solely by "an act or omission of a third party other than . . . one whose act or omission occurs in connection with [land contracts, deeds or other instruments transferring title or possession or other types of contractual relationships], existing directly or indirectly, with the defendant."[125] Although awkwardly connected to the language of the third-party defense, subsection 101(35)(A) confirms the judicial holdings that a waste site owner cannot avoid liability by leasing the property or otherwise transferring possession to a third party.[126]

Subsection 101(35)(A) takes a significant step beyond this accepted principle, however, by providing that a transfer of title, rather than of the mere right to possess, can defeat the third-party defense.[127] Again, the reference in subsection 101(35)(A) to a deed is awkwardly connected. Rarely, if ever, will a third party's dumping activities "occur . . . in connection with" a deed or other instrument transferring title. Subsections 101(35)(A)(i), 101(35)(B), and 101(35)(C), however, illustrate the meaning of this reference to title conveyances by imposing two significant new conditions to the third-party defense.

New subsections 101(35)(A)(i) and 101(35)(B) limit the third-party defense by imposing substantial requirements on a defendant when it acquires the property.[128] These subsections provide that a defendant qualifies for the third-party defense only if it

---

[124] *Id.* § 9601 (35)(A).
[125] 42 U.S.C. § 9607(b)(3) (1982); 42 U.S.C.A. § 9601(35)(A) (West Supp. 1987).
[126] *See supra* cases cited in notes 56–57.
[127] 42 U.S.C.A. § 9601(35)(A) (West Supp. 1987).
[128] *Id.* § 9601(35)(A)(i), (B).

can establish that, before it acquired the property, it conducted "all appropriate inquiry"[129] and "did not know and had no reason to know"[130] of the existence on the site of the leaking hazardous substances. Ironically, responsible business practices will usually preclude a lender/owner from proving that it satisfied these requirements, especially in light of the factors that a court is statutorily directed to consider in determining whether a defendant is eligible for the protection of the third-party defense.[131]

Commercial lenders virtually always require a title examination for any property offered as collateral. Although title examination reports normally are limited to a statement of the current state of title, the title examiner preparing the report generally examines the prior links in the chain of title. Therefore, information about prior owners and users would be available to the lender, enabling it to assess the likelihood that the property had been used in the past as a waste site or for a hazardous waste generating business, such as paint production. Furthermore, lenders extending loans to commercial borrowers often investigate the borrower's business operations in order to tailor the loan agreement to the borrower's particular need. Thus, the lender would have notice of the nature of the borrower's business, including the potential for storing or disposing hazardous waste on the property. Lenders also commonly employ inspectors to examine the property, particularly in connection with construction loans. The inspector should discover any visible signs of hazardous waste on the property. Finally, lenders often require a borrower to execute an affidavit at the loan closing certifying facts concerning the condition of the property and its title. Thus, of all potential property purchasers, a lender is among the most able to discover the potential for hazardous waste contamination.[132]

The third-party defense, however, will not necessarily protect a lender/owner even if it did not discover the existence of hazardous wastes in the course of normal pre-loan investigations.

---

[129] *Id.* § 9601(35)(B).

[130] *Id.* § 9601(35)(A)(i).

[131] Subsection 101(35)(B) directs the court to consider: (1) the defendant's specialized knowledge or experience; (2) the relationship of the purchase price the defendant paid to the value of the property if uncontaminated; (3) any commonly known or reasonably ascertainable information about the property; (4) the obviousness of the presence or likely presence of contamination at the property; and (5) the ability to detect such contamination by appropriate inspection. *Id.* § 9601(35)(B).

[132] *See generally* M.R. FRIEDMAN, CONTRACTS AND CONVEYANCES OF REAL PROPERTY (4th ed. 1984).

Despite lenders' arguments to the contrary, EPA has taken the position that a lender/owner will not satisfy the inquiry duty of subsection 101(35)(B) merely by investigating the property before making the loan. The lender/owner will qualify for the third-party defense only if it did not have actual or constructive notice when it acquired the property.[133] This position is firmly grounded in the statutory language.[134]

Moreover, subsection 101(35)(B) provides that a court determining whether a defendant qualifies for the third-party defense must consider whether the contamination could have been discovered by "appropriate inspection." This standard apparently requires an investigation that is more focused on determining the existence of hazardous waste on the site than a lender's normal loan investigation. The most "appropriate inspection" for determining the existence of hazardous waste contamination is by specifically testing the property for such contamination. Predictably, as a direct result of CERCLA and now of SARA, the number of private environmental inspectors is rapidly increasing. Additionally, lenders' attorneys are advising their clients to include such inspections as part of their routine loan processing and pre-foreclosure procedures.[135]

The lender's usual method of acquiring the property, foreclosure, enhances the risk that it will not qualify for the third-party defense. Subsection 101(35)(C) directs a court ruling on the availability of the third-party defense to examine the purchase price paid in relation to the property's value if uncontaminated. If the purchaser paid significantly less than the uncontaminated property value, the apparent presumption is that it had actual knowledge of the contamination or should have been alerted to the fact that something was amiss. Foreclosing lenders may run afoul of this provision because foreclosure sale prices often are substantially lower than the property's fair market value. Although that fact reduces the probative value of the price paid as an indicator of the purchasing lender's actual or constructive

---

[133] Leifer Paper, *supra* note 26, at 4.

[134] Subsection 101(35)(B) provides:

> To establish that the defendant had no reason to know, as provided in [subsection 101(35)(A)(i)], the defendant must have undertaken, at the time of acquisition, all appropriate inquiry into the previous ownership and uses of the property consistent with good commercial or customary practice in an effort to minimize liability.

42 U.S.C.A. § 9601(35)(B) (West Supp. 1987).

[135] *See, e.g.*, Angelo & Bergeson, *supra* note 11, at 101; Richman & Stukane, *supra* note 13, at 13; Shea, *supra* note 13, at 11.

knowledge of the property contamination, a court may not be inclined to rule in favor of the lender on this basis because of Congress' express intent to assess response costs against those economically benefited by the cleanup.

Perhaps the most serious obstacle for a lender/owner to surmount, however, in its attempt to claim the protection of the third-party defense is the subsection 101(35)(B) provision that the court, in determining whether the lender had reason to know of the contamination, consider whether the defendant has "any specialized knowledge or experience."[136] The court's analysis in *Maryland Bank and Trust*[137] does not bode well for lender/owners in this regard, as the court strongly indicated that it would hold a lender/owner to a high standard in determining whether the lender/owner had sufficient reason to know of the contamination to prevent it from asserting the third-party defense. In discussing the policy reasons for imposing liability on lender/owners, the court stated: "Mortgagees . . . have the means to protect themselves, by making prudent loans. Financial institutions are in a position to investigate and discover potential problems in their secured properties. For many lending institutions, such research is routine. CERCLA will not absolve them from responsibility for their mistakes of judgment."[138] As might be expected, EPA has echoed this view. An EPA official has indicated that EPA will hold commercial lenders to a higher standard of knowledge than many other types of property owners.[139] Therefore, for a lender to retain the protection of the third-party defense, it must document first that before acquiring the property it performed a thorough search of the property title and of the land's physical condition and second that its investigation did not reveal any evidence of hazardous waste on the site.

Even if the lender/owner can establish that it was an innocent purchaser of the property under SARA, it will qualify for the third-party defense only if it can prove that it also was an innocent seller under new subsection 101(35)(C).[140] That sub-

[136] 42 U.S.C.A. § 9601(35)(B) (West Supp. 1987).
[137] 632 F. Supp. 573 (D. Md. 1986).
[138] *Id.* at 580 (footnotes omitted).
[139] Leifer Paper, *supra* note 26, at 5–6.
[140] 42 U.S.C.A. § 9601(35)(C) (West Supp. 1987). Of course, this statutory requirement does not apply if the lender has not sold the land. Regardless of whether the lender has sold the land, the lender will lose the third-party defense if it "caused or contributed to the release or threatened release of a hazardous substance which is the subject of the action. . . ." *Id.* § 9601(35)(D).

section provides the second way in which SARA's contractual relationship definitional section restricts availability of the third-party defense. Subsection 101(35)(C) provides that a former property owner is liable for response costs, even if it did not contribute to the contamination by action or inaction, if it obtained actual knowledge of the hazardous waste contamination during ownership but did not inform the property buyer of that fact. This new provision creates an incentive that did not previously exist for an innocent property purchaser who discovers hazardous waste contamination to clean the property, rather than to attempt to avoid CERCLA liability by selling the property before an EPA cleanup.

Under new subsection 101(35)(C), when an owner discovers the hazardous condition, it has three options. It can (1) clean the site, (2) sell the property and disclose the condition, or (3) sell the property without disclosing the condition. If it sells and discloses, subsection 101(35)(A)(i) would prevent the purchaser from using the third-party defense even if a cleanup does not occur during its ownership. Few, if any, purchasers would be willing to buy under these circumstances. On the other hand, if the owner sells without disclosing the existence of the hazardous waste, it is liable for response costs under new subsection 101(35)(C)[141] whenever the cleanup occurs. Therefore, aside from fraudulent concealment, the only way to make the property marketable again and to avoid CERCLA liability is to clean the site. EPA and the public obviously benefit from this result because the property will be cleaned sooner than if EPA performed the cleanup, and the cleanup will not require Superfund financing or a response cost recovery action. From the lender's perspective, this new subsection 101(35)(C) creates an additional incentive for scrupulous examination of a parcel of land before accepting it as collateral for a loan and before acquiring title to it.

The third-party defense provides a lender/owner's only realistic escape hatch from CERCLA liability under a proper interpretation of the Act. Before SARA, a lender/owner who did not operate the site and exercised due care with respect to existing hazardous wastes on the property would not be liable for response costs even if the cleanup occurred during its owner-

---

[141] 42 U.S.C.A. § 9601(35)(C) (West Supp. 1987).

ship.[142] SARA has virtually eliminated this safe harbor for lenders, however, and further creates a circumstance under which a lender will be liable for response costs although the cleanup occurred after it sold the contaminated property. Under the current state of the law, even the most innocent lender/owners will have an uphill battle qualifying for the third-party defense because of their recognized expertise in evaluating property.[143] As a last resort, therefore, a lender might argue that this liability scheme is unconstitutional.

D. *Constitutionality*

Defendants in CERCLA response cost recovery actions have challenged the Act on several constitutional grounds, including arguments based on an alleged violation of substantive due process,[144] the taking clause,[145] and the contract clause.[146] To date, no court has ruled on the constitutionality of CERCLA as applied to a lender. Accordingly, this section of the Article will focus on the constitutional defenses that a lender might assert in a CERCLA action. CERCLA affects a lender's rights and liabilities both when it owns only the mortgage and if and when it subsequently acquires the encumbered land. Therefore, analysis of CERCLA's constitutionality as it affects lenders must focus both on a lender's mortgage ownership and on its land

---

[142] *See* Cadillac Fairview/California Inc. v. Dow Chem. Co., 14 Envtl. L. Rep. (Envtl. L. Inst.) 20,376, 20,378 (C.D. Cal. 1984).

[143] One commentator has argued that Congress enacted the contractual relationship definition as part of SARA "to make clear that innocent purchasers of land may rely on the 'third-party' defense despite the fact that they engaged in a land transaction with the previous owner who caused a release." Garber, *Federal Common Law of Contribution Under the 1986 CERCLA Amendments*, 14 ECOLOGY L.Q. 365, 378 (1987). As discussed above, the new "contractual relationship" definition has a very different effect.

[144] *See, e.g.*, NEPACCO, 810 F.2d 726, 732–34 (8th Cir. 1986); United States v. Conservation Chem. Co., 619 F. Supp. 162, 218–21 (D. Mo. 1985).

[145] *See, e.g.*, NEPACCO, 810 F.2d at 734; *Conservation Chem.*, 619 F. Supp. at 215–17.

[146] *See, e.g.*, *Conservation Chem.*, 619 F. Supp. at 214; United States v. Stringfellow, 14 Envtl. L. Rep. (Envtl. L. Inst.) 20,388, 20,389 (C.D. Cal 1984). At least one CERCLA defendant also has challenged CERCLA as violating the ex post facto clause and the equal protection clause. *Conservation Chem. Co.*, 619 F. Supp. at 214. The court rejected both challenges. The ex post facto clause, applies only to criminal laws that inflict punishment. United States Trust Co. of New York v. New Jersey, 431 U.S. 1, 17 n.13 (1977). The equal protection challenge also cannot be sustained under the rational relationship test. *See generally* J. NOWAK, R. ROTUNDA & J. YOUNG, CONSTITUTIONAL LAW § 14.3, at 530 (3d ed. 1986).

ownership. Such analysis reveals that the Act is constitutional as applied to lenders in both situations.

1. Substantive Due Process Challenges

An innocent lender/owner plausibly can defend a CERCLA recovery action on two substantive due process grounds. First, the lender/owner may argue that imposition of CERCLA liability on a person who did not create or contribute to the problem for which damages are assessed violates substantive due process. Second, it may assert that CERCLA constitutes unconstitutionally retroactive legislation. Although retroactive legislation is not expressly prohibited by a constitutional provision, the Supreme Court has held that retroactive legislation may violate substantive due process.[147] In light of the low level of scrutiny afforded economic and social welfare legislation,[148] both arguments should be dismissed.

The first argument, that CERCLA's liability scheme is unconstitutional as applied to innocent owners, merits only passing reference. Few would deny that leaking hazardous waste sites pose a grave danger to people and to the environment and that sound policy demands that these sites be cleaned. The pertinent question becomes "who should bear the economic burden of the clean- up?"

Imposing response costs on the person who controls the property and is directly benefited by the cleanup is reasonable and is consistent with a landowner's common law liability.[149] Unless a property owner acts to prevent hazardous waste releases from its land, it is contributing to the releases, albeit by inaction. By failing to act, thereby necessitating EPA intervention, the lender is not an innocent owner.[150] Moreover, Congress exempted the

---

[147] *See, e.g.*, Railroad Retirement Bd. v. Alton R.R., 295 U.S. 330 (1935).

[148] *See generally* J. NOWAK, R. ROTUNDA & J. YOUNG, *supra* note 146, § 11.4, at 356.

[149] *See infra* text accompanying notes 190–238.

[150] This was the basis for the court's holding in United States v. Price, 523 F. Supp. 1055, 1073 (D.N.J. 1981), *aff'd*, 688 F.2d 204 (3d Cir. 1983). In that case, the court held that defendants, who acquired a parcel of property several years after dumping on the property had stopped, were liable under the Resource Conservation and Recovery Act, 42 U.S.C. § 6973 (1982), even though they had not dumped or permitted others to dump wastes on the property. The reported opinion does not indicate that the court or the parties considered the constitutionality of this imposition of liability. *Compare* Department of Envtl. Protection v. Exxon Corp., 151 N.J. Super. 464, 376 A.2d 1339 (N.J. Super. Ct. Ch. Div. 1977), in which the New Jersey Superior Court construed the following definition of "discharge" in a state statute to require an act by the landowner:

least blameworthy landowners by means of the third-party defense. Thus, the courts should reject the argument that imposition of liability on a lender/owner is a prima facie violation of substantive due process.

The lender/owner's second due process argument, that CERCLA is unconstitutional retroactive legislation, is also flawed. The Act does operate retroactively in the sense that it may impose liability for actions taken before the Act's effective date. CERCLA defendants may claim that they are being held liable for conduct that they did not know would be actionable or harmful at the time it was taken. If the waste generators and transporters and the waste site owners acting before CERCLA's effective date had known that they would be liable in the future for response costs, they might have conducted their waste activities in a manner designed to reduce or to avoid liability.[151] Similarly, if a lender had known that it could be liable for response costs if it acquired the waste site upon default in a loan, it might have refused to lend against the security of that land and might not have acquired the land to recover its loan.

One response to these arguments that avoids the constitutional issue is that CERCLA does not impose liability for acts committed before its effective date. Rather, CERCLA imposes liability for the land's current condition and the damage it is causing or threatening to cause. Supreme Court precedent supports this view: "A statute is not rendered retroactive merely because the facts or requisites upon which its subsequent action depends, or some of them, are drawn from a time antecedent to the enactment."[152] Although some courts have interpreted CERCLA and similar environmental acts in this manner,[153] this

---

"'Discharge' shall mean, but is not limited to, any spilling, leaking, pumping, pouring, emitting, emptying or dumping." 151 N.J. Super. at 471, 376 A.2d at 1343. The court's decision was based on rules of statutory construction, however, and not on a determination that the act would be unconstitutional if applied to an innocent landowner. 151 N.J. Super. at 471, 376 A.2d at 1343–44.

[151] *See, e.g.*, United States v. Shell Oil Co., 605 F. Supp. 1064, 1072 (D. Colo. 1985); Reply Memorandum of "Generator" Defendants and Amicus Curiae in Support of Motion for Summary Judgment on Constitutional Grounds at 21, United States v. South Carolina Recycling & Disposal, Inc., 653 F. Supp. 984 (D.S.C. 1984).

[152] Reynolds v. United States, 292 U.S. 443, 449 (1934).

[153] *See, e.g.*, United States v. Miami Drum Services Inc., 25 Env't Rep. Cas. (BNA) 1469, 1477 (S.D. Fla. 1986); South Carolina Recycling & Disposal, Inc., 653 F. Supp. at 984. *See also Price*, 523 F. Supp. at 1071–72, in which the court discussed a retroactivity challenge to the Resource Conservation and Recovery Act, 42 U.S.C. § 6973 (1982):

> Defendants . . . observe that Section 7003 was not adopted and did not become effective until 1976, and they argue that the statute was not meant to apply

interpretation does not respond to the substance of the retroactivity argument—that the defendant's actions were legal when taken and that the defendant may have been unaware at the time of the harmful effects that could result from its conduct. Despite the argument's superficial appeal, it is inconsistent with the Supreme Court's interpretation of the scope of protection the due process clause provides against retroactive legislation.

Two relatively recent Supreme Court cases, *Usery v. Turner Elkhorn Mining Company*[154] and *Pension Benefit Guaranty Corporation v. R.A. Gray & Company*,[155] demonstrate that even if CERCLA is construed as applying retroactively, it is not unconstitutional for that reason. In *Usery*, the Supreme Court considered a substantive due process challenge that parallels the challenge to CERCLA's retroactive aspects. The Federal Coal Mine Health and Safety Act of 1969,[156] as amended by the Black Lung Benefits Act of 1972,[157] requires coal mine operators to compensate, among others, any former employee disabled by pneumoconiosis, even if the employee had terminated his or her work in the industry before Congress enacted the 1969 Act. Like defendants challenging CERCLA's retroactive liability

---

> retroactively to acts that preceded that date. Hence, because the dumping of toxic wastes at Price's Landfill ceased in 1972, defendants argue that the statute cannot be used to impose liability on them.
>
> We find this argument unpersuasive. . . . The gravamen of a Section 7003 action, as we have construed it, is not defendants' dumping practices, which admittedly ceased with respect to toxic wastes in 1972, but the present imminent hazard posed by the continuing disposal (i.e., leaking) of contaminants into the groundwater. Thus, the statute neither punishes past wrongdoing nor imposes liability for past acts. Rather, as defendants themselves argue, its orientation is essentially prospective. When construed in this manner, the statute simply is not retroactive. It merely relates to current and future conditions.

Tri-County Landfill Co. v. Illinois Pollution Control Bd., 41 Ill. App. 3d 249, 353 N.E.2d 316 (1976) (application of Environmental Protection Act not retroactive because discharges from the landfill occur presently); Freeman Coal Mining Corp. v. Illinois Pollution Control Bd., 21 Ill. App. 3d 157, 313 N.E.2d 616 (1974) (pollution occurs at time of discharge, not at time pollution source is created); United States v. Diamond Shamrock Corp., 12 Envtl. L. Rep. (Envtl. L. Inst.) 20,819 (N.D. Ohio 1981) (though defendant ceased disposal of hazardous waste four years prior to passage of Resource Conservation and Recovery Act, imposition of liability is not retroactive because the Act applies to presently existing conditions); Philadelphia Chewing Gum Corp. v. Dep't of Envtl. Resources, 35 Pa. Commw. 443, 387 A.2d 142 (1978) (although appellants created condition before passage of Clean Streams Law, they could be required to clean present condition); Commonwealth v. Barnes & Tucker Co., 455 Pa. 392, 319 A.2d 871 (1974), *appeal dismissed*, Barnes & Tucker Co. v. Pennsylvania, 434 U.S. 807 (1977) (imposing liability based on Clean Streams Law does not deny due process of law).

[154] 428 U.S. 1 (1976).

[155] 467 U.S. 717 (1984).

[156] 30 U.S.C. § 801 (1982).

[157] 30 U.S.C. §§ 901–962 (1982).

scheme, the coal mine operators in *Usery* argued that holding them liable for the disabilities of former employees violated the operators' due process rights because liability was predicated on completed acts that, when taken, were legal and, "at least in part," were not known to be dangerous.[158]

The *Usery* Court unequivocally rejected this argument. Although the Court stated that the justifications supporting prospective application of legislation might be insufficient to justify its retroactive application, the Court applied a lower level of scrutiny in reviewing the retroactive aspects of the legislation. The facts that the defendants were unaware of the dangerous nature of their activities and that they acted in reliance on the current state of the law were factored into the Court's analysis but were not deemed controlling in assessing the rationality of the legislative scheme. The *Usery* Court held that "the imposition of liability for the effects of disabilities bred in the past is justified as a rational measure to spread the costs of the employees' disabilities to those who have profited from the fruits of their labor. . . ."[159] CERCLA's retroactive impact can be justified on similar grounds.

In 1984, the Court reaffirmed the *Usery* opinion in the *R.A. Gray* case. In that case, the Court considered a substantive due process challenge to the retroactive application of the Multiemployer Pension Plan Amendments Act of 1980 (Pension Plan Amendment)[160] to employers that had withdrawn from pension plans before the Pension Plan Amendment's enactment. In rejecting the employers' due process challenge, the Court relied heavily on *Usery*, stating that the Court in *Usery* "had little difficulty" rejecting the due process challenge.[161] Although the Court repeated the statement from *Usery* that a legitimate legislative purpose must support retroactive application of legislation, the Court applied a lower level of scrutiny in reviewing the sufficiency of that purpose and held: "[W]e believe it was eminently rational for Congress to conclude that the purposes of the [Pension Plan Amendment] could be more fully effectuated if its withdrawal liability provisions were applied retro-

---

[158] *Usery*, 428 U.S. at 15.
[159] *Id.* at 18.
[160] 29 U.S.C. § 1461 (1982).
[161] Pension Benefit Guaranty Corporation v. R.A. Gray & Co., 467 U.S. 717, 729 (1984).

actively."[162] Thus, because the legislative purpose was enhanced by its retroactive application, the lower level of review was satisfied.

In light of the holdings in *Usery* and in *R.A. Gray*, the retroactive aspect of CERCLA does not violate substantive due process. Congress enacted CERCLA in response to the shortcomings of the Resource Conservation and Recovery Act.[163] The imposition of response costs on those who benefited by the dumping activities or by the cleanup is rational. Furthermore, the publicized dangers of hazardous waste dumping and the federal and state environmental legislation in force before CERCLA's enactment put lenders and other landowners on notice of the potential for liability based on ownership of contaminated land. By lending against the security of a waste site, the lender assumed the risk that the land would be rendered valueless by the release of hazardous pollutants. Protecting a defendant lender from liability, though it was aware of a potential danger when it made the loan, removes an important incentive for socially responsible behavior.[164] For these reasons, courts that have characterized CERCLA as retroactive legislation properly have held that it does not violate substantive due process.[165]

---

[162] *Id.* at 730.

[163] 42 U.S.C. § 6973 (1982). *See* Grad, *supra* note 15, at 35–36.

[164] *Toxic Waste Note*, *supra* note 4, at 1559.

[165] *See, e.g.*, NEPACCO, 579 F. Supp. 823 (W.D. Mo. 1984), *aff'd in part*, 810 F.2d 726 (8th Cir. 1986), *cert. denied*, 108 S. Ct. 146 (1987); United States v. Shell Oil Co., 605 F. Supp. 1064 (D. Colo. 1985); United States v. Conservation Chem. Co., 619 F. Supp. 162 (W.D. Mo. 1985).

The defendants in both the *Shell Oil* and *Northeastern* cases alternatively argued that the government could not constitutionally charge them for response costs incurred before CERCLA's effective date. Although the district courts in both cases indicated that Congress constitutionally could have legislated such liability, the district courts reached opposite results on the issue of pre-CERCLA response cost liability. Based on its review of CERCLA's language and legislative history, the court in *Shell Oil* held that Congress intended to hold responsible parties liable for pre-enactment government response costs. The court summarily dismissed the defendant's argument that to do so would be unconstitutional:

> Once it is accepted that Shell may be liable for its pre-CERCLA acts, it is irrelevant, from a due process perspective, whether the government commenced cleanup before or after the Act became law on December 11, 1980. There are no serious due process concerns in holding responsible parties liable for pre-CERCLA response costs.

605 F. Supp. at 1073.

The district court in *Northeastern*, on the other hand, rejected such liability, finding that CERCLA does not expressly authorize recovery of such costs. The court indicated, however, that Congress could have done so.

> Although it was possible for Congress to legislate the liability of past generators and transporters for pre-CERCLA response costs, they did not and this Court does not deem it advisable to engage in judicial legislation concerning

2. Taking Clause Challenges

In addition to a substantive due process claim, CERCLA's retroactive aspect could trigger a taking claim by a lender that acquired a mortgage before CERCLA's effective date. Although the Supreme Court stated in *Penn Central Transportation Co. v. New York City*[166] that taking cases involve "ad hoc, factual inquiries,"[167] an admission it recently repeated in *Keystone Bituminous Coal Association v. De Benedictis*,[168] a lender has strong authority for its assertion that federal law can cause an unconstitutional taking of lien rights. In *United States v. Security Industrial Bank*,[169] *Armstrong v. United States*,[170] and *Louisville Bank v. Radford*,[171] the Supreme Court reviewed taking challenges to laws that destroyed or substantially destroyed pre-existing liens. In *Armstrong* and in *Radford*, the Court held that an unconstitutional taking had occurred. In *Security Industrial Bank*, the Court sidestepped the issue by construing the statute to apply only prospectively because "substantial doubt" existed that retroactive application would survive the taking challenge.

Armed with this authority, a lender holding a pre-CERCLA lien on a waste site could argue that retroactive application of CERCLA is a taking of its mortgage lien. The most important right granted in a mortgage is the right to sell the property upon the borrower's default to recoup the loan amount.[172] If the mort-

---

a statute of such importance and controversy. All doubts of retroactive application must be resolved in favor of the defendants; therefore, the defendants are not liable for pre-CERCLA response costs.

579 F. Supp. at 843. On appeal, the Eighth Circuit reversed this portion of the district court decision on the basis of the district court's reasoning in *Shell Oil*. *See also* United States V. Wade, 577 F. Supp. 1326 (E.D. Pa. 1984) (holding that federal courts may create federal common law when necessary to protect uniquely federal interests).

[166] 438 U.S. 104 (1978).

[167] *Id.* at 124. The Court went on to say:

> [T]his Court, quite simply, has been unable to develop any "set formula" for determining when "justice and fairness" require that economic injuries caused by public action be compensated by the government, rather than remain disproportionately concentrated on a few persons. . . . Indeed, we have frequently observed that whether a particular restriction will be rendered invalid by the government's failure to pay for any losses proximately caused by it depends largely "upon the particular circumstances [in that] case."

Id.

[168] 107 S. Ct. 1232, 1247 (1987).

[169] 459 U.S. 70 (1982).

[170] 364 U.S. 40 (1960).

[171] 295 U.S. 555 (1935).

[172] If the lender cannot foreclose the mortgage, it is virtually valueless. The other rights the mortgage grants the lender, such as the right to ensure that property taxes and insurance premiums remain current, are of no consequence to the lender unless it

gaged land contains hazardous wastes, the lender will argue that the specter of section 107 liability effectively destroys its right to foreclose on the property by requiring the foreclosure sale purchaser to acquire CERCLA liability along with title. The purchaser potentially will be liable for response costs far in excess of the property's uncontaminated value unless the purchaser qualifies for the limited third-party defense. Because the lender is normally the only bidder at the foreclosure sale and because, as described above,[173] a lender will have difficulty qualifying for the third-party defense, a lender may argue that CERCLA effectively destroys its security.

A fundamental flaw in the lender's taking argument is that it focuses on the decrease in the value of its collateral, rather than on an allegation that CERCLA deprives it of a legally protected property interest. Although CERCLA may affect the value of the collateral, it does not affect the mortgage lien. This is not a mere technicality. The lender retains its lien until the debt is paid or the lender forecloses. It also retains the benefits of the mortgage covenants and the common law and statutory rights afforded mortgagees. For example, if the state where the land is located is a title theory or intermediate theory jurisdiction, in which a mortgage conveys title to the encumbered land to the lender, the lender can take possession of the property even before foreclosing. Additionally, the lender retains whatever rights are created by state law for the appointment of a receiver to take possession of the property. All the legal rights associated with the mortgage lien remain valid.

The continued vitality of the mortgage lien distinguishes a taking challenge to CERCLA from the successful taking challenges involved in the three major Supreme Court lien cases. In *Security Industrial Bank*, the Court reviewed a federal law that authorized complete destruction of pre-existing liens.[174] In *Arm-*

---

can foreclose. The mortgage covenants are almost all directed to maintaining the value of the collateral, which the lender is concerned with only as long as it retains the right to sell the property.

As further evidence of the importance of the right to foreclose, once a mortgage has been foreclosed, it is characterized as being *functus officio*. *See, e.g.*, Cross Cos., Inc. v. Citizens Mortgage Inv. Trust, 305 Minn. 111, 232 N.W.2d 114, 117–18 (1975). Thus, even if the foreclosure sale proceeds were insufficient to pay the loan in full, so that a portion of the debt remains outstanding, the mortgage is extinguished because it has served its primary function.

[173] *See supra* text accompanying notes 115–43.

[174] 459 U.S. 70 (1982). This case consolidated seven bankruptcy cases, in which each bankrupt debtor used an exemption created by the Bankruptcy Reform Act of 1978 to

strong, although the plaintiffs' pre-existing liens were not technically destroyed, they were rendered totally unenforceable as a matter of law.[175] *Radford* involved a federal law that stripped

---

void a lien that had been created before that Act's effective date. 11 U.S.C. § 522(f)(2) (1982). The lienors challenged this retroactive application of § 522(f)(2) as a violation of the taking clause. The Court did not decide the constitutional issue. Instead, it relied on the rule of statutory construction that, if "fairly possible," the court should construe a statute in a manner that avoids a constitutional issue. Although the statutory language gave no indication that Congress intended the Bankruptcy Reform Act to apply only prospectively, the Court so construed it because "there is substantial doubt whether the retroactive destruction of the appellees' liens in this case comports with the Fifth Amendment." 459 U.S. at 78.

The Court distinguished a creditor's contractual right to repayment of the debt from the property right it acquires in the collateral by virtue of the lien. Although Congress is constitutionally empowered to impair contract rights retroactively in connection with bankruptcy proceedings, the Fifth Amendment is a check on Congress' power in dealing with property rights. U.S. CONST. amend. V. "[H]owever 'rational' the exercise of the bankruptcy power may be, that inquiry is quite separate from the question whether the enactment takes property within the prohibition of the Fifth Amendment." 459 U.S. at 75. Because § 522(f)(2) provided for destruction of property rights that existed when the statute became effective, the Court determined that a substantial taking issue existed.

In an interesting concurrence, Justice Blackmun, joined by Justices Marshall and Brennan, argued that the Court should reach the taking issue. Despite the Bankruptcy Reform Act's clear operation to destroy the pre-existing liens involved in that case, Justice Blackmun stated that he would hold that the Bankruptcy Reform Act did not cause a taking:

> because the exemptions in question are limited as to kinds of property and as to values; because the amount loaned has little or no relationship to the value of the property; because these asserted lien interests come close to being contracts of adhesion; because repossessions by small loan companies in this kind of situation are rare; because the purpose of the statute is salutary and is to give the debtor a fresh start with a minimum for necessities; because there has been creditor abuse; because Congress merely has adjusted priorities, and has not taken for the Government's use or for public use; because the exemption provisions in question affect the remedy and not the debt; because the security interest seems to have little direct value and weight in its own right and appears useful mainly as a convenient tool with which to threaten the debtor to reaffirm the underlying obligation; because the statute is essentially economic regulation and insubstantial at that; and because there is an element of precedent favorable to the debtor in such cases as Penn Central Transp. Co. v. New York City and PruneYard Shopping Center v. Robins.

Id. at 84 (Blackmun, J., concurring) (citations omitted).

Blackmun's most persuasive argument is that a taking did not occur because the lien was an unenforceable part of an unconscionable transaction. The other factors stated in the concurrence, however, are contrary to established taking jurisprudence. For example, that the statute is "essentially economic legislation" is irrelevant in determining whether a compensable taking has occurred. *Id.* Similarly, the absence of governmental or public use is irrelevant as long as a public purpose exists. *See* Hawaii Hous. Auth. v. Midkiff, 467 U.S. 229 (1984); Berman v. Parker, 348 U.S. 26 (1954).

[175] 364 U.S. 40, 48 (1960). In *Armstrong*, the plaintiffs acquired mechanics' liens in property subsequently acquired by the federal government. Although the liens were still valid after the government's acquisition, they were unenforceable because of sovereign immunity. The Court held that, although the liens were still valid: "The total destruction by the Government of all value of these liens, which constitute compensable property, has every possible element of a Fifth Amendment 'taking'. . . ."

a mortgagee of its most important lien rights.[176] In contrast, CERCLA does not alter a mortgagee's lien rights. Because no property right is affected, the taking clause does not apply.

Even on its own terms, the lender's argument is fundamentally flawed. The substance of the lender's complaint is that CERCLA decreases the value of its collateral. The simple answer is that CERCLA did not destroy the property's value—the dumping activities did. Regardless of the existence of CERCLA, a leaking hazardous waste site is a liability rather than an asset. As described in Part III of this Article, a landowner has an independent common law duty to abate hazardous conditions on its property and is responsible for damages resulting from those conditions. The government is not in violation of the taking clause when it acts to abate a public nuisance.[177] Based on the above analysis, CERCLA does not effect a taking whether it is applied prospectively or retroactively.

---

[176] 295 U.S. 555, 601–02 (1935). In *Radford*, the Court held unconstitutional a Bankruptcy Act amendment that applied only retroactively. The amendment, enacted during the 1930s depression, permitted a bankrupt debtor under certain circumstances to eliminate a lien on its land for less than the outstanding debt amount. If the lender consented, the Bankruptcy Act permitted a borrower, upon being adjudged bankrupt, to purchase the encumbered property at its then appraised value and provided for deferred purchase payments bearing interest at the rate of one per cent per annum. If the lender did not consent to the purchase, the borrower could obtain a five year stay on all proceedings against the land and retain possession of the property in exchange for payment of annual rent. At any time during the five year period, the borrower could pay the lender the appraised value for the property and thereby eliminate the lien.

The Court held that the amendment was an unconstitutional taking because it deprived the lender of the following property rights:

1. The right to retain the lien until the indebtedness thereby secured is paid.
2. The right to realize upon the security by a judicial public sale.
3. The right to determine when such sale shall be held, subject only to the discretion of the court.
4. The right to protect its interest in the property by bidding at such sale whenever held, and thus to assure having the mortgaged property devoted primarily to the satisfaction of the debt, either through receipt of the proceeds of a fair competitive sale or by taking the property itself.
5. The right to control meanwhile the property during the period of default, subject only to the discretion of the court, and to have the rents and profits collected by a receiver for the satisfaction of the debt.

Id. at 594–95.

The Court found that the only right left to the lender under the mortgage was the right to retain the lien until the borrower exercised its right to release the lien by paying the property's appraised value, which was less than the outstanding debt amount. *Id.* at 596. The Court held that, because the amendment destroyed "rights in specific property which are of substantial value," it violated the taking clause. *Id.* at 601.

[177] *See, e.g.*, Miller v. Schoene, 276 U.S. 272 (1928). *See also* Keystone Bituminous Coal Ass'n v. De Benedictis, 107 S. Ct. 1232, 1246 n.22 (1987); R. EPSTEIN, TAKINGS: PRIVATE PROPERTY AND THE POWER OF EMINENT DOMAIN 112 (1985); Michelman, *Property, Utility, and Fairness: Comments on the Ethical Foundations of "Just Compensation" Law*, 80 HARV. L. REV. 1165, 1196–1201 (1967).

3. Contractual Impairment Challenges

The lender's final constitutional challenge to CERCLA's liability scheme asserts that the Act effects an unconstitutional impairment of contractual relations. While a challenge based on the contract clause can be readily dismissed because the contract clause does not apply to federal legislation,[178] the federal government's right to interfere with contractual rights is restricted by the Fifth Amendment due process clause.[179] This protection, however, is far less substantial than that which the contract clause provides against state legislation.[180] As stated by the Supreme Court in *National Railroad Passenger Corporation v. Atchison, Topeka and Santa Fe Railway Company*:[181]

> When the contract is a private one, and when the impairing statute is a federal one, . . . inquiry is especially limited, and the judicial scrutiny quite minimal. The party asserting a Fifth Amendment due process violation must overcome a presumption of constitutionality and "establish that the legislature has acted in an arbitrary and irrational way."

Because CERCLA's liability scheme is closely tailored to its legislative goals,[182] a lender will find it difficult, if not impossible, to establish that the scheme is either arbitrary or irrational.

Moreover, a contract impairment challenge is flawed in the same manner as a taking challenge. The basis for the lender's contract challenge is that the threat of CERCLA liability impairs its ability to enforce its loan contract. Such a challenge should fail, however, because CERCLA does not alter the terms of the loan agreement. The lender retains its right to pursue anyone personally liable for the debt. CERCLA merely allocates liability for the cost of eliminating a hazardous condition.

---

[178] The grounds for dismissal are contained in the clause's express terms: "No *State* shall . . . pass any . . . law impairing the obligation of contracts." U.S. CONST. art. I, § 10 (emphasis added). Because CERCLA is federal rather than state legislation, it is not subject to the prohibition of the contract clause. At least two federal district courts have rejected contract clause challenges to CERCLA by non-lenders on this ground. United States v. Conservation Chem. Co., 619 F. Supp. 162, 214 (D. Mo. 1985); United States v. South Carolina Recycling & Disposal, Inc., 14 Envtl. L. Rep. (Envtl. L. Inst.) 20,388 (C.D. Cal. 1984).

[179] U.S. CONST. amend. V.

[180] As the Supreme Court stated: "We have never held . . . that the principles embodied in the Fifth Amendment's Due Process Clause are coextensive with prohibitions existing against state impairments of pre-existing contracts." Pension Benefit Guar. Corp. v. R.A. Gray & Co., 467 U.S. 717, 733 (1984).

[181] 470 U.S. 451, 472 (1985).

[182] *See supra* text accompanying notes 5–8.

Lenders in six states, however, have a more substantial argument that CERCLA impairs their contractual rights to recover outstanding debts. California, Idaho, Montana, Nevada, New Jersey, and Utah have statutory one-action rules.[183] The one-action rule provides that, in the event of default, the lender must foreclose on the encumbered land as a prerequisite to a suit to recover the amount outstanding under the loan agreement.[184] The rule is designed to force lenders to attempt to recover the debt from the collateral securing it before pursuing the borrower's other assets.[185] Lenders in states with one-action rules might claim that CERCLA's imposition of liability on innocent property owners effectively prevents the lender from foreclosing because the foreclosing lender usually is the only bidder at the foreclosure sale. As the only bidder, the lender usually acquires title to the foreclosed property However, no lender will wish to acquire title if the property is a hazardous waste site potentially subject to a CERCLA cleanup action. Under these circumstances, the lender is effectively prevented from suing to recover the debt.

The above argument lacks force for three reasons. First, a lender is not legally required to bid for the property at the foreclosure sale, although it normally does so. As long as the lender conducts the foreclosure according to normal procedures, the one-action rule should be satisfied even if no one purchases the property because the lender attempted to satisfy the debt from the property.[186] The lender then can seek a judgment for the debt against anyone who is personally liable for it.

Second, a court may excuse the lender from conducting the sale before seeking a judgment on the note. Courts in five of the six states with a one-action rule have created an exception to the rule's application for property that has become value-

---

[183] CAL. CIV. PROC. CODE § 726 (West Supp. 1987) (although § 726, by its terms, applies only to mortgages, courts have construed it to apply to trust deeds as well, *see, e.g.*, Bank of California Nat'l Ass'n v. Leone, 37 Cal. App. 3d 444, 112 Cal. Rptr. 394 (1974)); IDAHO CODE § 6-101 (1979); MONT. CODE ANN. § 71-1-222 (1986); NEV. REV. STAT. § 40.430 (1985); N.J. STAT. ANN. § 2A:50-2 to -2.3 (West 1952 & Supp. 1987); UTAH CODE ANN. § 78-37-1 (1987). Although the New Jersey one-action rule applies only to certain types of residential properties, residences surprisingly have been built on former hazardous waste dump sites.

[184] G. NELSON & D. WHITMAN, REAL ESTATE FINANCE LAW § 8.2, at 598 (2d ed. 1985).

[185] *Id.*

[186] United States v. Caprice, 427 F. Supp. 1035, 1039 (D.N.J. 1976).

less.[187] This judicially created exception is designed to excuse the lender from an exercise in futility. Although no one yet has asserted this exception when pollution has destroyed the property value, the rationale for the rule will be served by extending it to this situation. Thus, a lender holding a security interest in a hazardous waste site should be able to avoid the foreclosure requirement if it can establish that the wastes have rendered the land valueless because of potential response costs.

Finally, if the property is generating income, a lender may be able to recover the debt without foreclosing and then suing those who are personally liable. Each state with a one-action rule provides lenders with the right to petition a court for appointment of a receiver.[188] A receiver can protect the land from

---

[187] The five states are:

1) California: The California one-action statute provides: "There can be but one form of action for the recovery of any debt . . . secured by mortgage. . . . In the action the court *may* . . . direct the sale of the property. . . ." CAL. CIV. PROC. CODE § 726(a) (West Supp. 1987) (emphasis added). Use of the word "may" indicates that the court has discretion not to require a foreclosure sale. *See also* Kaiser Indus. Corp. v. Taylor, 17 Cal. App. 3d 346, 353, 94 Cal. Rptr. 773, 776 (1971) (holding that where creditor has no security interest in debtor's real property or security becomes valueless, creditor is not required to foreclose by statute providing only one form of action, foreclosure, on debt secured by a mortgage); Comment, *Mortgages and Trust Deeds: Foreclosure Sale of a Portion of the Mortgaged Premises: Remedies Open to the Mortgagee When the Security is Valueless: Pleading the Existence of Security*, 25 CALIF. L. REV. 469, 473–77 (1937). *But see* Jeanese, Inc. v. Surety Title & Guar. Co., 176 Cal. App. 2d 449, 455, 1 Cal. Rptr. 752, 755 (1960) (holding that a purchase money lender always must foreclose even if the security has become valueless).

2) Idaho: Edminster v. Van Eaton, 57 Idaho 115, 117–19, 63 P.2d 154, 155 (1936) (holding that statutory one-action rule does not preclude mortgagee from suing in independent action on note for which security was given, where security has become valueless); Warner v. Bookstahler, 48 Idaho 419, 282 P. 862 (1929) (holding that foreclosure of first mortgage without redemption and issuance of sheriff's deed exhausted security as to junior mortgagee); Berry v. Scott, 43 Idaho 789, 255 P. 305 (1927) (holding that one action rule provides only remedy for recovery of debt and enforcement rights secured by mortgage, unless it can be shown that the security is valueless).

3) Montana: Bailey v. Hansen, 105 Mont. 552, 74 P.2d 438 (1937) (holding that one action rule does not prohibit a personal action where the security given the creditor has become valueless without any fault on his part, in which event he may secure an attachment).

4) Nevada: McMillan v. United Mortgage Co., 82 Nev. 117, 122, 412 P.2d 604, 606 (1966) (holding that attachment statute may be utilized if security, without fault of the mortgagee or beneficiary, has become valueless).

5) Utah: Lockhart Co. v. Equitable Realty, Inc., 657 P.2d 1333 (Utah 1983) (recognizing an exception to the one-action rule where the security interest in the property has been depleted and is valueless through no fault of the mortgagee).

[188] California: CAL. CIV. PROC. CODE § 564 (West Supp. 1987); Cormack & Irsfeld, *Applications of the Distinction Between Mortgages and Trust Deeds in California*, 26 CALIF. L. REV. 206, 211–18 (1938); Comment, *Comparison of California Mortgages and Trust Deeds and Land Sale Contracts*, 7 UCLA L. REV. 83, 91–94 (1960).

Idaho: IDAHO CODE § 8-601 (1979).

Montana: MONT. CODE ANN. § 27-20-102 (1986).

Nevada: NEV. REV. STAT. ANN. § 32.010 (Michie 1986); Bowler v. Leonard, 70 Nev.

further waste and can apply the income generated by the property to payment of property expenses, including the outstanding loan amount.[189] In this way, the lender may avoid the problem presented by the one-action rule and still recover the outstanding loan amount. For these reasons, the one-action rule will not provide lenders with a successful contract impairment defense.

Therefore, like the substantive due process and taking clause arguments, the lender's impairment of contract theory is unavailing. CERCLA does not eliminate or modify any of the lender's rights pursuant to the debt contract. Although CERCLA renders foreclosure and purchase a less desirable means of recouping the debt, the substantial public good achieved by CERCLA's statutory scheme militates against a finding that this scheme is arbitrary or capricious. Realistically, therefore, the only potentially successful defense that a lender can assert to avoid CERCLA liability as an owner is the narrowly defined third-party defense.

III. COMMON LAW LIABILITY

Lenders have reacted so strongly to potential CERCLA liability in part because courts rarely have held them liable for their borrowers' activities. Injured parties have attempted to recover damages from lenders for their borrowers' actions on a number of theories, including that the borrower had acted as the lender's instrumentality[190] or

---

370, 382–84, 269 P.2d 833, 839 (1954); Electrical Prod. Corp. v. Second Judicial Dist. Ct., 55 Nev. 8, 11–13, 23 P.2d 501, 503 (1933).

New Jersey: Sweeney v. Grant Silk Mfg. Co., 119 N.J. Eq. 321, 182 A. 484, *aff'd*, 120 N.J. Eq. 607, 187 A. 374 (1936).

Utah: UTAH R. CIV. P. 66 (1987).

[189] While beyond the scope of this Article, a judicially established receivership creates an interesting issue concerning the receiver's liability for response costs as an "operator" pursuant to 42 U.S.C. § 9607(a)(1) (1982).

[190] United States v. Jon-T Chem., Inc., 768 F.2d 686, 691–92 (5th Cir. 1985) (listing the factors courts consider when determining whether one entity is the instrumentality or alter ego of another); Krivo Indus. Supply Co. v. National Distillers & Chem. Corp., 483 F.2d 1098, 1105 (5th Cir. 1973) ("[I]n the cases resulting in instrumentality liability for the creditor, the facts have unmistakably shown that the subservient corporation was being used to further the purposes of the dominant corporation and that the subservient corporation in reality had no separate, independent existence of its own."); Chicago Mill & Lumber Co. v. Boatmen's Bank, 234 F. 41 (8th Cir. 1916) (holding that where one corporation owns or controls the entire property of another, and operates its plant and conducts its business as a department of its own business, it is responsible for the obligations of the controlled corporation); James E. McFadden, Inc. v. Baltimore Contractors, Inc., 609 F. Supp. 1102 (D. Pa. 1985) (holding that total and actual control

agent,[191] or that the borrower and lender had been joint venturers.[192] The case law that has developed with respect to these theories requires a plaintiff to establish such total control of the borrower or active lender involvement that the wrongful conduct, though nominally committed by the borrower, in fact was committed by the lender.[193] In the usual loan transaction, a lender's activities will be insufficient to trigger this common law liability, even though the lender's activities may have been sufficient to trigger CERCLA liability as an operator.

A vital distinction exists, however, between a lender's liability during the life of the loan and a lender/owner's CERCLA liability. The lender/owner's CERCLA liability is not based on the loan relationship. Liability is triggered by the lender's property ownership.[194] Viewed in this light, the lender/owner's CERCLA liability is consistent with well established common law princi-

---

must be established to render a general contractor the instrumentality of the bonding corporation). *See also* Duff v. Southern Ry. Co., 496 So. 2d 760, 762–63 (Ala. 1986) (holding that the question whether employer was an instrumentality of railroad was a question of fact precluding summary judgment); Miller v. Dixon Indus., 513 A.2d 597, 604–05 (R.I. 1986) (stating that the mere fact that there exists a parent-subsidiary relationship between two corporations is insufficient reason to impose liability on the parent either for the torts of the subsidiary or for a contract breached by the subsidiary); In re S.I. Acquisition, Inc., 58 Bankr. Ct. Dec. (CCR) 454 (Bankr. W.D. Tex. 1986) (holding that creditor's action against debtor's parent corporation and against principal of both corporations seeking to pierce corporate veil under alter ego theory was not property of the bankruptcy estate and was not a claim assertable by debtor or trustee). *But see* In re Mercer Trucking Co., 16 Bankr. 176 (Bankr. N.D. Tex. 1981) (holding trustees had standing to assert alter ego status of creditor vis-a-vis debtors).

[191] A. Gay Jenson Farms, Co. v. Cargill, Inc., 309 N.W.2d 285, 292 (Minn. 1981); RESTATEMENT (SECOND) OF AGENCY § 14-o comment a (1958). *See* Save Way Oil Co. v. Mehlman, 115 A.D.2d 721, 496 N.Y.S.2d 537 (1986); B.A. Buck v. Nash-Finch Co., 78 S.D. 334, 102 N.W.2d 84 (1960).

[192] Rehnberg v. Minnesota Homes, 52 N.W.2d 454, 457 (Minn. 1952) (listing the four elements that must be satisfied to find a joint venture). *See also* Atlanta Shipping Corp. v. Chemical Bank, 631 F. Supp. 335 (S.D.N.Y. 1986) (dismissing joint venturer claim on the basis that a transaction involving a loan of money and creating a debtor-creditor relationship will not of itself make lender and borrower joint venturers); Connor v. Great W. Sav. & Loan Ass'n, 73 Cal. Rptr. 369, 375, 447 P.2d 609, 615 (1968) (dismissing joint venturer claim because association between builder and contractor did not include agreement to share profits and losses either might realize or suffer); Meyers v. Postal Finance, 287 N.W.2d 614, 617–18 (Minn. 1979) (holding that assignee was not liable for club's alleged misrepresentations solely because of its status as assignee); Holts v. Tillman, 480 So. 2d 1134 (Miss. 1985) (holding that attorney's conduct did not establish intent to form a joint venture with property owners and that to establish such a joint venture would require a showing that attorney has a proprietary interest in lease agreement).

[193] For an exposition of these and other theories of lender liability, see S. Nickles, Lender Liability: Major Theories, Minnesota Continuing Legal Education 1986 Bankruptcy Institute (Sept. 18–19, 1986 Minneapolis, Minn.).

[194] 42 U.S.C. § 9607(b)(3) (1982).

ples regardless of the involuntary nature of the lender's acquisition of the land.

Three common law actions—strict liability, nuisance, and trespass—are the direct forebears of CERCLA liability and often provide additional sources of liability for a potential CERCLA defendant. In each of these common law causes of action, particularly strict liability and nuisance, a property owner is liable for injuries resulting from the use of its land. The owner's absence of moral culpability is virtually irrelevant. Instead, the theory for imposing such stringent liability is to prevent injuries and, failing that, to ensure compensation for victims.[195]

A. *Strict Liability*

The most direct common law forebear of the CERCLA liability scheme is strict liability. Although strict product liability has evolved primarily during the past quarter century, the concept of strict liability originated more than a century ago. The doctrine developed in large part from the 1868 English decision, *Rylands v. Fletcher*.[196] The facts in that case parallel the factual settings in which CERCLA liability accrues.

In *Rylands*, the owners of a mill retained a contractor to construct a reservoir on their land.[197] Unknown to the owners, the reservoir was located over abandoned mine shafts. When the reservoir was partially filled with water, the water broke through the shafts, traveled through connecting passages, and damaged the plaintiff's property.[198] Despite the defendants' absence of moral culpability, the Court held them liable, employing language that applies with equal force to a hazardous waste site:

> [I]f the defendants, not stopping at the natural use of their [land], had desired to use it for any purpose which I may term a non-natural use, for the purpose of introducing into

---

[195] Although a site owner liable under CERCLA also often will be liable under one of these common law actions, CERCLA potentially provides important litigation advantages. CERCLA authorizes action against a site owner whenever "there may be an imminent and substantial endangerment to the public health or welfare or the environment because of an actual or threatened release of a hazardous substance from a facility." 42 U.S.C. § 9606(a) (1982). In contrast, a plaintiff suing a site owner under one of the common law causes of action faces more difficult causation and procedural barriers. For a description of these problems, see Farber, *supra* note 22.

[196] L.R. 3 H.L. 330 (1868).

[197] Although the mill owners only leased the land on which their mill was located, the Court treated them as owning the land for purposes of the opinion. *Id.* at 338.

[198] Although the plaintiff's and defendants' properties were not adjoining, the court treated them as being so. *Id.* at 337.

> the [land] that which, in its natural condition, was not in or upon it—for the purpose of introducing water, either above or below ground, in quantities and in a manner not the result of any work or operation on or under the land, and if in consequence of their doing so, or in consequence of any imperfection in the mode of their doing so, the water came to escape and to pass off into the [land] of the plaintiff, then it appears to me that that which the defendants were doing they were doing at their own peril; and if in the course of their doing it the evil arose . . . of the escape of the water, and its passing away to the [land] of the plaintiff and injuring the plaintiff—then . . . the defendants would be liable.[199]

If the term "hazardous waste" is substituted for the word "water" in this quotation, this passage describes the basic CERCLA liability scheme and demonstrates that a site owner could be held strictly liable under the common law for injuries caused by hazardous wastes from its site.

The concept of strict liability for nonnatural uses of land quickly took root in America. In fact, the Massachusetts Supreme Judicial Court applied the doctrine to a case involving percolating water in the same year that *Rylands* was decided.[200] Three years later, the Minnesota Supreme Court applied the doctrine to another case involving underground water.[201] Although the course of the doctrine's development has not always been a straight one, the strict liability concept is now a well-established part of our liability jurisprudence.[202] The First and Second Restatements of Torts both provide that a landowner is strictly liable for harm caused by hazardous activities on its land.[203] The First Restatement defines this liability in terms of "ultrahazardous activity,"[204] whereas the Second Restatement defines it in terms of "abnormally dangerous activity."[205] Such

---

[199] *Id.* at 339.

[200] Ball v. Nye, 99 Mass. 582 (1868).

[201] Cahill v. Eastman, 18 Minn. 324 (1871).

[202] W. PROSSER & W. KEETON, TORTS § 78 (5th ed. 1984).

[203] RESTATEMENT OF TORTS § 519 (1938); RESTATEMENT (SECOND) OF TORTS § 519 (1979).

[204] Section 520 provides that:

> An activity is ultrahazardous if it
> (a) necessarily involves a risk of serious harm to the person, land or chattels of others which cannot be eliminated by the exercise of the utmost care, and
> (b) is not a matter of common usage.

RESTATEMENT OF TORTS § 520 (1938).

[205] Section 520 currently provides that:

> In determining whether an activity is abnormally dangerous, the following factors are to be considered:
> (a) existence of a high degree of risk of some harm to the person, land or chattels of others;

differences in terminology and in the precise parameters of liability under the First and Second Restatements are immaterial, however, when applied to a waste site that is releasing hazardous materials into waterways and neighboring lands. Dumping wastes that can seriously interfere with life and with the environment is an ultrahazardous and abnormally dangerous activity.[206]

The policy underpinnings of these common law concepts and of CERCLA are the same. Both attempt to ensure that situations that are dangerous to society are avoided or cured. The methods adopted to achieve this goal also are the same. By unavoidably tying liability to such an objective standard as property title, recourse and remedy are virtually guaranteed. Most important, the strict liability rule gives a property owner clear notice that it will be liable for injuries resulting from hazardous conditions on its land. The owner thus has a strong incentive to conduct its activities in a manner that will eliminate or reduce injuries. If injuries nevertheless result, the property owner is easily identifiable and owns at least one asset—the land—that can be applied to satisfy a judgment against it. In contrast, a person who does not own the land but merely conducts the operation, such as an independent contractor or lessee, may be difficult to locate or may be judgment proof. Thus, CERCLA may be viewed as

---

(b) likelihood that the harm that results from it will be great;
(c) inability to eliminate the risk by the exercise of reasonable care;
(d) extent to which the activity is not a matter of common usage;
(e) inappropriateness of the activity to the place where it is carried on; and
(f) extent to which its value to the community is outweighed by its dangerous attributes.

RESTATEMENT (SECOND) OF TORTS § 520 (1979).

Although the Second Restatement characterizes this form of liability as being strict, the defendant's blameworthiness is inherent in a determination that the activity was abnormally dangerous under the factors set forth in § 520.

> [The § 520 factors] consider the place where the activity is carried on, or consider alternatives to the activity. Such considerations permit notions of fault to get in by the back door, because the choice of place, the question of appropriateness of the activity to a particular place, or choices in the manner of conducting the activity invariably import notions of duty of care, responsibility and fault. This, in turn, places the heavy burden of proof back on the plaintiff.

REPORT OF THE SUPERFUND STUDY GROUP K-16, *quoted in* D. FARBER, TOXIC TORTS AND VICTIM'S COMPENSATION PROPOSALS 3 (1985).

[206] Department of Envtl. Protection v. Ventron Corp., 94 N.J. 473, 468 A.2d 150 (1983). The court interpreted the Second Restatement as incorporating the *Rylands* test. *Id.* at 157. *See also* THE LAW OF TORTS § 78 (W. Keeton 5th ed. 1984); Comment, *Successor Landowner Liability for Environmental Torts: Robbing Peter to Pay Paul?*, 13 RUTGERS L.J. 329, 330–31 n.6 (1982).

a codification of common law strict liability, although it provides a third-party defense for innocent owners that did not exist at common law.

B. *Nuisance*

CERCLA's other forebears is nuisance doctrine.[207] Like strict liability, nuisance imposes liability based on property ownership. Long before environmental protection statutes existed, nuisance doctrine prohibited a property owner from using its land in a way that unreasonably interferes with a right common to the general public (public nuisance)[208] or with another property owner's use and enjoyment of its land (private nuisance).[209] A leaking hazardous waste site constitutes both a public and a private nuisance under common law.

The Second Restatement's definition of a public nuisance states that an interference is unreasonable if: (1) it involves a significant interference with the public health, safety, peace, comfort, or convenience; (2) the conduct is proscribed by law; or (3) the conduct is continuing or has produced a long-lasting effect and, as the defendant knows or has reason to know, has a significant effect upon the public right.[210] Because CERCLA authorizes EPA to clean a site and to charge the owner for response costs only after the site has been identified as presenting an imminent and substantial danger to the public health,[211] any site subject to CERCLA cleanup also would constitute a public nuisance under the common law.[212] Thus, CERCLA liability is, in a sense, a codification of a specific instance of common law public nuisance and serves the same preventive and compensatory goals.

The Second Restatement of Torts defines a private nuisance as "a structure or other condition erected or created on [one's]

---

[207] One commentator has described nuisance as "the common law backbone of modern environmental law." THE SUPERFUND CONCEPT, *supra* note 1, at 23.

[208] RESTATEMENT (SECOND) OF TORTS § 821B (1979).

[209] *Id.* § 821D. Comment a to § 821D states that the private nuisance action dates to the twelfth century.

[210] *Id.* § 821B.

[211] 42 U.S.C. § 9604(1) (1982).

[212] *See* THE SUPERFUND CONCEPT, *supra* note 1, at 23–24; Comment, *supra* note 206, at 329–32; Comment, *An Analysis of Common Law and Statutory Remedies for Hazardous Waste Injuries*, 12 RUTGERS L.J. 117, 125–28 (1980) [hereinafter Comment, *Remedies for Hazardous Waste Injuries*].

land which causes a continuing invasion of [another's] land."[213] If a dump site has leaked hazardous wastes onto neighboring lands, the dump site owner may be held liable for damages caused on neighboring lands. Thus, neighboring landowners may sue the site owner on a private nuisance cause of action.

C. *Trespass*

Trespass is the third of CERCLA's common law antecedents and an additional cause of action upon which lender/owners may be sued. Though it is a more limited action than CERCLA, trespass liability more closely parallels the CERCLA liability of an innocent owner than does strict liability or nuisance. Like CERCLA, a trespass action focuses on land ownership. Liability in trespass, however, is less than strict. To establish a cause of action for trespass, a plaintiff must prove not only that pollutants physically invaded its land,[214] but also that the defendant knowingly or negligently caused the invasion.[215]

Some courts have held that constructive notice to the defendant may be sufficient to support liability in trespass.[216] However, at least one court has held that a defendant was not liable in trespass even though he intentionally discharged pollutants on his land, because he did not know that subterranean currents or other conditions would carry the pollutants to the plaintiff's land.[217] In light of the problems of proof and the availability of alternative theories of recovery, plaintiffs rarely rely on the trespass theory even if their land or water has been damaged by pollutants.[218] The trespass theory, however, like strict liability and nuisance, demonstrates that under traditional notions a property owner may be held liable for injuries caused by conditions on its land. CERCLA's liability scheme reflects this same notion.

---

[213] RESTATEMENT (SECOND) OF TORTS § 201 comment b (1979).

[214] W. PROSSER, THE LAW OF TORTS § 89, at 594 (4th ed. 1971).

[215] Hudson v. Peavey Oil Co., 279 Or. 3, 6–7, 566 P.2d 175, 177 (1977); *see* RESTATEMENT (SECOND) OF TORTS § 166 (1979).

[216] Elsey v. Adirondack & St. L. Ry. Co., 97 Misc. 273, 277, 161 N.Y.S. 391, 393 (Sup. Ct. 1916); Furrer v. Talent Irrigation Dist., 258 Or. 494, 513, 466 P.2d 605, 615 (1970).

[217] Phillips v. Sun Oil Co., 307 N.Y. 328, 331, 121 N.E.2d 249, 251 (1954).

[218] Comment, *Remedies for Hazardous Waste Injuries*, *supra* note 212, at 126.

D. *Negligence*

Unlike strict liability, nuisance, and trespass, negligence focuses less on property ownership in the hazardous waste context than on duty. A suit for negligence will lie against the owner of a hazardous waste site if the owner breached a duty of reasonable care owed to the plaintiff, thereby proximately causing a foreseeable injury to the plaintiff.[219] The focus on foreseeability distinguishes negligence from CERCLA's basic liability scheme, although foreseeability does affect the availability of the third-party defense. An innocent owner can qualify for that defense only if it exercised due care, which inevitably requires an analysis of the foreseeability of the injury.

The dangerous nature of hazardous waste sites and the gravity of the injuries that can result from their negligent operation or maintenance place a high standard of care on a site owner.[220] Thus, if hazardous materials are escaping from the site, particularly to the degree necessary to prompt an EPA cleanup, a plaintiff in a private action for negligence may be able to establish that the defendant did not satisfy its duty of reasonable care.[221] The unique characteristics of hazardous waste litigation, however, may enable the site owner to escape liability for negligence although it might be unable to do so under CERCLA. A defendant, particularly one that owned the site before the dangers of pollution were fully recognized and while disposal techniques were less effective, might avoid liability for negligence by establishing that it acted reasonably in light of the state of knowledge and technology at that time. Thus, the basic CERCLA liability scheme is more closely related to strict liability and to nuisance than to negligence.

E. *Successor Owner Liability*

As established by the above analysis, the common law imposes tort liability on a property owner that creates a hazardous waste site that causes injury. Whether a subsequent site owner also is liable, though no dumping occurred during its ownership,

---

[219] W. PROSSER & W. KEETON, *supra* note 202, § 30, at 164–65.

[220] *Id.* at 171; Comment, *Remedies for Hazardous Waste Injuries*, *supra* note 212, at 123.

[221] *See, e.g.*, Ewell v. Petro Processors of Louisiana, Inc., 364 So. 2d 604, 606 (La. Ct. App. 1978).

is more problematic. Like CERCLA, the common law does not exempt a property purchaser from liability solely because it did not create the hazard or because it owned the site for only a short time before the injury occurred. Under common law, the purchaser is liable if it knew or should have known of the wastes and if it had a reasonable opportunity to correct the condition.[222] In case after case, courts have held property purchasers liable for injuries resulting from artificial conditions that existed on the land when the purchasers acquired the property.[223]

The most frequently cited case on this issue is *Palmore v. Morris*.[224] In *Palmore*, a boy was injured when a negligently maintained gate fell on him. The accident occurred twenty hours after the property on which the gate was located had been sold. In holding that the purchaser, rather than the seller, was liable for the boy's injuries, the court stated:

> Before he purchased the real estate, the law presumes the grantee examined the property, and was cognizant of its situation, surroundings, the character of the structures upon it, and their condition of repair. Without an express covenant by the grantors, as between them and the grantee, there was no duty on the grantors to repair. The purchaser thereafter assumed that duty because he then became the owner and occupant. . . . There may be a case where the grantor conceals from the grantee a defect in a structure, known to him alone, and not discoverable by careful inspection, that the owner would be held liable, though out of possession; but that is not this case.[225]

The case presents a species of title liability. Under this analysis, whether the vendee created the condition or affirmatively contributed to it is irrelevant. Rather, the vendee's possession

---

[222] Section 336 of the Second Restatement of Torts provides:

> One who takes possession of land upon which there is an existing structure or other artificial condition unreasonably dangerous to persons or property outside of the land is subject to liability for physical harm caused to them by the condition after, but only after,
>
> (a) the possessor knows or should know of the condition, and
>
> (b) he knows or should know that it exists without the consent of those affected by it, and
>
> (c) he has failed, after a reasonable opportunity, to make it safe or otherwise to protect such persons against it.

RESTATEMENT (SECOND) OF TORTS § 366 (1979). *See also id.* § 839.

[223] *See, e.g.*, Central Consumers' Co. v. Pinkert, 122 Ky. 720, 92 S.W. 957 (1906); Palmore v. Morris, 182 Pa. 82, 37 A. 995 (1895); W. PROSSER, THE LAW OF TORTS (W. Keeton 5th ed. 1984); Annotation, *Liability of Purchaser of Premises for Nuisance Thereon Created by Predecessor*, 14 A.L.R. 1094 (1921).

[224] 37 A. 995 (1895).

[225] *Id.* at 999.

and control of the land is dispositive. Again, responsibility follows title without more:

> Any future possession [by the seller] in face of his deed, unless there be an independent stipulation to the contrary, would be a palpable trespass; and with his surrender of possession all the duties incident to ownership, as to him, were at an end. From the moment [the purchaser] took possession under his deed the duties theretofore incumbent on [the seller] were transferred to him, and he became answerable to the public for neglect in their performance.[226]

Because the vendee has exclusive control over the land through the acquisition of title, it is liable at common law for injuries resulting from harmful conditions on the land.

A few states have adopted statutes codifying this common law principle.[227] For example, section 7.48.170 of the Washington Revised Code provides: "Every successive owner of property who neglects to abate a continuing nuisance upon or in the use of such property caused by a former owner, is liable therefore in the same manner as the one who first created it."[228] Courts have not afforded lenders any special protection from the operation of these statutes. In *Pierce v. German Savings & Loan Society*,[229] for example, the California Supreme Court applied a similar statute to hold a mortgagee that acquired title to the encumbered land liable for damages resulting from a nuisance that existed on the land before the lender acquired it.[230] Thus, lenders that acquire title are no more shielded by their status as lenders than are any other property purchasers.

CERCLA's withholding of lender immunity parallels the common law. A lender that acquires title to a waste site normally will be liable under the common law rule of vendee liability in the same circumstances that it will be liable under CERCLA. As under CERCLA, an institutional lender's sophistication and expertise with respect to land will make it difficult for the lender to avoid common law liability by claiming absence of notice.[231] On the other hand, if the lender successfully establishes the lack of notice defense, it will be exempt from liability not only under

---

[226] *Id.* at 998–99.

[227] *See, e.g.*, CAL. CIV. CODE § 3483 (West 1970); IDAHO CODE § 52-109 (1979); WASH. REV. CODE ANN. § 7.48.170 (1961).

[228] WASH. REV. CODE. ANN. § 7.48.170 (1961).

[229] 72 Cal. 180, 13 P. 478 (1887).

[230] *Id.* at 479.

[231] *See supra* notes 123–39 and accompanying text.

the common law, but also under the third-party defense to CERCLA liability.

CERCLA and the common law conflict in an important manner, however. If a lender/owner acquires the contaminated property without actual or constructive notice of the waste, it can retain its immunity from CERCLA liability, even if, after discovering the waste, it exercises the due care required by the third-party defense provision.[232] Apparently, the statutory due care standard does not impose a duty on the owner to eliminate the hazard, because the owner will not be liable for cleanup costs if the government or another authorized party cleans the site. Although the third-party defense provision expressly exempts an innocent owner from liability only for response costs,[233] thereby creating an inference that the third-party defense will not protect the owner from a judicial order to clean the site issued pursuant to section 106,[234] such an interpretation is unlikely because it could render the third-party defense meaningless.[235] Under the common law, on the other hand, the lender/owner will be liable for injuries caused by the hazardous condition after it has had a reasonable opportunity to discover the condition and to correct it.[236] Additionally, the new owner potentially will be subject to a common law action to abate the public nuisance.[237]

A court could circumvent this conflict by holding that CERCLA preempts the common law with respect to hazardous waste sites. CERCLA, however, does not address personal injury actions. It deals only with the procedures for cleaning hazardous waste sites and the mechanisms for financing cleanups, whether the cleanup is conducted by the owner or by some other entity. Moreover, imposing liability for personal injury and private property damage creates an incentive for even an innocent owner to clean the site, though this factor brings the preemption argument back full circle. The abatement question also presents a difficult preemption issue because EPA has identified only a fraction of the country's waste sites for CERCLA cleanups.[238]

---

[232] W. PROSSER & W. KEETON, *supra* note 221, § 64, at 449–50.
[233] 42 U.S.C. § 9607(b)(3) (1982).
[234] 42 U.S.C. § 9606 (1982).
[235] *See* 1A F. GRAD, TREATISE ON ENVIRONMENTAL LAW § 4A.02[e] (1987).
[236] *See* RESTATEMENT (SECOND) OF TORTS § 366 (1979).
[237] *See id.* § 821C.
[238] *EPA Seeks Comments on 64 Proposed Sites to be Added to the National Priorities List*, *supra* note 4, at 1725.

The remaining waste sites could constitute public nuisances even though EPA has not targeted them for cleanup.

The salient point is that CERCLA imposes no greater burden on innocent landowners than the common law and, in some regards, may impose a lesser burden. Subsection 107(a) liability for response costs parallels the common law doctrines of strict liability and nuisance by tying liability to property ownership without regard to fault. The third-party defense of subsection 107(b)(3) potentially eliminates this liability for a lender/owner by providing a circumstance under which it will be held to only a standard of due care, rather than of strict liability. The exact parameters of the statutory due care standard for an innocent lender/owner, however, have yet to be defined.

IV. Scope of Lender/Owner's CERCLA Liability

Identifying lender/owners as potential CERCLA defendants is only the first step in assessing the Act's impact on them. In the usual case, the lender/owner is only one of many parties that is liable under CERCLA for response costs. Apportionment of liability, therefore, is a salient issue and, in view of the enormity of response costs, a very important one. Because lenders normally are perceived as having deeper pockets than other potential defendants, lenders will be particularly attractive targets to a cost- conscious EPA. Part IV of this Article examines the scope of a lender/owner's liability for response costs on two levels: (1) the amount of response costs that a court can assess against a lender/owner in a cost recovery action; and (2) the portion of those costs that the lender/owner can recover from others.

In lieu of an express provision concerning the scope of section 107 liability, CERCLA applies the same standard of liability imposed in another federal environmental act, the Clean Water Act.[239] Like CERCLA, however, the Clean Water Act leaves several basic liability issues open. Therefore, courts deciding CERCLA cost recovery actions must formulate the necessary common law. From the earliest CERCLA cases, courts have held that federal, rather than state, common law should con-

[239] Clean Water Act, 33 U.S.C. § 1321 (1982).

trol.[240] Therefore, courts deciding CERCLA cases have been free to establish the liability rules that will best serve CERCLA's purposes unconstrained by state law.

A. *Amount of Response Costs*

The courts uniformly have held that a CERCLA defendant may be held jointly and severally liable for response costs.[241] Therefore, a lender/owner sued for response costs may be liable for the entire amount of the costs. Although some courts have stated that the applicability of joint and several liability in a CERCLA cost recovery action should be determined on a case-by-case basis,[242] imposition of joint and several liability undoubtedly will be the norm when EPA institutes a cost recovery action because this standard provides important advantages to EPA.[243]

First, joint and several liability significantly simplifies EPA's pretrial investigations and its burden of establishing liability during the trial. Rather than finding and suing each potential

---

[240] *See, e.g.*, Colorado v. ASARCO, 608 F. Supp. 1484, 1489 (D. Colo. 1985); United States v. Chem-Dyne Corp., 572 F. Supp. 802, 808–10 (S.D. Ohio 1983). The courts have cited a variety of sources for this proposition. *See Toxic Waste Note*, *supra* note 4, at 1526. A waste site owner is entitled to contribution not only when it has paid EPA's response costs, but also when it has cleaned the site. *See, e.g.*, Sand Springs Home v. Interplastic Corp., 670 F. Supp. 913 (N.D. Okla. 1987).

[241] *See, e.g.*, United States v. Shell Oil Co., 605 F. Supp. 1064, 1083–84 & n.9 (D. Colo. 1985); United States v. A & F Materials Co., 578 F. Supp. 1249, 1256 (S.D. Ill. 1984); United States v. Wade, 577 F. Supp. 1326, 1338–39 (E.D. Pa. 1983); *Chem-Dyne*, 572 F. Supp. at 810.

[242] For economic analyses of joint and several liability, see Landes & Posner, *Joint and Multiple Tortfeasors: An Economic Analysis*, 9 J. LEGAL STUD. 517 (1980); Ulen, Hester & Johnson, *Minnesota's Environmental Response and Liability Act: An Economic Justification*, 15 Envtl. L. Rep. (Envtl. L. Inst.) 10,109, 10,113 (1985); Note, *Joint and Several Liability for Hazardous Waste Releases Under Superfund*, 68 VA. L. REV. 1157, 1182–95 (1982); Note, *Allocating Costs*, *supra* note 1, at 584.

[243] Although early versions of CERCLA expressly provided for joint and several liability, the provision was deleted from the final version to ensure the Act's passage. See 126 CONG. REC. 30,932 (1980) (statement of Sen. Randolph (D-W. Va.)); *Sand Springs Home*, 670 F. Supp. at 915. SARA's legislative history indicates, however, that Congress agreed with the courts' imposition of joint and several liability. *See, e.g.*, 131 CONG. REC. H11,073 (daily ed. Dec. 5, 1985) (statement of Rep. Eckart (D-Ohio)):

> The Committee on Energy and Commerce and the other committees involved in this bill fully subscribe to the reasoning of the court in the seminal case of *United States v. Chem-Dyne Corporation*, 572 F. Supp. 802 (S.D. Ohio 1983), which has established a uniform Federal rule allowing for joint and several liability in appropriate CERCLA cases. . . . Thus, nothing in this bill is intended to change the application of the uniform Federal rule of joint and several liability enunciated by the Chem-Dyne court.

See also 1A F. GRAD, *supra* note 235, § 4A-89 ("SARA, rather than directly addressing a central issue such as the strict, joint and several standard of liability in statutory language, makes the judicial opinions on this subject part of its legislative history.").

section 107 defendant, EPA can choose one or more defendants based on accessibility, ability to satisfy a judgment, and on the facility with which EPA can establish their liability.[244] Although a section 107 defendant may implead other potentially responsible parties, it will bear the burden and expense of locating and joining those parties and of proving their liability.[245]

Second, joint and several liability dramatically simplifies EPA's burden of proving causation. In the absence of joint and several liability, EPA could recover damages from a defendant only to the extent that EPA could prove the actual amount of damage that a particular defendant caused. This burden of proof usually would be impossible to satisfy because it would require EPA to establish the precise amount of each type of hazardous substance that has been released from the waste site as well as the identities of the generators and transporters of the hazardous substances, the site operators, and the site owners. The variety of activities in which potential section 107 defendants engage compounds the problem of establishing the amount of damage that each defendant caused.[246]

Finally, joint and several liability enhances EPA's ability to recover all of its response costs. In the absence of joint and several liability, EPA could recover the full amount of its response costs only if all responsible parties could be located and were sufficiently solvent to pay their portion of the response costs. For all the above reasons, joint and several liability is an appropriate common law liability doctrine for EPA response cost recovery actions.[247]

---

[244] *See* Dubuc & Evans, *Recent Developments Under CERCLA: Toward a More Equitable Distribution of Liability*, 17 Envtl. L. Rep. (Envtl. L. Inst.) 10,197, 10,197 (1987).

[245] The legislative history indicates that this was a reason for imposing joint and several liability:

> [T]he effect of the decision [imposing joint and several liability] was to require the defendants, rather than the plaintiff, to show that other tortfeasors contributed to the harm and in what quantities they so contributed. This incentive to locate all other responsible parties is one of the prime considerations underlying use of joint and several liability in pollution suits. In fact, in addition to shifting the burden of proving the cause of plaintiff's injury, the court placed on the sued defendants ultimate responsibility for bringing any other defendants into the suit as codefendants. Facing the prospect of either proving that other parties were also responsible for the injury or paying the full judgment themselves, defendants would henceforth have incentive to insure that no parties have been inadvertently omitted from the suit.

126 CONG. REC. 26,784 (1980) (statement of Rep. Gore (D-Tenn.)).

[246] Note, *Hazardous Waste and the Innocent Purchaser*, 38 FLA. L. REV. 253, 263 (1986).

[247] If the waste site owner or another entity specified in 42 U.S.C. § 9607(a) (1982)

B. *Lender/Owners' Right to Contribution*

The right to contribution may be considered the corollary of joint and several liability. If a court holds a lender/owner jointly and severally liable for response costs, CERCLA's goals normally will be furthered by permitting the lender/owner to recoup all or a portion of the damages from the other responsible parties specified in subsection 107(a).[248] Thus, in addition to any other remedies the lender/owner may have, such as insurance coverage or a cause of action against the former owner for misrepresentation or failure to disclose a material defect in the property,[249] the lender/owner can sue other responsible parties for contribution. The right to contribution is particularly important in light of the discretion joint and several liability affords EPA.[250]

Before SARA, no clearly recognized right to sue for contribution for response costs existed, particularly in light of the relatively recent recognition in America of a right to contribution among joint tortfeasors.[251] CERCLA did not expressly create a right of contribution, although subsection 107(e)(2) arguably contemplated its availability: "Nothing in this subchapter . . . shall bar a cause of action that an owner or operator or any other person subject to liability under this section . . . has or would have, by reason of subrogation or otherwise against any person."[252] Although this section could be construed as applying only to contractual rights, a few courts have permitted a con-

---

cleans the site and brings a cost recovery action, on the other hand, a court might determine that the equities of the case do not justify imposition of joint and several liability.

[248] 42 U.S.C. § 9607 (1982).

[249] Obde v. Schlemeyer, 56 Wash. 2d 449, 353 P.2d 672 (1960); W. PROSSER & W. KEETON, *supra* note 202, § 106; Keeton, *Fraud—Concealment and Non-disclosure*, 15 TEX. L. REV. 1, 18–21 (1936). If the lender acquired the property at a foreclosure sale, the former owner probably will not have made any such representations and thus will not have a duty to disclose defective conditions on the property.

[250] In Sand Springs Home v. Interplastic Corp., 670 F. Supp. 913, 914 (N.D. Okla. 1987), for example, EPA instituted a CERCLA administrative proceeding against a site owner and refused the owner's request that EPA join the waste generators because of "the emergency nature of the situation and the fact it would require a lengthy time period to identify and join the generators in such proceedings." Similarly, in Colorado v. ASARCO, Inc., 688 F. Supp. 1484, 1485 (D. Colo. 1985), EPA did not join hundreds of potentially liable parties. In United States v. Conservation Chem. Co., 589 F. Supp. 59 (W.D. Mo. 1984), EPA sued only seven potentially responsible parties. Three of the defendants each had contributed less than two percent of the waste. The defendants filed third-party claims against 154 generators, 16 insurance companies, and 14 federal agencies.

[251] Northwest Airlines, Inc. v. Transport Workers Union, 451 U.S. 77, 86–88 & nn.16–17 (1981); W. PROSSER & W. KEETON, *supra* note 202, § 50.

[252] 42 U.S.C. § 9607(e)(2) (1982).

tribution action on the basis of this language alone.[253] Other courts have authorized such suits as a matter of federal common law.[254] The latter have reasoned that contribution should be a normal concomitant of joint and several liability when the defendant is not guilty of willful wrongdoing.[255]

SARA has settled the debate on the right to contribution for response costs by expressly providing for a right of contribution against "any other person who is liable or potentially liable under section 107(a). . . ."[256] Moreover, the new subsection 113(f) expressly provides that CERCLA contribution actions are governed by federal law and must be brought in accordance with the Federal Rules of Civil Procedure. This clarification is especially helpful because not all states permit an action for contribution among parties that are jointly and severally liable.[257]

SARA's contribution provision provides relief to a CERCLA defendant in two ways. First, it reduces a defendant's litigation costs by eliminating the need to litigate the issue of whether a contribution action is available.[258] Second, the liberal joinder provisions of the Federal Rules of Civil Procedure, in conjunction with the contribution provision from SARA, provide CERCLA defendants with a relatively efficient and cost effective means of pursuing other potentially liable parties.[259] The joinder and contribution provisions minimize the time between the de-

---

[253] Wehner v. Syntex Agribusiness, Inc., 616 F. Supp. 27, 31 (E.D. Mo. 1985); United States v. South Carolina Recycling & Disposal, Inc., 653 F. Supp. 984, 995 n.8 (D.S.C. 1984); United States v. Chem-Dyne Corp., 572 F. Supp. 802, 807 n.3 (S.D. Ohio 1983).

[254] United States v. New Castle County, 642 F. Supp. 1258, 1265–69 (D. Del. 1986); *ASARCO*, Inc., 608 F. Supp. at 1484; United States v. Ward, 14 Envtl. L. Rep. (Envtl. L. Inst.) 20,804 (E.D.N.C. 1984).

[255] *See CERCLA 1985: A Litigation Update*, 15 Envtl. L. Rep. (Envtl. L. Inst.) 10,395, 10,403 (1985).

[256] 42 U.S.C. § 9613(f) (1982).

[257] Northwest Airlines, Inc. v. Transport Workers Union, 451 U.S. 77, 87 n.17 (1981).

[258] Although courts uniformly have granted a right to contribution for CERCLA response costs, "the issue was continually being contested by third-party defendants." Dubuc & Evans, *supra* note 244, at 10,200.

[259] Rule 14(a) of the Federal Rules of Civil Procedure provides that, at any time after an action is commenced, a defendant may implead "a person not a party to the action who is or may be liable to him for all or part of the plaintiff's claim against him." FED. R. CIV. P. 14(a). Thus, even though the right to contribution does not vest until a person actually pays more than her share of the damages, Rule 14(a) enables a third-party plaintiff to establish the liability of other responsible parties in the same suit in which EPA establishes the third-party plaintiff's liability. *See, e.g.*, Adalman v. Baker, Watts & Co., 599 F. Supp. 752 (D. Md. 1984), *aff'd in part, rev'd in part*, 807 F.2d 359 (4th Cir. 1986); Tri-Ex Enter., Inc. v. Morgan Guar. Trust Co. of New York, 586 F. Supp. 930 (S.D.N.Y. 1984); Vaughn v. Terminal Transp. Co., 162 F. Supp. 647, 648–49 (E.D. Tenn. 1957).

fendant's payment to EPA and its recovery of a portion of those costs from the other responsible parties. Additionally, the cost of a second trial is avoided, and the defendant avoids the possibility of inconsistent holdings on the evidence.[260]

In addition to codifying the right to contribution, new subsection 113(f) adopts an equitable method of apportioning damages among the responsible parties,[261] which should aid lender/owners. Contrary to the equitable apportionment method prescribed by subsection 113(f), the majority of courts apply a rule of pro rata contribution.[262] The pro rata method, in contrast, often would result in a lender/owner bearing a share of damages disproportionate to its fault. Unlike the other entities liable under subsection 107(a)—waste generators, waste site operators, and waste transporters[263]—the innocent lender/owner will not have actively contributed to the hazardous waste problem. Although the lender/owner does benefit from the cleanup, that benefit often will be of significantly less value than the lender's pro rata share of damages.

Because of the failings of the pro rata method, the recent trend in tort damage apportionment has been to assess damages according to the defendant's relative degree of fault.[264] New

---

[260] Bosin v. Minneapolis, St. Paul & S. Ste. M. R.R. Co., 183 F. Supp. 820, 823 (E.D. Wis. 1960), *aff'd sub nom.* Minneapolis, St. P. & S. Ste. M. R.R. Co. v. Fond du Lac, 297 F.2d 583 (7th Cir. 1961); 3 J. MOORE, W. TAGGART & J. WICKER, MOORE'S FEDERAL PRACTICE ¶ 14.04, at 14–26 (2d ed. 1985). One commentator has argued that courts usually should sever the contribution action from the government's CERCLA suit in order to permit the government to recover its response costs in the shortest possible time. Note, *The Right to Contribution for Response Costs Under CERCLA*, 60 NOTRE DAME L. REV. 345, 367 (1985).

[261] 42 U.S.C. 9613(f) (1982).

[262] RESTATEMENT (SECOND) OF TORTS § 886A n.h (1979). *See also* UNIFORM CONTRIBUTION AMONG TORTFEASORS ACT § 1(b). Under this method, if ten parties are liable for an injury, each is responsible for ten percent of the damage award, regardless of its relative fault.

[263] 42 U.S.C. § 9607(a) (1982).

[264] RESTATEMENT (SECOND) OF TORTS § 886A n.h (1979). Thus, if two people cause an injury and one is 75 percent responsible, that person will be liable for 75 percent of the damages. At least one state trial court has apportioned damages in a contribution action under state hazardous waste cleanup act. In Advance Circuits, Inc. v. Carriere Properties, File No. 84-3316 & -4591 (Minn. Dist. Ct. 1987), waste generators that had generated wastes dumped at a particular site paid the cost of cleaning the site and filed a contribution action against the site's owners and operators. The court held that the generators were entitled to recover 70 percent of the cleanup costs from the defendants. The court based its apportionment on its findings that: (1) the defendants' "actions and inactions were the substantial and material contributing cause of the release and threat of release of hazardous substances at the site"; (2) almost all the hazardous materials on the site were by-products of the defendants' operations on the site; (3) the defendants were "completely and uniquely responsible for the care exercised in the treatment and storage of materials"; (4) the defendants were uncooperative with the government agen-

subsection 113(f) adopts this more equitable approach. It provides: "In resolving contribution claims, the court may allocate response costs among liable parties using such equitable factors as the court determines are appropriate."[265] This language evidently contemplates that courts will consider relative fault, which is consistent both with EPA's position concerning allocation of liability[266] and with the developing area of federal contribution actions.[267] This result also addresses the concern expressed by some legislators in considering CERCLA that joint and several liability for response costs "may be expedient, but it certainly flies in the face of fundamental fairness and equity."[268]

---

cies in dealing with the problem; and (5) one of the defendants had represented himself as being "well-trained and expert in the area of recycling" and "implicitly aware of the hazardous nature of the substances he was accepting and treating." On the other hand, the court found that the plaintiff generators had (1) contributed to the release only by relying on the defendants' representations concerning proper handling of the hazardous materials; (2) delivered their hazardous wastes to the defendants; and (3) cooperated with the government officials.

[265] 42 U.S.C. § 9613(f) (1982).

[266] *See infra* notes 271–73 and accompanying text. In a speech before the Environmental Risks for Lenders Conference, an EPA Deputy Associate Enforcement Counsel stated that, before suing a lender/owner under § 107, EPA will attempt to determine whether the lender is an innocent owner. If EPA believes that a lender is innocent, EPA will not file suit against it or will enter into a de minimis settlement with it pursuant to § 122(g). 42 U.S.C.A. § 9622(g) (West Supp. 1987). Leifer Paper, *supra* note 26, at 5.

[267] At least one federal district court has indicated that it would apportion CERCLA damages in a contribution action. United States v. South Carolina Recycling & Disposal, Inc., 653 F. Supp. 984, 995 (D.S.C. 1984) ("[A]rbitrary or theoretical means of cost apportionment do not diminish the indivisibility of the underlying harm, and are matters more appropriately considered in an action for contribution between responsible parties after plaintiff has been made whole."). *See also* Kohr v. Allegheny Airlines, 504 F.2d 400, 405 (7th Cir. 1974), *cert. denied*, 421 U.S. 978 (1975) ("[T]he better rule is that of contribution and indemnity on a comparative negligence basis."); Gomes v. Brodhurst, 394 F.2d 465, 469 (3d Cir. 1968) ("On balance the equities inherent in a comparative negligence system convince us of its ultimate merit."); McLean v. Alexander, 449 F. Supp. 1251, 1272 (D. Del. 1978), *rev'd on other grounds*, 599 F.2d 1190 (3d Cir. 1979) ("[E]quity demands consideration of the relative degrees of fault painfully evident in this case."). *Cf.* Professional Beauty Supply, Inc. v. National Beauty Supply, Inc., 594 F.2d 1179, 1182 n.4 (8th Cir. 1979) (using pro rata apportionment of damages because the determination of exact degrees of fault in a complicated antitrust case would be difficult and because pro rata contribution serves as a more effective deterrent than comparative shares); Herzfeld v. Laventhol, Krekstein, Horwath & Horwath, 378 F. Supp. 112, 136 (S.D.N.Y. 1974), *aff'd in part, rev'd in part on other grounds*, 540 F.2d 27 (2d Cir. 1976) (using pro rata apportionment because the determination of comparative shares would be too confusing in a securities fraud case in which one defendant prepared and disseminated misleading report and was encouraged to do so by the other defendant who was equally aware of the misleading nature of the report).

Two commentators have argued that, in addition to assessing the relative degrees of fault in causing the harm, courts apportioning liability also should consider the degree to which each party cooperated with the government in cleaning the site. Garber, *supra* note 143, at 387–88; Note, *The Right to Contribution for Response Costs Under CERCLA*, *supra* note 260, at 365–66.

[268] 1980 U.S. CODE CONG. & ADMIN. NEWS 6143. *See also* 126 CONG. REC. 30,972

SARA permits a court to go beyond the regular apportionment of damages rule, however, to distribute CERCLA liability more equitably. Under the usual apportionment rule, the plaintiff in a contribution action can recover only that portion of the damages that the defendants caused.[269] Thus, if a potential defendant cannot be joined because it no longer exists, is insolvent, or is otherwise unavailable, that party's potential contribution is simply lost. Under these circumstances, the plaintiff must bear the portion of damages attributable to that party's actions, as well as the portion attributable to its own share of fault.

Subsection 113(b), however, may provide courts with sufficient latitude to avoid this result. In some cases, it will be equitable and appropriate for a court to apportion the full amount of damages among the parties to the contribution action even if not all responsible persons have been joined. Such a result is particularly appropriate when an attractive, though relatively blameless, defendant, such as a lender/owner, is saddled with a judgment for the full amount of response costs. Furthermore, the possibility that the court will spread the entire liability among the parties to the contribution action creates an incentive for the defendants to locate other responsible parties. Expansion of the apportionment rule will help to assure that the costs of hazardous waste releases are borne by the parties that cause them and will spread some of the transaction costs that the defendant in the original response cost recovery action otherwise would have to bear. Moreover, the probability of a negotiated settlement increases as the percentage of potentially responsible parties joined in the contribution action increases.

A liberal interpretation of subsection 113(f) is consistent with new section 122, which authorizes EPA to negotiate settlements. To promote expeditious settlements, SARA authorizes EPA to prepare a nonbinding preliminary allocation of responsibility (NBAR) that allocates 100 percent of the responsibility among the potentially responsible parties specified in subsection 107(a).

(1980) (statement of Sen. Helms (R-N.C.)) ("Retention of joint and several liability in S. 1480 received intense and well-deserved criticism from a number of sources, since it could impose financial responsibility for massive costs and damages awards on persons who contributed only minimally (if at all) to a release or injury.").

[269] RESTATEMENT (SECOND) OF TORTS § 886A(2) (1979) ("No tortfeasor can be required to make contribution beyond his own equitable share of the liability."). In comment c, however, the Restatement states that "the court may be expected to do what is fair and equitable under the circumstances" if, for example, one of the joint tortfeasors is insolvent or is outside of the court's jurisdiction. *Id.* § 886A at comment c.

The statute directs EPA in preparing the NBAR to consider *inter alia* "ability to pay" and "inequities and aggravating factors."[270] The inclusion of these considerations contemplates an apportionment of liability based on considerations other than fault and confirms that the entire liability may be assessed against fewer than all of the potentially responsible parties.

Although EPA is still in the process of preparing guidelines governing the allocation of liability for owners and operators,[271] it has indicated that liability for settlement purposes will not necessarily be determined on a strict fault basis. EPA has stated, for example, that allocation of liability to owners and operators "is a case-specific decision based upon consideration of the settlement criteria," such as those listed above.[272] Of course, the settlement procedures and considerations are designed to benefit EPA by expediting settlements, thereby reducing transaction costs, and not necessarily to benefit potentially responsible parties. New section 122 does reflect, however, that response costs may be allocated among less than all the responsible parties.[273]

Even if a court deciding a CERCLA contribution action is willing to apportion liability based on relative degrees of fault and to spread the entire liability among the parties before it, a lender/owner that files a contribution action may still confront significant obstacles. Hundreds of potentially responsible parties may have contributed to the waste site. Joining a substantial percentage of them will place an enormous discovery and litigation burden on the lender/owner. Additionally, if the lender/

---

[270] 42 U.S.C.A. § 9622(e)(3)(A) (West Supp. 1987). *See also* Superfund Program; Non-binding Preliminary Allocations of Responsibility (NBAR), 52 Fed. Reg. 19,919 (1987).

[271] The EPA Office of Enforcement and Compliance Monitoring has stated that guidelines for owners will be published in the Federal Register in April of 1988. Telephone interview with Anna Thode, EPA Office of Enforcement and Compliance Monitoring, Waste Enforcement Division (Mar. 8, 1988).

[272] Superfund Program; Non-binding Preliminary Allocations of Responsibility, *supra* note 270, at 19,920. The EPA comments reflect an intent to incorporate the types of factors normally considered by a court applying a pure apportionment theory. EPA's request for comments on the NBAR procedure states, for example, that:

> In general, owner/operator culpability is a significant factor in determining the percentage of responsibility to be allocated. For example, a commercial owner and/or operator that managed waste badly should receive a higher allocation than a passive, noncommercial landowner that doesn't qualify as innocent. . . . The relative allocation among successive owners and/or operators may be determined, where all other circumstances are equal, by the relative length of time each owned and/or operated the site.

Id. New section 122, however, reveals that these factors are to be tempered by equitable considerations. 42 U.S.C.A. § 9622 (West Supp. 1987).

[273] 42 U.S.C.A. § 9622 (West Supp. 1987).

owner impleads these parties in an EPA cost recovery action, a court probably will sever the contribution action from the main action pursuant to Rule 42(b),[274] on the ground that EPA needs to recover its response costs as quickly as possible to replenish its fund for cleanups. This efficiency rationale was a primary reason for creating joint and several liability. Thus, the lender/owner's recovery will be delayed, and it incurs the cost of a second trial. In this situation, the lender/owner has strong incentive to avoid litigation by entering into a negotiated settlement.

Because CERCLA does not specifically address the scope of liability, courts have the opportunity to create federal common law tailored to serve CERCLA's goals. The enormity of the task of rehabilitating thousands of waste sites justifies imposing joint and several liability in an EPA cost recovery action because this liability scheme minimizes both litigation costs and the time for recovery of response costs, thereby furthering the cleanup effort. Perhaps most important, the specter of joint and several liability creates a strong incentive for a party that was only minimally involved with the site, such as a lender/owner, to settle.

If a party is held jointly and severally liable, waste site cleanups and the related cost recovery actions will not be adversely affected if courts creating the new CERCLA contribution rules are flexible in their awards. Courts should reject pro rata apportionment of liability as being too inflexible. Instead, courts should apportion liability according to the parties' relative degrees of fault. In this way, a party is more likely to internalize the full costs of its conduct, and would not be forced to internalize more than the share of costs attributable to its conduct. Courts also should reject the equally inflexible rule that the defendant in the initial cost recovery action must ultimately bear liability not only for its actions, but also for the actions of all other responsible parties that were not joined in the contribution action. By adopting a flexible, fact-oriented approach to liability apportionment, courts can increase the effectiveness and fairness of the CERCLA liability scheme.

---

[274] FED. R. CIV. P. 42(b).

V. CONCLUSION

Dumping hazardous wastes formerly appeared to be a relatively cheap and convenient disposal method. It was cheap and convenient, however, only because of the limited information available concerning the effects of dumping. Participants in hazardous waste-related industries were not required to absorb the cost of negative externalities generated by dumping or, more important, to prevent or to reduce such externalities by employing safe disposal methods. Catastrophes such as Love Canal have focused national attention on the problem by dramatically demonstrating the threat to human life and to the environment posed by many waste sites. CERCLA authorizes EPA to clean a leaking hazardous waste site and to recover its response costs from those that have benefited from the dumping and cleanup activities, including lender/owners.

The prospect of CERCLA liability surprised most lenders, although the potential for environmentally related liability existed under other federal and state acts. The prospect is particularly alarming to lenders because the magnitude of that liability often will substantially exceed a lender's investment in the land. Furthermore, some lenders will feel trapped by the liability provisions because they did not acquire the property in a voluntary, arms-length transaction. Normally, the lender will have acquired the land only after other attempts to collect the debt failed or appeared to be fruitless because of the debtor's poor financial position. Moreover, the lender may have acquired the land before CERCLA's effective date, so that the liability was unanticipated.

While creating an uproar in the lending industry, CERCLA imposes no greater liability than does established common law. A lender is not liable under CERCLA because of its security interest in the polluted land; it is liable because of its ownership and control of the land. In fact, far from imposing greater liability than the common law, CERCLA imposes a lower standard because it provides a third-party defense. A truly innocent lender/owner, one that acquired the land without notice of the hazardous materials, exercised due care after discovering them, and disclosed their existence when selling the land, is not liable under CERCLA. In contrast, the common law generally imposes liability on an owner for hazardous artificial conditions

on its land regardless of whether it created them or exercised due care.

SARA has made it significantly more difficult for lenders to qualify for the safe harbor of the third-party defense. Institutional lenders often will have difficulty satisfying SARA's sliding standard for pre-acquisition inquiry, which is weighted to reflect an owner's knowledge and expertise. Therefore, to avoid CERCLA liability as an owner, a lender must conduct a thorough title and environmental inspection of a parcel of land before acquiring it.[275] Unless it does so, it may find itself stuck in a quagmire from which it will be costly to escape. It is unlikely that CERCLA liability can be avoided by acquiring the property in the name of a nominee because courts have not permitted defendants to hide behind corporate veils.[276]

SARA established that lender/owners are potentially liable for CERCLA response costs. The exact scope of that liability, however, has not been defined. This Article suggests that joint and several liability is the most appropriate and effective liability scheme to apply in response cost recovery actions. That scheme must be tempered, however, by flexibility in the reallocation of liability in the related contribution action. In recognition of the differing degrees of moral culpability and active contribution to the waste site, courts should apportion liability based on relative degrees of fault, rather than on a rigid pro rata basis. Additionally, courts should allocate the full amount of damages among the parties to the contribution action, rather than require the plaintiff to bear the portion of damages attributable to insolvent or otherwise unavailable responsible parties. The courts possess the requisite power to adopt these rules as they create the federal common law of CERCLA, and in so doing they will further CERCLA's goals.

---

[275] For a discussion of the precautions a lender should take before acquiring land through foreclosure or a negotiated settlement, see, e.g., Levitas & Hughes, *Hazardous Waste Issues in Real Estate Transactions*, 38 MERCER L. REV. 581 (1987); Richman & Stukane, *supra* note 13, at 13; Shea, *Protecting Lenders Against Environmental Risks*, May 1987 PRAC. REAL EST. LAW. 11.

[276] *See supra* note 190.

ARTICLE

LEGALITY AND DISCRETION IN THE DISTRIBUTION OF CRIMINAL SANCTIONS

PAUL H. ROBINSON*

The judicial system now responds to criminal conduct in two rather divergent steps. A judge or jury first determines if a defendant should be held liable for a criminal offense. If so, then the judge or jury goes on to choose a penalty. Precise rules, designed to ensure fairness and predictability, govern the first stage, liability assignment. In the second stage, sentencing, however, judges and juries exercise broad discretion in meting out sanctions.

In this Article, Professor Robinson argues that both liability assignment and sentencing are part of a single process of punishing criminal behavior and should be made more uniform. Decisionmakers would receive greater discretion when they assign liability and more guidance when they prescribe sanctions. Criminal codes, he contends, should explicitly give judges and juries the flexibility to incorporate normative judgments into their decisions on liability. Sentencing guidelines, however, should be more detailed and precise. Professor Robinson offers several strategies to address the practical difficulties of drafting sentencing guidelines that firmly direct judges and juries in dealing with complex issues, and he illustrates how criminal codes could give decisionmakers more discretion in certain circumstances.

The criminal law is unique in allowing courts to impose sanctions only when a defendant has violated a precise and unambiguous written rule enacted before the defendant committed the acts that constitute an offense. Courts have enforced this "legality principle" through doctrines that prohibit ex post facto and vague laws,[1] judicial creation of offenses,[2] and common law offenses,[3] as well as through the rule that penal statutes be

* Distinguished Professor of Law, Rutgers University School of Law—Camden. B.S., Rensselaer Polytechnic Institute, 1970; J.D., University of California at Los Angeles, 1973; LL.M., Harvard University, 1974; Dipl. Legal Stud., Cambridge University, 1976; former Commissioner, United States Sentencing Commission.

An earlier draft of this Article was presented at the Conference on Reform of the Criminal Law, The Inns of Court, London, England, in July, 1987.

[1] *See, e.g.*, Calder v. Bull, 3 U.S. (3 Dall.) 386 (1798) (ex post facto laws); Connally v. General Constr. Co., 269 U.S. 385 (1926) (vague laws).

[2] *See, e.g.*, Viereck v. United States, 318 U.S. 236 (1942). Many states have codified a prohibition against judicial creation of offenses. *See, e.g.*, DEL. CODE ANN. tit. 11, § 202(a) (1979); HAW. REV. STAT. § 701-102(1) (1976); N.J. STAT. ANN. § 2C:1-5(a) (West 1982).

[3] *See, e.g.*, ILL. ANN. STAT. ch. 38, para. 1–3 (Smith-Hurd 1972 & Supp. 1985); LA. REV. STAT. ANN. § 14:7 (West 1974); N.J. STAT. ANN. § 2C:1-5(a) (West 1982).

strictly construed.[4] Courts and commentators have ascribed a variety of purposes to the principle. It is meant to assure fair prior notice of criminal sanctions and to enable the public to predict to some small extent how courts will apply the law.[5] It also is designed to assure that the legislative branch decides what conduct should be criminal[6] and to limit the potential for abuse of discretion.[7] While commentators do not always include it as a traditional purpose of the legality principle, another important effect is to assure some degree of uniformity among decisionmakers—both judges and juries—in imposing criminal sanctions in similar cases.

Unfortunately, the virtues of the legality principle create their own vices. Prior and precise written rules make decisions on liability more predictable and uniform than they would be if decisionmakers enjoyed more discretion, but such rules also tend to leave decisionmakers less able to adapt solutions to individual circumstances. Fixed rules may leave the criminal law unable to account for new forms of criminal conduct[8] and

[4] *See, e.g.*, Busic v. United States, 446 U.S. 398, 406–07 (1980); *see generally* Hall, *Strict or Liberal Construction of Criminal Statutes*, 48 HARV. L. REV. 748 (1935).

[5] Fair prior notice requires only the fair opportunity for notice, not actual notice. The requirement that a defendant have such opportunity for notice in order to be prosecuted both advances fairness notions and enhances the deterrent effect of sanctions. *See* Palmer v. City of Euclid, 402 U.S. 544, 546 (1971) (per curiam) (the "underlying principle is that no man shall be held criminally responsible for conduct which he could not reasonably understand to be proscribed," quoting United States v. Harriss, 347 U.S. 612, 617 (1954)); G. WILLIAMS, CRIMINAL LAW: THE GENERAL PART § 184, at 575 (2d ed. 1961) (if citizen cannot ascertain the law beforehand, punishment for a breach of that law is "purposeless cruelty").

[6] A vague or ambiguous statute shifts the authority to decide what conduct should be criminal to the judicial branch as courts are called upon to determine the statute's scope. Courts more explicitly assert the power to brand conduct as criminal when they punish common law and other judicially created offenses.

[7] *See, e.g.*, Smith v. Goguen, 415 U.S. 566, 572–73 (1974) (statute prohibiting contemptuous treatment of the American flag declared void for vagueness under fourteenth amendment, because Constitution requires legislatures to set "reasonably clear guidelines for law enforcement officials and triers of fact in order to prevent 'arbitrary and discriminatory enforcement'" of criminal statutes); H. PACKER, THE LIMITS OF THE CRIMINAL SANCTION 91–102 (1968) (institutional mechanisms to limit judicial law-making include vagueness doctrine, canon of strict construction, and presumption against preventive detention). Broad discretion invites abuse by the judicial branch. *See, e.g.*, Papachristou v. City of Jacksonville, 405 U.S. 156, 162 (1972) (vagueness encourages arbitrary convictions as well as arbitrary arrests). The ultimate goal is equality in application. *See* State v. Ragland, 105 N.J. 189, 209, 519 A.2d 1361, 1371 (1986) (a fundamental principle of Western democracy is a "government of laws and not of men"; to satisfy this principle, all elements of the criminal justice system strive toward "equal application of the law to all accused"); *Papachristou*, 405 U.S. at 171 (1972) ("The rule of law, evenly applied to minorities as well as majorities, to the poor as well as the rich, is the great mucilage that holds society together.").

[8] A precise listing of prohibited weapons, for example, serves legality interests well but may prohibit punishment of a person who develops a "creative" weapon. *See, e.g.*, State v. Lee, 96 N.J. 156, 475 A.2d 31 (1984).

therefore unable to keep up with the "malicious ingenuity of mankind."[9] Similarly, the principle often does not keep pace with society's growing understanding of human frailties, for which the criminal law attempts to account by properly shaping mitigating factors and excusing conditions.[10] Precise prior written rules also make it difficult for the legal system to account for unusual offenses,[11] common offenses committed in unusual contexts,[12] and the infinite combinations of factors that may be relevant in determining liability and its degree.[13]

The constraints dictated by the legality principle burden courts even in simple, commonplace cases. Precision tends to spawn technicalities that frustrate effective justice.[14] The legality principle frequently makes proving guilt more difficult and costly.[15] It tends to exclude moral judgments, which typically

---

[9] Commonwealth v. Taylor, 5 Binn. 277, 281 (Pa. 1812).

[10] *See, e.g.*, State v. Diana, 24 Wash. App. 908, 913, 604 P.2d 1312, 1315 (1979) (remanding conviction so that trial court could hear evidence on medical necessity defense, a defense the defendant had not raised at trial, because "substantial medical and legal developments relating to the medical attributes of marijuana" had been "widely publicized" since the trial); *cf.* Eddings v. Oklahoma, 455 U.S. 104 (1982) (in the penalty phase of a capital case, jurors cannot be limited to consideration of a predetermined statutory list of mitigating factors).

[11] *See, e.g.*, Baker v. State, 215 Ark. 851, 223 S.W.2d 809 (1949) (posing corpse to make it appear alive found to be criminally indecent treatment of dead body); State v. Bradbury, 136 Me. 347, 9 A.2d 657 (1939) (cremation of corpse in house furnace).

[12] For example, while fraud does not normally endanger human life, it may if committed by a person who purports to provide medical treatment. *See, e.g.*, People v. Phillips, 64 Cal. 2d 574, 414 P.2d 353, 51 Cal. Rptr. 225 (1966) (in order to induce patient to accept treatment from him, chiropractor advised eight-year-old girl who suffered from eye cancer to forego medical treatment that might have cured her or prolonged her life).

[13] The criminal law often leaves courts poorly equipped to account for combinations of exculpatory factors, for example. *See, e.g.*, United States v. Hill, 655 F.2d 512 (3d Cir. 1983) (subnormality combined with entrapment); Marx v. State, 291 Ark. 325, 724 S.W.2d 456 (1987) (coercion combined with other personal problems).

[14] *See, e.g.*, People v. Nunez, 162 Cal. App. 3d 280, 208 Cal. Rptr. 450 (1984). The court in *Nunez* considered whether the defendant had violated the California Penal Code, which provides: "Every person undergoing a life sentence in a state prison of this state, who, with malice aforethought, commits an assault upon the person of another . . . is punishable with death or life imprisonment without possibility of parole." CAL. PENAL CODE § 4500 (West 1982). The court read the statute literally and held that it did not apply to the defendant because although he assaulted someone while he was serving a life sentence, he was not yet serving the sentence *in state prison* at the time; he was still in county jail. *Nunez*, 162 Cal. App. 3d at 283, 208 Cal. Rptr. at 452. The statute was subsequently amended to cover such cases. CAL. PENAL CODE § 4500 (West Cum. Supp. 1988) (as amended effective January 1, 1987). *Cf.* United States v. Gray, 633 F. Supp. 1311 (D. Mont. 1986) (construing statute that prohibited threats against judges' family members to apply also to threats against judges themselves; illogical for Congress to have intended to exclude judges themselves from the statute's protection).

[15] For a discussion of the relative advantages of increasing the burden of proof to reflect a heightened concern for precision, see McMillan v. Pennsylvania, 477 U.S. 79, ———, 106 S. Ct. 2411, 2416 (1986). In *McMillan*, the defendant's "visible possession of a firearm" gave rise to a mandatory minimum prison sentence. The Commonwealth

cannot be expressed in precise language, in favor of intricately detailed fixed rules that frequently give results inconsistent with the community's moral assessment.[16] Finally, because a fixed rule can accomodate only a limited degree of factual detail, it tends to group meaningfully different cases into a single factual category and treat them as if they were identical.[17] Affording discretion to decisionmakers avoids many of the difficulties that precise written rules create.

Given competing interests of legality and discretion, one might expect the legal system to strike a balance between them at each stage of criminal adjudication.[18] Instead, current practice creates a dichotomy. Decisionmakers have extremely limited discretion when they assign liability, while they enjoy very broad discretion when they sentence. This Article questions whether this dramatic difference is justifiable. Liability assignment is, after all, only the first of two necessary steps in distributing criminal sanctions. As the traditional statement of the legality principle provides, "*nullum* crimen *sine lege*, *nulla* poena *sine lege*."[19] The rationales that support precise written

---

had characterized the circumstance of possessing a visible firearm as a sentencing factor rather than an element of the crime charged. A bare majority of the court rejected the argument that the state had to prove this circumstance beyond a reasonable doubt. *Id.* In dissent, four justices argued that a specific component of a criminal transaction that gives rise to a special stigma and a special punishment should be treated as if it were an element of the offense, i.e., subject to the requirement of proof beyond a reasonable doubt. *Id.* at 2421 (Marshall, J., dissenting).

[16] *Compare* LA. REV. STAT. ANN. § 14:19 (West 1986) (reasonable and apparently necessary force is justified to prevent a forcible offense against the person) *with* N.J. STAT. ANN. § 2C:3-4 (West 1982), *as amended by* Act of May 15, 1987, ch. 120, 1987 N.J. SESS. LAW SERV. 44 (West) (detailed listing of circumstances that justify the use of force and deadly force in self-defense).

[17] *See, e.g.*, N.J. STAT. ANN. § 2C:3-7(b)(2)(d) (West 1982) (law enforcement officer justified in using deadly force to effect an arrest for burglary of a dwelling); Tennessee v. Garner, 471 U.S. 1 (1985) (deadly force not justified to prevent the escape of all burglars).

[18] For example, the legal system gives judges some discretion to create liability when they interpret criminal statutes. The system bars judges from interpreting statutes so as to create new liability, but it permits them to impose a supposedly pre-existing, legislatively-created liability for the first time. The distinction between these two types of statutory interpretation—one permissible and one not—obviously is fuzzy, but the legal system tolerates it because it gives decisionmakers needed flexibility without openly repudiating the legality principle. Professor Kadish illustrates these alternative explanations in his classic comparison of two cases decided by the California Supreme Court within a few years of each other: People v. Sobiek, 30 Cal. App. 3d 458, 106 Cal. Rptr. 519 (Cal. Dist. Ct. App. 1973) and Keeler v. Superior Court, 2 Cal. 3d 619, 470 P.2d 617, 87 Cal. Rptr. 481 (1970). S. KADISH & M. PAULSEN, CRIMINAL LAW AND ITS PROCESSES: CASES AND MATERIALS 173–81 (3d ed. 1975).

[19] "[N]o conduct may be held criminal unless it is precisely described in a penal law. . . . [N]o person may be punished except in pursuance of a statute which prescribes a penalty." J. HALL, GENERAL PRINCIPLES OF CRIMINAL LAW 28 (2d ed. 1960) (emphasis added).

rules governing assignment of liability and its degree apply as well to criminal sentencing. A lack of precise, prior, written sentencing rules undercuts the interests that the legality principle protects as severely as does a lack of precise, prior, written liability rules. Broad sentencing discretion, ranging from probation to the statutory maximum, can do violence to notions of fair notice, as when a judge imposes a severe sanction on a defendant who reasonably expects only a trivial one.[20] A judge can improperly usurp the legislature's criminalization decision by imposing a trivial sanction for a crime for which the legislature has recently called for a more severe one, for example.[21] Further, discretionary sentencing can lead to abuse, as each judge is free to impose any sentence that he or she feels is appropriate.[22] Finally, discretion in sentencing can produce disparity in the distribution of criminal sanctions in similar cases just as easily as can discretion in the assignment of liability.[23]

Indeed, beyond the formal finding of conviction, the determination of criminal liability and its degree means little without the imposition of sanctions. The two are interdependent stages in the single process of determining who will receive how much sanction. There is little reason to insist that decisionmakers strictly adhere to legality when they determine whether a particular offender is liable for a Grade B felony or a Grade B misdemeanor, for example, if the court has discretion to give the same sentence no matter which grade of offense it assigns

---

[20] The "proportionality" cases give many examples of severe sanctions for conduct that may reasonably be considered to be of minor seriousness. *See, e.g.*, Hutto v. Davis, 454 U.S. 370 (1982) (forty years imprisonment and $20,000 fine for sale of nine ounces of marijuana); O'Neil v. Vermont, 144 U.S. 323 (1892) (fifty-four years imprisonment for selling intoxicating liquor without authority); Ford v. State, 182 Ind. App. 224, 394 N.E.2d 250 (1979) ($10,000 fine for tending a store in which obscene materials were sold until the owner returned).

[21] A judge might, for example, impose a sentence that fails to reflect recent changes in public perceptions of the seriousness of drunk driving or insider trading.

[22] *See infra* note 124.

[23] Indeed, unwarranted disparity has been more fully documented in sentencing than in the assignment of liability; sentencing disparity was the driving force behind the current sentencing reform movement. *See, e.g.*, S. REP. No. 225, 98th Cong., 1st Sess. 65 (1983), *reprinted in* 1984 U.S. CODE CONG. & ADMIN. NEWS 3182, 3248 ("The shameful disparity in criminal sentences is a major flaw in the existing criminal justice system, and makes it clear that the system is ripe for reform.") [hereinafter S. REP. No. 225]; Seymour, *1972 Sentencing Study for the Southern District of New York*, 45 N.Y. ST. B.J. 163, *reprinted in* 119 CONG. REC. 6060 (1973); A. PARTRIDGE & W. ELDRIDGE, THE SECOND CIRCUIT SENTENCING STUDY, A REPORT TO THE JUDGES (1974).

to the defendant's acts.[24] If the distribution of sanctions is to be discretionary at the sentencing stage, society has benefited little from the strict adherence to legality at the liability assignment stage. It is true that current practice bestows the benefits of legality on innocent people and denies it only to criminals. But that seems a questionable distinction upon which to rely.[25] Legality advances societal interests, rather than merely those of innocent defendants. By assuring a fair opportunity for notice of what conduct is prohibited, minimizing the potential for abuse of discretion, and reserving the criminalization authority to the legislature, rather than de facto delegation to the judicial or executive branches, it maintains the credibility of the criminal justice process. And it deters crime as much as possible by making clear what conduct is prohibited.

A skeptic might speculate that the legal system's requirement that decisionmakers observe legality rules when they assign liability is only "window dressing" designed to shield a disregard for legality concerns in sentencing. One litigant has suggested, for example, that legislatures and courts sometimes characterize issues as sentencing rather than liability matters in order to sidestep stringent legality requirements.[26] Alternatively, one might speculate that a need for vast discretion in sentencing stems from the legal system's strict adherence to legality principle when liability is assigned. Because liability is assigned under a fixed-rule system of criminal law that does not reflect the sophistication of our sanctioning judgments, decisionmakers need discretion in sentencing to make needed adjustments to the results generated by the rigid process of assigning liability. Each of these critiques of the legality regime justifies an investigation into whether the processes of assigning liability and sentencing can be made more uniform in generally adhering to

---

[24] Obviously, the greater offense will carry a greater maximum sentence, and a judge will be able to impose a more severe sanction for the felony than for the misdemeanor. But judges so rarely impose maximum sentences that the maximum associated with a particular crime grade does not significantly constrain sentencing discretion.

[25] Jerome Hall writes:

> It is easy to imagine why some proposed solutions of difficult problems took the form of urging the retention of *nullum crimen* but abandonment of *nulla poena*, for it is the former which protects the mass of respectable citizens, while the latter, which bars complete individualization of "treatment," affects only proven criminals. But if anything can be done to any convicted person, the guarantee of legality has in fact vanished entirely.

J. HALL, *supra* note 19, at 55–56.

[26] The defendant hinted at this argument in McMillan v. Pennsylvania, 477 U.S. 79, 106 S. Ct. 2411 (1986). *See supra* note 15.

the legality principle and yet permitting limited discretion where appropriate. This Article undertakes such a study. It will examine both the possiblity of introducing greater legality into sentencing and the possibility of allowing greater flexibility in liability assignment.

While the legal system long has formally committed itself to the principle of legality, the past few centuries have seen a gradual but powerful trend toward a higher standard of legality. The American criminal code reform movement of the last twenty-five years is in large part a movement toward greater satisfaction of legality: codifying previously common law doctrines,[27] increasing the precision with which offense elements are defined,[28] and consolidating and broadening offense definitions to provide more comprehensive coverage.[29] The modern American sentencing reform movement also has produced increased adherence to legality. Vast discretion in sentencing is being limited and, within those limits, guided.[30] However, these two modern reform movements are only the most recent evidence of a longstanding trend of apparently glacial force.[31] The trend toward greater legality has existed from early common law[32] and is evident in other societies, such as the Soviet system, which initially rejected the "bourgeois principle" of legality.[33]

If greater legality is an inevitable trend in mankind's attempt to govern itself better, it is a matter of interest to speculate on how the criminal codes and sentencing systems of the future will move toward increased legality and at what rate.[34] If past developments are any indication, increased legality will not be

---

[27] *See, e.g.*, MODEL PENAL CODE §§ 2.06 (complicity), 3.02–.09 (justifiable use of force), 5.01 (attempt) (1985).

[28] *See* Robinson & Grall, *Element Analysis in Defining Criminal Liability: The Model Penal Code and Beyond*, 35 STAN. L. REV. 681, 685–702 (1983).

[29] *See, e.g.*, MODEL PENAL CODE § 223.1 (1985) (consolidated theft offenses).

[30] *See generally* State v. Yarbough, 100 N.J. 627, 498 A.2d 1239 (1985) (discussing attempts to limit discretion over consecutive and concurrent sentences); State v. Roth, 95 N.J. 334, 345–56, 471 A.2d 370, 375–82 (1984) (discussing sentencing reform generally).

[31] Historical development has rarely seen a permanent reduction in legality. It underwent a temporary setback, for example, earlier in this century when observers who viewed the criminal law primarily as a rehabilitative vehicle advocated broader discretion in sentencing. *See generally* J. HALL, *supra* note 19, at 55–58.

[32] *See, e.g.*, Robinson, *A Brief History of Distinctions in Criminal Culpability*, 31 HASTINGS L.J. 815 (1980). *See generally* J. HALL, *supra* note 19, at 25–55 (discussing the origins and development of the legality principle).

[33] P. SOLOMON, SOVIET CRIMINOLOGISTS AND CRIMINAL POLICY 22–27 (1978).

[34] *See, e.g.*, Robinson, *Four Predictions for the Criminal Law of 2043* (forthcoming in volume 19 RUTGERS L.J. (1988)); Robinson, *A Sentencing System for the 21st Century?*, 66 TEX. L. REV. 1 (1987) [hereinafter Robinson, *21st Century?*].

obtained by means that undercut effective justice,[35] but rather by new mechanisms that promote legality without sacrificing effective justice.

This Article explores existing and potential mechanisms that can better accommodate important legality concerns with the discretion necessary for effective justice. Part I outlines the types of issues in liability assignment and sentencing that judges and juries can best resolve by exercising a carefully circumscribed discretion. Part II describes techniques that drafters of criminal codes and sentencing guidelines can employ to minimize the discretion that judges and juries need by simplifying the way codes and guidelines invite those decisionmakers to analyze the cases that they face. Part III elaborates methods that drafters can use to guide how decisionmakers exercise the discretion they should be given. Finally, Part IV illustrates how criminal codes can allow judges and juries to make normative judgments, which require discretion, while promoting legality concerns as much as possible.

I. Avoiding Unnecessary Discretion

The virtues of the legality principle justify a presumption against discretion. Discretion ought to be available, but only in those cases where circumstances justify it, because discretion has definite and significant costs. It permits decisionmakers to impose criminal sanctions without fair notice, to usurp legislative authority over what conduct should be criminal, to abuse power, and to reach inconsistent results in similar cases.

In certain instances, however, effective and fair distribution of criminal sanctions requires decisionmakers to have some discretion. First, some types of cases and factors are simply unforeseeable.[36] Second, the circumstances affecting the distribution of sanctions, and the manner in which offenses combine, are sometimes so complex that legislators may experience difficulty drafting workable rules to take them into account.[37] Finally, proper distribution of criminal sanctions commonly requires decisionmakers to make judgments about abstract

[35] "Effective justice" or similar phrases here refer to two objectives of the criminal justice system, just punishment and utilitarian crime control.

[36] *See supra* notes 11–12.

[37] Criminal conduct often comprises a combination of various component actions. *See infra* notes 61–64 and accompanying text.

factors, judgments that a concrete rule cannot fully guide. The "vulnerability" of a victim, for example, may be highly relevant, but is an abstract concept that may not be easily reduced to specific factual criteria. Similarly, determining when a risk that a defendant has created is "unjustifiable" involves a normative value judgment that cannot be reduced to a set of specific rules. Discretion may be necessary to account for such unforeseeable, complex, or abstract factors. However, if these are the only factors that create a need for discretion, discretion ought *not* be permitted when these factors are not involved. That is, it should not be allowed when the circumstances are foreseeable, when codifiers can draft governing rules that are sufficiently simple to be workable, and when abstract factors or normative judgments are not at issue. Cases that involve such factors are the exception rather than the rule, so discretion should also be infrequent.

Overall, criminal codes succeed in limiting discretion to instances where it is necessary. The rules and doctrines that implement the legality principle for liability assignment generally assure that conviction is permitted only for offenses defined in the code, and only upon proof of the specific criteria provided in the offense definition. There are, however, some significant pockets of discretion, even in substantive criminal law. These are discussed in Part III of this Article.

Sentencing systems are less effective at limiting discretion to cases in which it is necessary. Frequently, statutes limit discretion only by prescribing a maximum sentence. Less often, they also prescribe a statutory minimum. Unnecessary sentencing discretion could be avoided by the adoption of binding sentencing rules that determine usual sentences but allow discretion where such flexibility is needed, i.e., for unforeseeable, complex, or abstract issues.

Binding sentencing rules that retain discretion for *unforeseeable* issues are relatively easy to construct by allowing a departure from the rules if an unforeseen circumstance arises. The Sentencing Reform Act of 1984,[38] which creates a sentencing commission to promulgate sentencing guidelines for the federal system, adopts such a mechanism.[39] Decisionmakers may depart from the sentences prescribed in the guidelines only if a circum-

[38] Pub. L. No. 98-473, 98 Stat. 1837 (codified as amended in scattered sections of 2, 5, 7, 8, 10, 12, 14, 15, 16, 18, 19, 20, 21, 22, 23, 25, 26, 28, 29, 30, 31, 41, 42, 43, 48, 49, 50 U.S.C. (Supp. III 1985)).

[39] 28 U.S.C.A. § 991 (West Supp. 1987).

stance arises that the Sentencing Commission did not consider adequately when it drafted the guidelines.[40] Because Congress also directed the Commission to take "every important factor relevant to sentencing" into account in drafting its guidelines,[41] departures due to unconsidered factors should rarely arise.

In order to indicate when a decisionmaker may appropriately depart from one of the guidelines, the Commission must describe the factors it considered when it formulated each. If the guidelines specify the factors that the Commission considered, then when others arise that it did not adequately address judges can identify them. Unfortunately, the federal guidelines frequently fail to provide this information.[42]

Binding sentencing rules that give decisionmakers discretion when *complex* issues arise are more difficult to draft than those that provide for discretion when unforeseen circumstances exist. Because comprehensive sentencing systems tend to present decisionmakers with many complex issues, affording decisionmakers discretion in every case that involves complexity may leave them bound by the sentencing guidelines in relatively few cases. In practice, however, much of the complexity commonly associated with comprehensive and binding sentencing rules may be unnecessary. Part II of this Article describes some of the structural devices that reduce complexity and thereby minimize the need for discretion.

A sentencing system can allow decisionmakers the discretion needed to deal effectively with *abstract* factors, yet remain binding in other respects, by describing those factors in an appropriately abstract form. For example, sentencing guidelines can instruct decisionmakers to consider such abstract factors as a victim's vulnerability or the justifiability of a risk that the defendant took, without insisting on detailed, concrete criteria for these concepts. Thus, decisionmakers will retain discretion to determine the presence and degree of the factor, but all other aspects of the sentencing system will be fixed and binding. This approach, of selective "pockets" of discretion, provides greater protection for legality interests than broad, generalized sentencing discretion, yet provides discretion for the specific issues for

---

[40] 18 U.S.C. § 3553(b) (Supp. IV 1986).

[41] S. REP. No. 225, *supra* note 23, at 169.

[42] *See* Dissenting View of Commissioner Paul H. Robinson on the Promulgation of Sentencing Guidelines by the United States Sentencing Commission, 52 Fed. Reg. 18,121, 18,123, 41 Crim. L. Rep. (BNA) 3174, 3177 (1987).

which it is needed. Moreover, where discretion is provided, the guidance mechanisms described in Part III can effectively control and guide decisionmakers' discretion in a way that minimizes injury to legality interests.

Decisionmakers similarly need selective discretion for specific issues when they assign liability, but state codes often do not give it to them. The common use of specific rules, based on objective facts, may increase consistency and may even generate proper results in many cases. Nevertheless, such fixed rules can also generate improper results in many other cases. For example, many states prescribe specific rules to govern when and how much force a person may use defensively. Some indirectly authorize homeowners to use deadly force against burglars.[43] While such a specific rule should produce relatively uniform liability findings, it also makes an improper result likely. For instance, use of deadly force may be grossly disproportionate if it is used in response to a twelve-year-old thief breaking into a first-floor garage. A more general standard, that a person may use force proportional to the harm threatened, would extend only limited discretion to decisionmakers but would give them the flexibility they need to avoid improper results. Modern codes frequently use specific rules rather than general, more conceptual standards to govern abstract concepts.[44] In each instance, such rules may further the legality interests of precision and uniformity,[45] but they may distort the proper distribution of criminal sanctions.

Studies on jury nullification indicate that jurors frequently exercise their nullification power to circumvent specific rules when they believe that applying them would conflict with broad normative notions of justice. For example, patterns of nullification have been identified in cases involving self-defense, trivial harms, accidental or inadvertent conduct, and extreme in-

---

[43] *See, e.g.*, N.J. STAT. ANN. § 2C:3-4(b)(2) (West 1982). *See generally* 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 142(f)(5) n.48 (1984) (use of deadly force for arrest for "any felony").

[44] *See, e.g.*, MODEL PENAL CODE §§ 3.04(2)(a) -(b), 3.05(2), 3.06(3) (1985). *See generally* 2 P. ROBINSON, CRIMINAL LAW DEFENSES §§ 131(d)(2) & nn.46 & 49, 142(f) (1984) (discussing difficulties).

[45] One may question, however, whether such precise rules further the legality interest of providing fair notice. Only criminal law professors are likely to master and remember the many special rules, and it is unlikely that even they would be guided by those rules on the spur of the moment, as, for example, when an assailant attacks them.

toxication.[46] It is just these factors—discussed in Part III—that thoughtful code drafters have chosen to express as abstract concepts rather than as specific rules. The correlation is no accident. Presumably, the code drafters saw the strong moral component of these issues and sought to preserve and incorporate, rather than override, the jury's moral judgment.

A drafting approach that allows juries to exercise their judgment in cases that involve moral issues is not only insightful but also practical. If the criminal code fails to permit moral judgments where appropriate, the system risks being ignored or subverted on these issues and, after losing credibility here, is likely to be ignored or undermined in other instances, as well. Sentencing guidelines entail at least as great a risk of subversion. Sentencing judges, who are accustomed to nearly absolute sentencing discretion, may attempt to subvert guideline sentences that they believe are improper.[47] Indeed, because sentencing judges may appreciate the extent of their power more fully than do trial jurors, judges may be more able and willing to undercut the formal rules. Jurors may be in greater awe of their oath to uphold the law than judges are.

Although abstract standards are useful for commonly understood concepts such as a victim's vulnerablility or the propor-

---

[46] H. KALVEN & H. ZEISEL, THE AMERICAN JURY 221–41, 258–85, 324–38 (1966). The data that Kalven and Zeisel compiled shows that much jury nullification is the result of: the jury's "expanded version of the law of self-defense," the tendency of the jury to excuse the actor when the victim participates in the crime, and the jury's perception that some offenses are too minimal to be dealt with by law. *Id.* at 221–85. Other factors that influence juries to acquit involve fact situations in which the defendant has already been sufficiently punished, where the punishment threatened is too severe, where a co-defendant has received preferential treatment, where the police or prosecutor has employed improper practices, where the crime involved accidental or inadvertent conduct, and where an intoxicated defendant acted completely out of character. *Id.* at 301–38. Juries are also more likely to refuse to convict a defendant charged with relatively unpopular laws such as game laws, liquor laws, and gambling laws. *Id.* at 286–97.

[47] For example, the District of Columbia has a mandatory sentencing statute that requires automatic prison terms for drug dealers. Judge Bruce Beaudin of the District's Superior Court views the law as

> rob[bing] him of his role as a judge who fits punishment to the crime, and he considers it his duty to fight back, sentencing people the way he sees fit and not according to the blanket call for prison that underlies the sentencing statute. The judge said, he chafes at anything "that is going to tie my hands on something where I can't make an exception in an individual case."

Washington Post, May 13, 1986, at A8, col. 1. A similar situation is found in Minnesota where a 1986 study by The Minneapolis Star and Tribune indicated that 49% of those offenders who might have been sentenced to prison avoided it through plea bargains with prosecutors or lenient sentences by judges. Minneapolis Star and Tribune, Oct. 6, 1986, at A1, col. 1.

tionality of defensive force, such standards should not be used for a non-intuitive public policy issue. Unless the issue is intuitive, the decisionmaker is not likely to know what the competing policy interests are, or how they are to be assessed and compared.[48] A statute that makes "unreasonable restraint of trade" a crime,[49] for example, would be likely to generate wildly disparate verdicts unless the statute precisely defines "unreasonable restraint."[50] The issue is not intuitive; it depends on economic conditions and analyses, and a judge or juror cannot reasonably be expected to master it without a good deal of guidance. If legislators prefer not to supply detailed definitions for non-intuitive concepts, then they should not deal with them through the criminal law but through other mechanisms.

Laudably, American criminal codes generally use only abstract concepts of a kind that judges and juries intuitively understand. Many code provisions concerning abstract concepts use explicitly conceptual code language,[51] while others use commentary to make clear that, whatever the face of the statute may suggest, an intuitive judgment is expected.[52] Modern codes rarely ask the decisionmaker to make a non-intuitive policy judgment from a vague standard.[53] Sentencing systems, by contrast, rarely if ever distinguish between abstract and non-abstract concepts or their intuitive or non-intuitive nature. They generally give decisionmakers broad discretion on all issues.

---

[48] A standard that calls for an intuitive judgment may further the notice interests of the legality principle even though it may be vague, if the community shares a common intuitive sense of justice on the point. "Statutes which embody well-settled common law norms stand on a footing somewhat different than statutes which attempt to circumscribe conduct newly proscribed. . . . Because the former group merely reiterate customary normative standards, the prohibitory language need not be drawn with the precision that a newly-conceived interdiction might require." Commonwealth v. Lewis, 307 Pa. Super. 468, 475, 453 A.2d 982, 985 (1982).

[49] *See* 15 U.S.C. § 1 (1982); Standard Oil Co. of New Jersey v. United States, 221 U.S. 1 (1911).

[50] Even the courts disagree as to the proper interpretation to be given this concept. *See* 7 P. AREEDA, ANTITRUST LAW ¶¶ 1500–1509 (1986).

[51] *See infra* text accompanying notes 70–77.

[52] *See, e.g.*, MONT. CODE ANN. § 45-1-102 (1983); N.Y. PENAL LAW § 1.05 (Consol. 1984).

[53] The Model Penal Code generally provides very specific formulations for non-intuitive or counter-intuitive rules. For example, on the issue of the scope of self defense, the Code provides specific defensive force limitations designed to minimize the net physical harm in an encounter even at the expense of the right that the defendant would otherwise have to use necessary and proportionate force. *See, e.g.*, MODEL PENAL CODE § 3.04(2)(a)(i) (1985) (prohibition against the use of force to resist an unlawful arrest); *id.* § 3.04(2)(b) (retreat rule); *id.* § 3.07(2)(b)(ii) (precluding use of deadly force by a citizen to effect arrest unless the citizen is assisting a peace officer) (1985). *See generally* 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 142(f)(5) (1984).

To summarize, criminal codes generally minimize unnecessary discretion. Indeed, they may go too far by providing detailed rules that fail to embody the normative judgments that justice demands. In contrast, sentencing systems generally give decisionmakers unnecessarily broad discretion. Policymakers can further interests of legality and still maintain necessary sentencing discretion by drafting sentencing rules that (1) allow departures from the usual sentence for unforeseen circumstances, (2) permit discretion on issues of complexity but minimize the need for such discretion by simplifying sentencing rules through the drafting and structuring techniques described in Part II, (3) define factors in conceptual rather than specific factual terms where necessary to capture the essential meaning of an intuitive abstract concept, but use the mechanisms described in Part III to guide the discretion inherent in such conceptual standards, and (4) otherwise require decisionmakers to follow precise sentencing rules.

At least in the early years of a binding sentencing system, modest sentencing guideline ranges that allow a limited exercise of judicial discretion might serve appropriately as a "catch-all" mechanism. The Sentencing Reform Act of 1984, for example, permits modest guideline ranges, in which the maximum sentence cannot exceed the minimum by more than 25% of the minimum.[54] Unfortunately, the more common approach to sentencing guidelines—to provide broad sentencing ranges, easy departures, only presumptive sentences, or a combination of these—is not fully sensitive to the virtues of the legality principle.

Similarly, the common practice of allowing decisionmakers to depart completely from sentencing guidelines as soon as one aspect of a case they face differs from the typical case those guidelines envision tends to thwart legality interests. A better approach is to limit a departure to the extent to which the atypical characteristic is relevant to the sentence. In many instances, decisionmakers might be given restricted discretion to alter the value of that component of the sentencing formula or the weight it may be given in calculating the final guideline sentence. The federal guidelines could take this approach, for example, in directing how judges should treat a criminal history score, a device that quantifies the extent of an offender's prior

[54] 28 U.S.C.A. § 994(b)(2) (West Supp. 1987).

criminal conduct.[55] The guidelines now provide that after elaborate calculation of the score, the sentencing judge may depart from the guideline range if he or she believes that the score calculated does not "adequately reflect" the seriousness of the defendant's prior criminal conduct.[56] Yet there seems to be little justification for dropping all restrictions on the judge's sentencing discretion because some aspect of one factor relevant to determining the appropriate sentence does not comport with the typical case the guidelines envision. It would be more appropriate to allow the judge discretion to alter the criminal history score within a restricted range, but to ensure that the judge remains bound by all aspects of the sentencing system unrelated to the problematic criminal history score.

II. STRUCTURAL MECHANISMS FOR REDUCING THE NEED FOR DISCRETION

As Part I concedes, some discretion in the distribution of criminal sanctions is necessary. Some cases cannot be foreseen, and abstract and normative factors are frequently distorted if reduced to a more specific and concrete form. In many cases, however, sentencing guidelines can use mechanisms other than discretion to address the problem of complexity. Legislators can develop mechanisms that simplify both sentencing issues and sentencing guidelines. This Part explores some of those structural mechanisms.

A. *The Use of Non-Overlapping Components of Criminal Conduct*

One of a sentencing judge's most significant discretionary powers is the authority to decide whether terms for multiple

[55] The federal guidelines' current scoring system for criminal history is contained in UNITED STATES SENTENCING COMMISSION, FEDERAL SENTENCING GUIDELINES MANUAL § 4A1.1, .2 (1987). *But see infra* note 56.

[56] UNITED STATES SENTENCING COMMISSION, FEDERAL SENTENCING GUIDELINES MANUAL § 4A1.3 (1987). The criminal history score is determined by a point system that takes into account prior sentences, whether the defendant committed the instant offense while under any criminal justice sentence, and whether the defendant committed the instant offense less than two years after release from imprisonment on certain prior sentences. *Id.* § 4A1.1.

offenses are to be served concurrently or consecutively.[57] That single decision can wipe out a host of finely drawn judgments that may have gone into calculating the sentence for each offense. Presently, a judge balances the nuances of each offense that a defendant has committed and arrives at an appropriate sentence for each. The judge must then decide whether the offender's sentences should run consecutively, and thus provide a sentence much higher than either sentence alone, or run concurrently, and thus provide a sanction no more severe than the higher sentence alone. Neither choice is likely to be fully satisfactory. A consecutive sentence is frequently too harsh, while imposing the more severe sentence alone trivializes the offense with the lesser sanction.[58]

The problem of concurrent versus consecutive sentences[59] is commonly viewed as a sentencing problem, but its cause and its solution can both be found in the definition of criminal offenses. The problem is common because even modern criminal offenses are frequently defined so that they overlap. Assume, for example, that a bank employee forged a check on another's account to buy an expensive car, and then falsified the bank records to hide the transaction. Under the United States Code, for example, he could be charged and convicted of any, all, or some combination of: forgery, uttering a forged check, fraud, false statements, and embezzlement.[60] To decide which of the sentences for these offenses should be consecutive and which concurrent requires a largely arbitrary judgment that vests broad discretionary power in the sentencing judge. An alternative approach is to isolate the basic components of the conduct that are relevant to a determination of its seriousness. In this case,

---

[57] *See, e.g.*, CAL. PENAL CODE § 669 (West 1970); D.C. CODE ANN. § 23-112 (1981) (consecutive sentences "unless the court . . . expressly provides otherwise"); GA. CODE ANN. § 17-10-10 (1987) (concurrent sentences "unless otherwise expressly provided"); MINN. STAT. § 609.15 (1986); N.Y. PENAL LAW § 70.25 (McKinney 1987).

[58] Many, if not most, judges recognize this problem and covertly manipulate the system to avoid improper results. They impose a sentence on the first count that is higher than they would have imposed if there were only one count and then impose that same higher sentence on the second count to run concurrently.

[59] For a discussion of the related but distinct problem of making an adjustment for multiple offenses of the same or similar type, as in the case of multiple instances of fraudulent checks, see Robinson, *21st Century?*, *supra* note 34, at 45–47.

[60] *See* 18 U.S.C. § 472 (1982) (uttering a "falsely made, forged, counterfeited, or altered obligation or other security of the United States"); *id.* §§ 471–495 (forgery); *id.* § 656 (theft, embezzlement, or misapplication by bank officer or employee); *id.* § 1001 (fraud, false statements or entries generally); *id.* § 1005 (fraud and false statements by a Federal Reserve bank employee); *id.* § 1344(1) (defrauding a federally chartered or insured financial institution).

the bank employee has unlawfully *taken* a certain amount. The employee's *breach of fiduciary duty* makes this taking worse than a taking of the same amount where there is no such breach. One might also argue that the means of the taking, through *false statements* in the falsification of records, also makes the conduct worse than a taking through a breach of duty without such falsification. In other words, the taking, the breach, and the falsification might be analyzed as the underlying components of the criminal conduct, so that a separate, additional sanction—a consecutive sentence—could properly be applied to each component. A sentencing system that begins with such non-overlapping components has a good chance of defining rules that avoid broad consecutive versus concurrent discretion.

Analyzing criminal conduct through its non-overlapping components also minimizes the complexity sentencing officials must face by reducing the wide variety of existing offenses to combinations of a few basic components. Current offenses essentially combine a limited number of common components. The components of an unlawful taking, a breach of fiduciary duty, and a false statement, which make up the offenses described above, are also components of an array of other offenses, each of which is comprised of a unique combination of components.

The number of possible combinations of components increases exponentially with the number of components. Five basic components can be combined in different ways to generate thirty-one different combinations; ten different components can make 1,024 combinations.[61] The reverse is also true; a total of 1,024 combinations can be reduced to ten basic components combined in different ways. Thus, a thousand offenses might be reducible to a manageable number of basic offense components.[62] The use of common basic components to describe most criminal conduct could dramatically simplify the system and thus reduce the discretion that is currently used to compensate for the complexity of current offense definitions.

The use of non-overlapping offense components has a third important advantage; it allows for independent manipulation of

---

[61] If N = the number of factors and C = the maximum number of factors in a given combination, then the number of possible combinations of N factors equals the summation of N!/C!(N-C)! for each value of C (C through 0).

[62] The ratio of reduction obviously will not be as dramatic as 1,024 to 10, because the system of offenses does not represent every combination of components; it contains only selected combinations.

different components of a single offense. In the case of robbery, for example, the amount of force used or the personal injury caused, risked, or threatened is normally the most important determinant of the overall seriousness of the crime. Of less significance is the value of the property taken. All other things being equal, a robbery of $5 million is generally viewed as more serious than a robbery of $50. Consequently, while other factors may also be relevant,[63] the overall seriousness of a robbery may be assessed with some accuracy by taking account of these two primary components—the extent of the injury caused, risked, or threatened and the value of the property taken. Each of these two components can have an almost infinite number of values, from trivial injury to death, from valueless to priceless. Analyzing these two components separately would allow sentencing officials to resolve a large variety of robbery cases with little complexity. A criminal system that requires decisionmakers to categorize an entire offense in one guideline category or another either generates overwhelming complexity, if it attempts to be comprehensive, or fails to account for variation of individual components, if it attempts to be selective. Such systems commonly are compelled, therefore, to give courts broad discretion in assessing the seriousness of the conduct in each case.[64]

B. *The Use of General Principles of Adjustment*

A notion analogous to generalized components can be applied to issues other than the objective components of an offense. The generalization of common offense and offender characteristics can simplify the distribution of criminal sanctions in a similar manner. In modern American criminal codes, for example, a general principle of liability commonly sets the maximum sanction for an unsuccessful attempt at a given proportion of the maximum sanction for the completed offense.[65] A sentencing system might take a similar approach, so that the amount of sanction for an unsuccessful attempt would be set at a certain percentage of the sanction for the completed offense. The pre-

[63] The seriousness of a robbery may also vary with such factors as the extent of psychological injury and the extent of restraint on freedom of movement, for example.

[64] The use of non-overlapping components of criminal conduct is discussed in greater detail in Robinson, *21st Century?*, *supra* note 34, at 27–32.

[65] *See, e.g.*, ALASKA STAT. § 11.31.100(e) (1983); CAL. PENAL CODE § 664 (West 1970).

sumptive value of that fraction, which represents the general relation between the attempted and completed offense, may be determined by intuitive judgment or empirical study or speculation. Whatever the basis, once the relation has been determined for a few situations, such as for attempted rape versus rape and attempted murder versus murder, the same fraction may be used in other situations as well, such as attempted theft versus theft, unless there is some reason to change it.[66] A rule that a person who completes a crime will receive a more severe sanction than someone who voluntarily stops before completion can logically be expected to deter some completions. The general principle could be defined once, with great sophistication and detail if necessary, and then applied to the many situations in which it arises, without the need to redefine it each time.

Similar general principles of adjustment may be defined for many other common factors as well. The sanction might vary, for example, according to an offender's role in the offense. Thus, a general adjusting fraction could be applied according to whether an offender was an organizer, manager, participant, or supporter. Such an approach is simpler than listing all the possible role choices under each possible offense element, and creating an unworkably long and complicated document. The definitions of the "role" categories, with assigned adjustment values, can be specified once and then applied to any offense element as needed. A base, unadjusted offense value might apply to typical participants, while their cohorts receive an adjustment up or down according to their roles.

A general adjustment for the offender's culpable state of mind could also be applied. Rather than list the sanction for each possible culpability level for every individual offense or offense component, different levels of culpability—for example, intentionally, recklessly, negligently, or blamelessly—could be defined once, and their respective adjustment factors indicated. The appropriate general adjustment could then be applied.[67]

Together, non-overlapping components of offense conduct and general principles of adjustment can significantly simplify the process of distributing criminal sanctions and thus reduce

---

[66] The policy maker might conclude that certain attempts, such as attempts to sell drugs or attempts to offer bribes, should be punished as severely as the completed act. Thus, the fraction would not apply to those cases.

[67] For illustrative provisions and a more detailed discussion of the use of general principles of adjustment, see Robinson, *21st Century?*, *supra* note 34, at 34–38, 48–53.

the discretion needed to make a more sophisticated system workable.

III. Guiding Necessary Discretion

Current sentencing systems provide very broad discretion, well beyond that which is necessary. Thus, the proposals offered in Parts I and II are useful means by which such discretion can be reduced. Even with the avoidance of unnecessary discretion described in Part I and the reduction in the need for discretion through the simplification mechanisms described in Part II, however, some sentencing discretion will remain necessary to deal with unforeseeable, abstract or normative, and inherently complex issues that cannot be reduced to workable rules. This Part outlines the mechanisms that may be used to guide the exercise of such sentencing discretion.

Current criminal codes are generally consistent with the discretion-avoidance proposals in Part I. Complexity and its accompanying discretion can be reduced with the proposals described in Part II, but it is primarily the discretion guidance mechanisms of this Part that will most significantly alter current liability assignment practice.

Concern for guiding discretion in substantive criminal law may seem peculiar because we tend to think of criminal law as the epitome of legality and the antipathy of discretion. This view is not altogether true, however. Imprecision and abstraction touch many of the most important provisions of American criminal codes. Consider the following key provisions from the Model Penal Code of the American Law Institute. It is likely that one of these provisions will be relevant in nearly every criminal case,[68] and this list of imprecise and abstract provisions is not exhaustive.[69]

---

[68] Six of the seven examples are General Part provisions that apply to a wide range of offenses. The remaining provision defines the reckless form of murder—the most egregious harm that the Code recognizes. The provisions defining culpable risk-taking in negligence and recklessness terms will themselves apply in a majority of cases. *See* Robinson & Grall, *supra* note 28, at 701 & n.87 (1983) (offenses commonly require recklessness as to one or more elements of the offense).

[69] *See, e.g.*, Model Penal Code § 5.05(2) (1985) (mitigation for inchoate offenses if conduct was so "inherently unlikely" to result in crime that neither the actor nor his conduct "presents a public danger"); *id.* § 210.3(1)(b) (murder reduced to manslaughter where committed "under the influence of extreme mental or emotional disturbance for which there is reasonable explanation or excuse"); *id.* § 211.1(2)(a) (assault by causing serious bodily injury "under circumstances manifesting extreme indifference to the value

Risk-taking is culpable under the Code if the risk is "a *substantial and unjustifiable* risk [that is] of such a nature and degree that . . . its disregard involves a *gross deviation* from the standard of conduct that a law-abiding person would observe in the actor's situation."[70] While murder normally requires an intentional or knowing killing, a reckless killing will be held murder under the Code if committed under circumstances "manifesting *extreme indifference* to the value of human life."[71] An actor's preparation to commit an offense will constitute a criminal attempt once he has taken a "*substantial step* in a course of conduct planned to culminate in his commission of the crime."[72] The Code excludes from liability conduct that causes a harm or evil "*too trivial* to warrant the condemnation of conviction."[73]

In addition to these provisions, which define or specify the grade of a criminal harm or evil, equally vague language exists in General Part provisions that specify the circumstances under which a criminal harm or evil may be attributed to an actor or his conduct. An actor's conduct is a proximate cause of a criminal result only if it "is not *too remote or accidental* in its occurrence to have a *just bearing* on the actor's liability or on the gravity of his offense."[74] An actor's commission of an offense is justified if "*the harm or evil* sought to be avoided by

---

of human life"); *id.* § 211.3 (terroristic threats include causing "serious public inconvenience"); *id.* § 212.1 (kidnapping provision prohibits removals of "substantial" distance and confinements of "substantial" duration); *id.* § 212.3 (false imprisonment involves restraint that interferes "substantially" with another's liberty); *id.* § 213.1(2), .2(2)(a) (gross sexual imposition and deviate sexual intercourse by imposition where submission is compelled by threats that would "prevent resistance by a woman of ordinary resolution").

[70] *Id.* § 2.02(2)(c), (d) (emphasis added). The Code commentary concedes that the formulation remains vague. "There is no way to state this value judgment that does not beg the question in the last analysis; the point is that the jury must evaluate the actor's conduct and determine whether it should be condemned." *Id.* § 2.02 comment 3, at 237.

[71] *Id.* § 210.2(1)(b) (emphasis added). Of this language the commentary concludes, "Whether recklessness is so extreme that it demonstrates [such] indifference is not a question that, in our view, can be further clarified; it must be left directly to the trier of the facts." MODEL PENAL CODE § 210.2 comment 2, at 29 (Tent. Draft No. 9, 1959). For similar statement, see MODEL PENAL CODE § 210.2 comment 4, at 22 (1980).

[72] MODEL PENAL CODE § 5.01(1)(c) (1985) (emphasis added). As the Code commentary notes, "Whether a particular act is a substantial step is obviously a matter of degree. To this extent, the Code retains the element of imprecision found in most of the other approaches to the preparation-attempt problem." *Id.* § 5.01 comment 6(a), at 329.

[73] *Id.* § 2.12(2) (emphasis added).

[74] *Id.* § 2.03(2)(b), (3)(b) (brackets omitted, emphasis added). As with previous provisions, the Code commentary admits that the formulation "deals only with the ultimate criterion. . . ." *Id.* § 2.03 comment 3, at 261.

such conduct *is greater than* that sought to be prevented by the law defining the offense."[75] An actor's commission of an offense may be excused if his cognitive or control functions are sufficiently impaired—for example, if he was subject to coercion that "a person of *reasonable firmness* in his situation would have been unable to resist"[76] or if, because of involuntary intoxication or mental disease or defect, he lacks "*substantial capacity* either to appreciate the criminality [wrongfulness] of his conduct or to conform his conduct to the requirements of the law."[77]

Each of these Code provisions calls upon the factfinder to do more than find facts. Each requires the judge or juror to make a judgment about an essentially normative or moral issue by drawing upon his or her own intuitive sense of justice. Thus, each of these provisions may be criticized as failing to satisfy the rationales of the legality principle. Such provisions fail to provide the public with adequate notice of specifically what will be punished; they permit a de facto delegation of the criminalization decision to judges or juries; they introduce the potential for abuse of discretion; and they increase the likelihood of inconsistent disposition of similar cases.[78] Despite such criticisms, however, provisions of this type are necessary. Such normative judgments are essential to the moral force of the criminal law, and to take away the discretion needed for such normative

---

[75] *Id.* § 3.02(1)(a) (emphasis added). The commentary observes, "It is true, of course, that the principle formulated in the Section is invariably lacking in precision. There is room for disagreement on what constitutes an evil and which of two evils is the greater." MODEL PENAL CODE § 3.02(1)(a) comment 4, at 9 (Tent. Draft No. 8, 1958). In the final version, 27 years later, the first sentence of the above quote was altered to read, "The principle formulated in this section is general." MODEL PENAL CODE § 3.02(1)(a) comment 4, at 17 (1985). The commentary argues, however, that "the lack of precision in the general rule is unavoidable if the rule is not to be improperly constricted; the situation is akin to the imprecision of such concepts as recklessness and negligence, which also call in part for weighing of conflicting values." *Id.*

[76] *Id.* § 2.09(1) (emphasis added).

[77] *Id.* § 4.01(1) (insanity) (emphasis added); *id.* § 2.08(4) (involuntary intoxication). The commentary explains, "To identify the degree of impairment with precision is, of course, impossible both verbally and logically." MODEL PENAL CODE § 4.01(1) comment 4, at 159 (Tent. Draft No. 4, 1955). The revised commentary makes a similar point.

> It was recognized, of course, that "substantial" is an open-ended concept, but its quantitative connotation was believed to be sufficiently precise for purposes of practical administration. The law is full of instances in which courts and juries are explicitly authorized to confront an issue of degree. Such an approach was deemed to be no less essential and appropriate in dealing with this issue than in dealing with the questions of recklessness and negligence.

MODEL PENAL CODE § 4.01 comment 3, at 172 (1985) (footnotes omitted).

[78] *See supra* notes 5–7.

judgments would seriously distort the proper distribution of criminal liability.

If decisionmakers are to retain necessary discretion in liability assignment and sentencing, however, the legal system must guide their exercise of that discretion in a way consistent with the underlying goals of the legality principle. In Part IV, the principles and general proposals offered here are illustrated in possible code provisions for the liability assignment provisions noted above. But many of the proposals in this Part can apply as well to to the exercise of sentencing discretion.

A. *Explicit Invitation to Make Normative Judgments on Normative Issues*

The virtues of using an abstract standard to embody a normative issue are discussed in Part I. Drafters of criminal codes and sentencing guidelines can increase decisionmakers' consistency in making moral assessments by impressing on them that the standard is meant to be a normative one that calls not merely for factfinding, but for a value judgment. If decisionmakers are not explicitly invited to exercise their judgment, they may feel restricted from doing so by other perceived legal, non-intuitive rules.[79] For example, if defense counsel suggests that a prisoner was justified in escaping to avoid an assault by another prisoner, a jury may well conclude that since the prisoner's imprisonment was lawful, it could hardly be lawful for him or her to escape. A jury instruction that simply acknowledges the existence of a defense if conduct is "justified" may not correct the jury's misperception. An instruction telling the jury that it must weigh the opposing interests to determine if the conduct is justified, is more likely to indicate to the jury that a guilty verdict is not mandated by the "lawfulness" of the imprisonment. In other words, absent a special dispensation, jurors may feel compelled to return a verdict that stays within the perceived limits of the

[79] Fletcher appears to make this claim about all rules: "relying on rules runs against the rigours of individualization." Fletcher, *The Individualization of Excusing Conditions*, 47 S. CAL. L. REV. 1269, 1305 (1974). This assumes that rules must be devoid of normative elements. As Part IV illustrates, however, one can construct rules that explicitly call for normative judgments yet guide the exercise of that judgment.

legal rules,[80] rather than one that is consistent with their intuitive notions of justice, even where the issue is meant to be resolved through a normative judgment. They may feel obliged to suppress their own moral judgments in order to fulfill their oath to "render a true verdict according to the law and the evidence."[81]

When modern codes seek a normative judgment rather than a factual finding,[82] they too often fail to say so, or say so only in the code's commentary, which is unavailable to the jury.[83] In the sentencing context, on the other hand, discretion is generally so broad that this issue never arises. Judges are rarely required to make judgments on specific issues, abstract or otherwise.

B. *Judges versus Juries as Normative Decisionmakers*

Judges and juries each display strengths and weaknesses in consistently applying abstract, normative standards consistently. Juries are likely to be less consistent than judges. Each juror typically sits for only a few cases during his or her term of service and rarely sits on similar cases during that term. As a group, the jury sits only for the case at hand. In contrast, a judge applies his or her judgment to many more cases, although even this source of uniformity is compromised in judicial districts with more than one criminal judge or where criminal case responsibility rotates frequently among judges.

On the other hand, a judge is less likely than a jury to render a normative judgment that reflects the shared community sense

---

[80] Studies indicate that jurors often misperceive the law. *See, e.g.*, Robinson, *Causing the Conditions of One's Own Defense: A Study in the Limits of Theory in Criminal Law Doctrine*, 71 VA. L. REV. 1, 52 nn.168–69, 53–54 nn.170–76 (1985).

[81] *See, e.g.*, TEX. CRIM. PROC. CODE ANN. § 35.22 (Vernon 1965) ("You and each of you do solemnly swear that in the case of the state of Texas against the defendant, you will a true verdict render according to the law and the evidence, so help you God."); *accord* ALASKA R. CRIM. P. 24(f); ARIZ. R. CRIM. P. 18.6(b). The evidence suggests that jurors take their oath very seriously and that the rate of instances of "jury nullification" is low. *See* H. KALVEN & H. ZEISEL, THE AMERICAN JURY 498 (1966); Scheflin, *Jury Nullification: The Right to Say No*, 45 S. CAL. L. REV. 168, 211–12 (1972).

[82] *See supra* notes 68–77 and accompanying text.

[83] With respect to Model Penal Code § 3.02 (lesser evils), for example, the commentary indicates that a judgment is called for and notes that "[d]eep disagreements are bound to exist over some moral issues, such as the extent to which values are absolute or relative and how far desirable ends may justify otherwise offensive means." MODEL PENAL CODE § 3.02 comment 4, at 17 (1985); *see also supra* note 75. Similar statements are contained in the commentaries to MODEL PENAL CODE § 210.2 (1980) (reckless killing "manifesting extreme indifference to value of human life" is murder), quoted *supra* text accompanying note 71, and in those to MODEL PENAL CODE § 2.03 (1985) (definitions of recklessness and negligence), quoted *supra* text accompanying note 74.

on an issue; and it is this community sense that normative standards seek to incorporate into the process.[84] Magistrates and judges are not typical members of the community. They differ significantly from the general population in education, intelligence, economic status, and political views.[85] Further, their judgments are likely to be distorted by the experience of becoming a lawyer and judge; common sense may be the first casualty of legal training. Moreover, the criminal court judge is exposed to a daily parade of the worst side of human behavior. Such exposure is likely to alter the judge's perceptions about the standard of unacceptable conduct. In judging whether conduct is "too trivial to warrant the condemnation of conviction,"[86] for example, the criminal court judge may be one of the persons in the community least able to represent the community's normative judgment at all reliably.[87]

Beyond a judge's atypical qualities and experience, he or she is at a disadvantage compared to a juror in making normative

---

[84] The community sense of justice is critical to both the efficacy and morality of punishment.

> [L]aw is ineffective in the deepest sense, indeed . . . it is hypocritical, if it imposes on the actor who has the misfortune to confront a dilemmatic choice, a standard that his judges are not prepared to affirm that they should and could comply with if their turn to face the problem should arise. Condemnation in such a case is bound to be an ineffective threat; what is, however, more significant is that it is divorced from any moral base and is unjust.

MODEL PENAL CODE § 2.09 comment 2, at 7 (Tent. Draft No. 10, 1960). Recognition of the significance of the community's norms is frequently expressed, for example, in the context of the insanity defense. *See, e.g.*, Bazelon, *The Morality of the Criminal Law*, 49 S. CAL. L. REV. 385, 396 (1976) (as the community is both the beneficiary and the victim of the insanity defense, the community, through "its traditional surrogate, the jury," should "assess criminal responsibility in light of the community standards").

[85] Judges "have the defects of their qualities. They are led away from the naked facts by the rules of evidence, by ideas which they got from books or trials, by notions on the motives of witnesses, by theories of psychology or some other 'ology' riding on the top of the current at the time, and by various mental bents." Norton, *What a Jury Is*, 16 VA. L. REV. 261, 265–66 (1930). As the Supreme Court has stated, juries therefore offer defendants special protection:

> Those who wrote our constitutions knew from history and experience that it was necessary to protect against . . . judges too responsive to the voice of higher authority. The framers of the constitution strove to create an independent judiciary but insisted upon further protection against arbitrary action. Providing an accused with the right to be tried by a jury of his peers gave him an inestimable safeguard . . . against the compliant, biased, or eccentric judge.

Duncan v. Louisiana, 391 U.S. 145, 156 (1967).

[86] MODEL PENAL CODE § 2.12(2) (1985); *see supra* text accompanying note 73.

[87] *See* H. KALVEN & H. ZEISEL, THE AMERICAN JURY 8 (1966). In summarizing arguments in favor of jury trials, the authors note that the jury's inexperience is sometimes considered an asset because it secures a fresh perception of each trial and avoids stereotypes said to infect the judicial eyes. Data collected on disagreement between judges and juries in criminal cases indicates that much disagreement is attributable to the jury's application of a "de minimis" infraction principle. *See id.* at ch. 18.

judgments because of the judge's isolated position when deciding cases. Where a shared community normative judgment is at issue, the process of expression, reaction, and response to others is critical, and much of one's judgment on such matters depends upon one's assessment of others' reactions. Juries, in contrast, are ideally suited in these respects for making normative judgments. Some writers suggest that, although magistrates would be more reliable at the sometimes technical job of factfinding, we use a lay jury system because of the importance of the more reliable normative judgments that the jury provides.[88]

On the other hand, judges have tools that increase the likelihood that their normative judgments will reflect what the legislature intended. Legislative histories, official commentaries, and prior case law applying a statute all help to explain the intended meaning of a legislature's abstract standard. These materials may describe the factors to be considered, and the formulae or method by which they are to be compared or evaluated, and may also give illustrations of fact patterns that fall within or outside the standard. Each of these mechanisms can give judges a greater opportunity to understand the intended concept and its application and thereby increase the reliability and consistency of the judgment.

One might look upon the choice between judges and juries as normative decisionmakers as an opportunity to perfect the balance between legality and the discretion needed for effective justice. To tip the scale toward greater consistency, decisionmaking might be allocated more to judges than to juries; to tip it toward better normative judgments, the decisionmaking might be allocated more to juries. Thus, the allocation of a particular

---

[88] The greater reliability of a jury in making normative judgments has been generally recognized. "It must be remembered that the jury is designed not only to understand the case, but also to reflect the community's sense of justice in deciding it." H.R. REP. No. 1076, 90th Cong., 2d Sess. 8 (1968) (the House Committee Reports accompanying the Federal Jury Selection and Service Act of 1968, Pub. L. No. 90-274, 82 Stat. 53 (codified at 28 U.S.C. §§ 1861–1871 (1982))). The jury, on the other hand, can be criticized because jurors generally lack the education and training that a judge has. *See, e.g.*, In re Boise Cascade Securities Litigation, 420 F. Supp. 99 (W.D. Wash. 1976) (striking jury demand on ground that trial to the court assured greater fairness and thereby furthered due process requirements of fifth and fourteenth amendments). As one federal trial judge concluded, "[a] jury, applying its collective wisdom, judgment and common sense to the facts of a case . . . is brighter, more astute, and more perceptive than a single judge." Zenith Radio Corp. v. Matsushita Electrical Industrial Co., 478 F. Supp. 889, 935 (E.D. Pa. 1979), *vacated sub nom.* In re Japanese Electronics Litigation, 631 F.2d 1069 (3d Cir. 1980).

issue could balance the advantage of the judge's access to guidance devices, such as legislative histories, official commentaries, or prior case law, against the problem of distorted judicial judgment on the issue. Under such balancing, one may question the wisdom of the Code's allocation of normative decisionmaking to the judge for the de minimis infraction defense.[89] On this issue, for example, few guidance mechanisms are available to the judge, and the danger of distorted judicial judgment may be at its greatest.[90]

An alternative to choosing between the respective advantages of judges and juries is to combine the advantages of each. If a jury is an inherently better normative decisionmaker but lacks the explanations and guidance available to judges, the better approach may be to leave the normative judgments to the jury but to have the jury instructions include the available guidance information. Such detailed jury instructions may not be appropriate in all cases but may be justified where a vague standard presents the central issue in a case. To avoid improper delegation of the criminalization authority to the courts, it would be best to have the criminal code, rather than the individual judge, provide the additional explanation or guidance that is to be given to the jury.[91] Such guidance might take the form of a series of illustrative applications of the provision or a description of the factors to be considered and their interrelation. The use of such mechanisms is discussed in the next two subsections and is illustrated in Part IV.

C. *The Use of Illustrations*

Many abstract or conceptual standards may benefit from illustrations that establish a point of reference. If, for example, the standard is "too remote or accidental," "too trivial," "extreme indifference," or "substantial step,"[92] a decisionmaker must first determine how remote, trivial, extreme, or substantial

---

[89] MODEL PENAL CODE § 2.12 (1985).

[90] *See infra* text accompanying notes 136–59. Ironically, the de minimis defense is the only one of the doctrines under discussion where the Code allocates the decision to the judge.

[91] If the legislature fails to provide the additional explanation and guidance, however, perhaps judges should. To the extent that such instructions would be reviewable on appeal, the practice would, over time, ameliorate the inconsistent application that vague standards inherently produce.

[92] MODEL PENAL CODE §§ 2.03(2)(b), 2.03(3)(b), 2.12(2), 210.2(1)(b), 5.01(1)(c) (1985).

the conduct or result must be to reach what the standard describes. That is, even if the decisionmaker understands the core concept of, for example, remoteness or triviality and can distinguish degrees of remoteness or triviality, the decisionmaker still must determine the quantum that the standard demands. It may be true, as the Model Penal Code commentary frequently insists, that many of the Code's standards cannot be expressed with greater specificity or clarity,[93] but illustrations of what is and is not sufficiently remote, trivial, extreme, or substantial can provide additional guidance.[94] If the decisionmaker is told when the critical point on the quantum continuum is reached, he or she can better determine whether the circumstances of the case at hand reach that point.

Illustrations have little precedent in sentencing systems.[95] They are even less common in criminal codes, but there is some precedent. The Model Penal Code, for example, lists examples of the disabilities that are to be considered as causing an involuntary act.[96] The Code similarly gives illustrations of what might be a "substantial step"—the point at which preparation to commit a crime becomes a criminal attempt.[97] Macaulay's Indian

---

[93] *See supra* notes 70–77.

[94] The Supreme Court has recognized that illustrations may save an otherwise vague provision from the void-for-vagueness doctrine. Parker v. Levy, 417 U.S. 733, 753–54 (1974).

[95] A few jurisdictions sometimes give illustrations for when the court should or should not depart from the sentencing guidelines. *See, e.g.*, UNITED STATES SENTENCING COMMISSION, SENTENCING GUIDELINES AND POLICY STATEMENTS 5K2.1 (departure policy statements on death), 5K2.7 (disruption of governmental function), 5K2.8 (extreme conduct) (1987); MINNESOTA SENTENCING GUIDELINES COMMISSION, SENTENCING GUIDELINES AND COMMENTARY, Commentary to § II.D.03 (1980) (factors that should not be used as reasons for departure).

[96] MODEL PENAL CODE § 2.01(2) (1985). At the conclusion of the list, subsection (d) provides a catch-all that is essentially the Code's standard for conduct constituting an involuntary act: "a bodily movement that otherwise is not a product of the effort or determination of the actor, either conscious or habitual." *Id.* § 2.01(2)(d).

[97] *Id.* § 5.01(2). Actually, the provision illustrates conduct that may not be held to be insubstantial. Many jurisdictions adopting attempt provisions patterned after Model Penal Code § 5.01 have not adopted the list of conduct that is not insufficient to constitute a substantial step as a matter of law. *See, e.g.*, N.J. STAT. ANN. § 2C:5-1 (West 1982); 18 PA. CONS. STAT. ANN. § 901 (Purdon 1983); WASH. REV. CODE § 9A.28.020 (1985). The Washington experience is noteworthy. In rejecting a vagueness challenge to the statute, the Supreme Court of Washington noted that "further definition of 'substantial step' by this court may be helpful." State v. Workman, 90 Wash. 2d 443, 450, 584 P.2d 382, 387 (1978). The Court cited with approval the Model Penal Code's "strongly corroborative" test and its list of examples. *Id.* at 451 & n.2, 584 P.2d at 387 & n.2. In describing jury instructions to be used in the future, however, the court merely stated: "We therefore hold it would be proper for a trial court to include in its instruction to a jury on the crime of attempt the qualifying statement that in order for conduct to be a substantial step it must be strongly corroborative of the actor's criminal purpose." *Id.* at 452, 584 P.2d at 388.

Penal Code of 1837 frequently provided concrete illustrations for most provisions.[98]

Illustrations are given a variety of weights; some are legally binding, while others are purely informative.[99] Even where the illustrations are binding on the judge, however, they do not cause jurors to reach more uniform decisions because jurors are denied access to them. Illustrations such as those given in Part IV could well prove quite useful in guiding juries in certain situations.

D. *The Use of Factors and Formulae*

The value and feasibility of illustrations is greatest for those types of issues that require a judgment regarding degree. Such questions involve a single continuum and require that a critical cut-off point be established where the core concept has been reached or exceeded. Where the issue involves a more complex evaluation illustrations are less likely to be useful since they only describe the result generated by the particular facts of the illustration. If, for example, the issue involves a complex balancing of factors, knowing the final result for one situation is of little help in other situations, unless underlying determinations about the various factors are also described. In the "lesser evils justification," for example, the decisionmaker must identify all the competing interests, assign a value to each, add the values for each of the opposing groups, and then compare these sums.[100] Knowing the facts and proper results of several justification situations is of limited usefulness in evaluating new justification situations. A single result may be consistent with any number of variations in the interim steps of the process, and the distillation of a pattern from several examples can be a difficult task. In these instances, a code or guidelines may best serve decisionmakers by identifying relevant factors and describing their interrelation.

---

[98] *See* T. MACAULAY, A PENAL CODE PREPARED BY THE INDIAN LAW COMMISSIONERS AND PUBLISHED BY COMMAND OF THE GOVERNOR GENERAL OF INDIA IN COUNCIL (1837).

[99] The involuntary act examples are binding: "The following are not voluntary acts. . . ." MODEL PENAL CODE § 2.01(2) (1985). The substantial step illustrations are only binding in a negative sense: "the following . . . shall not be held insufficient as a matter of law." *Id.* § 5.01(2). The illustrations do not direct what is to be held sufficient by the jury. The Macaulay Code illustrations, by contrast, have no binding effect.

[100] *See infra* notes 224–40 and accompanying text.

Several of the abstract Code provisions present complex issues of this type, such as the definition of culpable risk-taking.[101] To determine whether an actor's risk-taking was culpable, the decisionmaker must assess the nature of the risk by identifying, valuing, and comparing the competing interests, as in assessing justification. Once the risk is assessed, the judge or jury can then determine the unreasonableness of the risk taken and compare it to an objective standard. Such a reasonable conduct standard requires a similarly complex calculation. The actor's risk-taking is culpable only if it constitutes a "gross deviation" from conduct in which a "reasonable" or "law-abiding" person would engage.[102]

The judgment of the degree of impairment that is required to excuse culpability is also a complex assessment. The decisionmaker must assess the degree of the actor's actual cognitive or control dysfunction. This dysfunction assessment is, itself, the product of many other preliminary judgments. The decisionmaker also must assess the amount of resistance to, or compensation for, such disabilities that a reasonable person could be expected to offer. These two assessments must then be compared to determine whether the actor's failure to compensate for, or to resist the effects of, such disabilities is exculpatory. In the context of duress, for example, the required analysis includes an assessment of the extent of the coercion on the actor, the extent of coercion that "a person of reasonable firmness in [the actor's] situation would have been unable to resist,"[103] and a comparison of the two.

To assess such complex concepts as the balancing of lesser evils, the reasonableness of risk-taking, or the degree of impairment required for exculpation, the decisionmaker must be given some minimal procedure for analyzing the issue. It is inadequate to use conclusory labels as some codes do and simply advise the decisionmaker that an actor is to be exculpated if, for example, his conduct was "justified," "reasonable," or "involuntary," or the actor is an "idiot."[104]

---

[101] *See infra* notes 240–46 and accompanying text.

[102] MODEL PENAL CODE § 2.02(2)(c), (d) (1985).

[103] *Id.* § 2.09(1).

[104] *See, e.g.*, LA. REV. STAT. ANN. § 14:18(1) –:18(4) (West 1986) (defense of justification can be claimed when the conduct is a "reasonable fulfillment" of duties of public office, a "reasonable accomplishment" of an arrest, "authorized by law," or "reasonable discipline" of minors); MINN. STAT. ANN. § 609.06(1), (3) (West 1986) ("reasonable force" may be used by a public officer effecting a lawful arrest or enforcing a court

First, a code or guidelines should give decisionmakers a list of factors to consider, though it need not be exhaustive. The appropriate level of abstraction for the factors may depend on the issue. Examples of what would not be appropriate factors to consider also may prove useful. But in all cases, an identification of relevant factors can educate and guide decisionmakers in handling complex issues, just as illustrations can aid them with simpler judgments.

An explicit description of how the factors interrelate can provide further guidance. Such a formula allows the decisionmaker to allocate relevant factors to their appropriate point of relevance in the analysis. Consider the actor who causes a car accident while driving a sick child to the hospital at high speed. A decisionmaker may consider it relevant that the sick passenger is the actor's child rather than that of a stranger. But how is this relevant in determining whether the actor's risk-taking is culpable? The risk created is the same whether the passenger is his child or a stranger's. The harm sought to be avoided is the same in both cases. The kinship may be relevant, however, in determining the standard of conduct that a "reasonable" or "law-abiding" person would observe in the actor's position.[105] A reasonable person may be more inclined to drive faster if his or her own child needs medical attention. The decisionmaker may ultimately conclude that kinship with the passenger should not alter the standard of conduct in this particular situation and that the risk-taking was culpable, but the structure provided by the formula puts the issue in the proper context for such a determination.

---

order, or by a person resisting an offense against the person); VT. STAT. ANN. tit. 13, § 2305(1) (1974) (homicide is justifiable if committed "in the just" defense of one's life); V.I. CODE ANN. tit. 14, § 41 (1964) (a person about to be injured may make resistance "sufficient to prevent" an illegal attempt to take or injure property or an offense against his person).

Several modern codes adopt voluntary act provisions similar to those of the Model Penal Code, but fail to provide any definition of "voluntary" or "involuntary" acts. With regard to "idiots," *see, e.g.*, CAL. PENAL CODE § 26(2), (4) (West 1970) (idiots and persons who commit act unconsciously); NEV. REV. STAT. § 194.010(3),(6) (1977) (idiots and persons who commit act unconsciously); *accord* OKLA. STAT. ANN. tit. 21, § 152(3),(6) (West 1983). *See generally* 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 171(a), at 259 n.1 (1984) (listing statutes and identifying those that fail to define the critical term). Courts have elaborated upon these statutory provisions. *See* People v. Freeman, 61 Cal. App. 2d 110, 142 P.2d 435 (1943) (reversing, in part, because trial court failed to explain sufficiently varying degrees of unconsciousness that might provide exculpation).

[105] MODEL PENAL CODE § 2.02(2)(c), (d) (1985).

Finally, it is also useful to state any special rules or assumptions that govern the analysis. For example, it may be appropriate for judges or juries to weigh only legally-recognized interests in a lesser evils defense.[106] Similarly, in balancing competing interests, the community's assessment of the proper balance may be more relevant than the actor's personal assessment, even if the actor shows the best of faith.[107] The fact that jurisdictions may disagree on such underlying assumptions[108] suggests that decisionmakers also may disagree and, unless explicitly instructed, may adopt a view contrary to that chosen by the legislature and different from that taken by other decisionmakers in similar cases.

Few sentencing systems provide such guidance; sentencing discretion is generally so broad that the issues never arise. Some modern codes use a formula to provide guidance on such complex issues.[109] It is less common for codes to guide discretion by identifying relevant factors or through illustrations.[110]

Beyond guiding discretion at trial and sentencing, illustrations and factor lists, in particular, may increase the accuracy and consistency of other discretionary judgments in the criminal

---

[106] *See, e.g.*, 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 124(b) (1984).

[107] *See, e.g., id.* § 124(d)(1); Cleveland v. Municipality of Anchorage, 631 P.2d 1073 (Alaska 1981) (the actor's assessment of the balance of harms and evils is not relevant; the relevant assessment is that of the community).

[108] For example, many jurisdictions treat as "justified" conduct resulting from a reasonable mistake as to a justification defense while others maintain a distinction between justified (objectively proper) conduct, and excused (mistakenly justified) conduct. *Compare* State v. Kelly, 97 N.J. 178, 478 A.2d 364 (1984) (reasonably mistaken conduct is justified) *with* State v. Leidholm, 334 N.W.2d 811 (N.D. 1983) (conduct is justified where actual circumstances make it legal; it is excused where the actor believes that justifying circumstances exist when they do not).

[109] For jurisdictions adopting the Model Penal Code's lesser evils formulation, for example, see 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 124(a) n.1 (1984). Many jurisdictions adopt causation provisions similar to the Model Penal Code formula. For examples of jurisdictions modeled after sections 2.03(2)(b) and 2.03(3)(b) of the Model Penal Code see DEL CODE ANN. tit. 11, §§ 262(2), 263(2) (1987); HAW. REV. STAT. § 702-215 (1985); MONT. CODE ANN. § 45-2-201(2)(b), (3)(b) (1987); N.J. STAT. ANN. § 2C:2-3(b),(c) (West 1982); 18 PA. CONS. STAT. ANN. § 303(b), (c) (Purdon 1983). *See also* DEL. CODE ANN. tit. 11, § 231 (1987) (Model Penal Code distinctions and ordinary negligence); HAW. REV. STAT. § 702-206 (1985); MONT. CODE ANN. § 45-2-101 (33), (37), (58) (1987) (Model Penal Code distinctions minus criminal liability for culpable inadvertance); N.J. STAT. ANN. § 2C:2-2(b) (West 1982); 18 PA. CONS. STAT. ANN. § 302(b) (Purdon 1983).

[110] Where the Model Penal Code has provided illustrations, many jurisdictions otherwise adopting the Model Penal Code formulation have not adopted the illustrations. *See supra* note 97. Courts, however, have apparently found the factors approach useful. *See, e.g.*, Brown v. United States, 256 U.S. 335, 343 (1921) (retreat is a factor relevant to the defense of self-defense); State v. Hunt, 2 Ariz. App. 6, 406 P.2d 208 (1965) (discussing factors relevant to an assessment of the propriety of force used in disciplining a child).

justice process. Their use may improve decisions by police officers, grand jurors, prosecutors, and judges at such stages of prosecution as the arrest decision, indictment, and plea bargaining.[111] Given the percentage of criminal cases that never go to trial, some observers may consider this effect of guided judgments to be of greater practical importance than the effect on judges and juries at trial.[112]

These guidance mechanisms could also have an important effect in encouraging constructive legislative reform. Legislatures frequently try to assure that a particular factor is taken into account when imposing liability. The fact that an actor is in his home when he is confronted by an attacker, for example, may be viewed as important to justify the use of deadly defensive force. Similarly, the fact that a person attacked in public has an opportunity to retreat may be important to disqualify the use of deadly defensive force. More often than not, a legislative desire to consider such factors as an attack at home or an opportunity to retreat results in the creation of special rules to expand or limit the availability of self defense. Thus, special provisions often make it lawful to kill an intruder in one's home,[113] or bar deadly force if there is an opportunity to retreat.[114] Such special rules can easily result in improper outcomes, however. Killing an intruder, for example, may not always be necessary to avoid, or be proportionate in degree to, the harm threatened.[115] Likewise, use of deadly defensive force

---

[111] Vagueness encourages arbitrary arrests and prosecutions as well as arbitrary convictions. *See supra* note 7 and accompanying text.

[112] One study involving 21 states and the District of Columbia revealed that the percent of convictions obtained without trial (the guilty-plea rate) ranges from 66.5% in Pennsylvania (1974) to 97.9% in Vermont (1975). D. JONES, CRIME WITHOUT PUNISHMENT 44 (1979). Data from 24 states and the District of Columbia on 54,426 dispositions showed that only 15% of criminal prosecutions were tried before a jury (75% disposed of by guilty plea; 10% by bench trial). H. KALVEN & H. ZEISEL, THE AMERICAN JURY 18 (1966) (statistics taken from U.S. Bureau of the Census, Judicial Criminal Statistics (1945)). Of course, the decision to enter a guilty plea is often largely influenced by the defense counsel's estimation of the client's chances before a jury.

[113] *See, e.g.*, Gainer v. State, 40 Md. App. 382, 391 A.2d 856 (1978); ALA. CODE § 13A-3-23(b)(1)(A) (1982); DEL. CODE ANN. tit. 11, § 464(e)(2)(a) (1979); N.J. STAT. ANN. § 2C:3-4(b)(2)(b)(i) (West 1987).

[114] *See, e.g.*, MODEL PENAL CODE § 3.04(b)(ii)(1) (1985); ALA. CODE § 13A-3-23(b)(1) (1975); ARK. STAT. ANN. § 5-2-607(b)(1) (1987); CONN. GEN. STAT. ANN. § 53A-19(b)(1) (West 1958); DEL. CODE ANN. tit. 11, § 464(e)(2) (1987); HAW. REV. STAT. § 703-304(5)(b) (1985), ME. REV. STAT. ANN. tit. 17-A, § 108(2)(c)(3)(a) (1964).

[115] For example, the intruder may be only a tired next-door neighbor who mistakenly enters the wrong row house.

without having made an attempt to retreat may sometimes be both necessary and proportionate.[116]

Guidance devices such as illustrations and factor lists give the legislature a convenient and more appropriate mechanism by which they can express views on such specific matters. The fact that an actor is in his or her home when attacked, or has an opportunity to retreat before using deadly force, may be important factors for a jury to consider in judging the justification for the use of defensive force.[117] Thus, adding these considerations to a list of relevant factors will provide useful guidance and help avoid unduly inflexible rules.

E. *A Guiding Statement of Purposes*

A final guidance technique, traditionally used more by criminal codes than by sentencing systems, is the articulation of a set of objectives, frequently termed the "purposes of punishment," to guide the decisionmakers in interpreting and applying sanction-distribution rules. Most modern American criminal codes follow the lead of the Model Penal Code in providing such a guiding statement of objectives.[118] When faced with an ambiguous provision, courts are to adopt the interpretation that best furthers the objectives of the code.[119]

While the broad discretion that exists in sentencing makes the need for guidance especially great, modern sentencing systems

---

[116] For example, retreat may create its own risk of injury or death.

[117] As Justice Holmes concluded, "the failure to retreat is a circumstance to be considered with all the others in order to determine whether the defendant went farther than he was justified in doing; not a categorical proof of guilt." Brown v. United States, 256 U.S. 335, 343 (1921).

[118] MODEL PENAL CODE § 1.02(2) (1985); *accord* ALA. CODE § 13A-1-3 (1982 & Cum. Supp. 1984); ALASKA STAT. § 11.81.100 (1980); ARIZ. REV. STAT. ANN. § 13-101 (1978); COLO. REV. STAT. § 18-1-102 (1978 & Cum. Supp. 1982); DEL. CODE ANN. tit. 11, § 201 (1979); FLA. STAT. ANN. § 775.012 (West Cum. Supp. 1982); GA. CODE ANN. § 16-1-2 (1984); HAW. REV. STAT. § 701-103 (1976); IDAHO CODE § 18-100 (1987); ILL. ANN. STAT. ch. 38, para. 1-2 (Smith-Hurd 1972); MINN. STAT. ANN. § 609.01 (West 1987); MONT. CODE ANN. § 45-1102 (1987); NEB. REV. STAT. § 28-102 (1979); N.Y. PENAL LAW § 1.05 (McKinney 1975); OR. REV. STAT. § 161.025 (1987); 18 PA. CONS. STAT. ANN. § 104 (Purdon 1983); TEX. PENAL CODE ANN. § 1.02 (Vernon 1974); UTAH CODE ANN. § 76-1-104 (1978); WASH. REV. CODE § 9A.04.020 (1985); GUAM CODE ANN. § 1.14 (1982).

[119] *See, e.g.*, COLO. REV. STAT. § 16-1-103 (1978); DEL. CODE ANN. tit. 11, § 203 (1979); MONT. CODE ANN. § 45-1-102 (1987); N.Y. PENAL LAW § 5.00 (Consol. 1984); UTAH CODE ANN. § 76-1-106 (1978); WASH. REV. CODE § 9A.04.20 (1985).

commonly lack a statement of objectives.[120] Minnesota and Washington alone have a statement of sentencing objectives.[121] Systems that provide such statements are generally viewed as the most successful to date at reducing unwarranted sentencing disparity.[122]

A statement of objectives can reduce sentencing disparity, because judges' differing philosophies are a prime cause of such inconsistency. As the Canadian Sentencing Commission notes: "[J]udges approach similar cases in different ways. These different approaches to cases—based on different views of what [sentencing] principles should be paramount—lead to different sentences being handed down for similar offences committed by similar offenders in similar circumstances."[123] For example, where a young addict has been selling drugs to acquaintances to support his own drug habit, one judge may impose a lengthy prison sentence to deter other potential dealers. Another judge may impose probation with a condition of community drug treatment in the belief that the most effective means of avoiding future dealing by this offender is to rehabilitate him by treating his addiction. Similarly, one judge may give an embezzling bank teller a significant prison term because he deserves it for his breach of trust, and because such a sentence provides a useful deterrent threat to other tellers. But another judge may conclude that no term of imprisonment is needed since the offender will never again be in a position of trust; that is, she may focus on an incapacitative rationale for sentencing and find it lacking in this case.[124] With no principle to guide conflicting philosophies,

[120] The Model Penal Code and many state codes provide such a statement of objectives to guide sentencing, *see supra* note 118, but few modern sentencing guidelines provide anything similar. Minnesota and Washington provide such a statement, but the statements are less comprehensive than the typical criminal code provision on the subject. *See infra* notes 121–26.

[121] MINNESOTA SENTENCING GUIDELINES COMMISSION, SENTENCING GUIDELINES AND COMMENTARY (1984) ("Statement of Purpose and Principles"); Sentencing Reform Act of 1981, WASH. REV. CODE § 9.94A.340–.420.

[122] *See* A. VON HIRSCH, K. KNAPP & M. TONRY, THE SENTENCING COMMISSION AND ITS GUIDELINES, 59, 74–79, 99–101 (1987).

[123] CANADIAN SENTENCING COMMISSION, SENTENCING REFORM: A CANADIAN APPROACH xxiii (1987) [hereinafter CANADIAN REPORT]. *See also* Kennedy, *Foreward*, P. O'DONNELL, M. CHURGIN & D. CURTIS, TOWARD A JUST AND EFFECTIVE SENTENCING SYSTEM viii (1977); S. REP. NO. 225, *supra* note 23, at 38–47, *reprinted in* 1984 U.S. CODE CONG. & ADMIN. NEWS 3182, 3221–30.

[124] Similarly, just punishment concerns may prompt one judge to impose a light sentence on an unemployed young woman who passes a $90 bad check but another judge, focusing on the high recidivism rate and large number of potentially deterrable similar offenders, may follow incapacitation and deterrent rationales to impose a significant prison sentence. *See generally* Robinson, *21st Century?*, *supra* note 34, at 8–10, 15.

the inevitable result is sentencing disparity, as different judges each make different personal, ad hoc choices among conflicting objectives. For similar reasons, appellate courts also need a guiding statement of objectives when they review the appropriateness of a given sentence.[125]

There are two preconditions, however, to using a statement of objectives to guide discretion. First, the same objectives should logically guide application and interpretation of both the code and sentencing guidelines. Preferably, the drafters should begin with an articulation of the governing principles and then derive the code and guidelines from them. The effectiveness of such objectives will be undercut if different policy decisions govern liability assignment and sentencing, or if policy decisions that determine how a code or set of guidelines is drafted differ from those that underlie how the code or guidelines will be applied by decisionmakers. Logical purity in development, drafting, and application, however, is not necessary. It is enough that in the end a common principle or understanding binds all aspects of the distribution of criminal sanctions.[126]

A second and equally important precondition is the existence of an over-arching framework to govern the interrelation among the objectives sought. Unfortunately, codes and sentencing sys-

---

[125] *See, e.g.*, L. WILKINS, J. KRESS, D. GOTTFREDSON, J. CAPLIN & A. GELMAN, SENTENCING GUIDELINES: STRUCTURING JUDICIAL DISCRETION 2 (1978) ("[A]ppellate review, without an attached explicit statement of court policy, is likely to substitute the judgement of an appellate court as to the 'correct' sentence in a given case for that of the lower court.").

[126] In fact, the Minnesota guidelines' statement of philosophy was added only toward the end of the drafting process but seems consistent with the earlier drafting. For a discussion of the process by which the Minnesota Sentencing Guidelines Commission ultimately adopted a statement of sentencing objectives, see MINNESOTA SENTENCING GUIDELINES COMMISSION, THE IMPACT OF THE MINNESOTA SENTENCING GUIDELINES: THREE YEAR EVALUATION 11–14 (1984). The Model Penal Code similarly reflects a consistent philosophy throughout its provisions, even though its "purposes" provision was not the first part drafted. Section 1.02 of the Code first appeared in Tentative Draft No. 2 and was first presented to the National Law Institute's May 1954 meeting. It was not approved until 1961. *See* MODEL PENAL CODE § 1.02 history note (1980). *See generally* Kadish, *Codifiers of the Criminal Law: Wechsler's Predecessors*, 78 COLUM. L. REV. 1098 (1978); Packer, *The Model Penal Code and Beyond*, 63 COLUM. L. REV. 594 (1963); Wechsler, *The Challenge of a Model Penal Code*, 65 HARV. L. REV. 1097 (1952).

This is not to say, however, that the same principles for meting out criminal punishment should be used for every aspect of a sentencing system. *See* Robinson, *Hybrid Principles for the Distribution of Criminal Sanctions*, 82 NW. U.L. REV. 19, ——— (1987) [hereinafter Robinson, *Hybrid*] (arguing that principles that govern the method of imposing sanctions should differ from those that determine how severe a sanction should apply). Nonetheless sentencing systems should avoid inadvertent or unjustified inconsistencies in their underlying principles.

tems rarely provide this framework. Different objectives frequently conflict in application since they rely on different criteria and thus frequently suggest conflicting sentences or statutory formulations. When faced with conflicting objectives, judges have nothing to guide their decision to follow one objective at the expense of another. Thus, decisionmakers commonly may choose different objectives in different situations. That is, faced with the young drug dealer described above, the judge may opt to follow general deterrence at the expense of rehabilitation and give a long prison term. But faced with the bank embezzler, the same judge may choose special deterrence principles, at the expense of general deterrence, and give a short term.

Legislatures themselves are inconsistent in choosing among conflicting objectives when they draft criminal codes. For example, most state criminal codes maintain an insanity defense[127] because it exculpates the blameless and thus furthers just punishment, even though abolishing the defense might more effectively incapacitate dangerous offenders. The same codes, however, sacrifice just punishment for increased deterrence when they recognize strict liability.[128] At the same time, rather than increasing the threatened sanction when the temptation or inclination is greater, as a deterrence principle suggests, the codes frequently lessen punishment, in cases of provocation,[129] for example, because of the offender's reduced blameworthiness. In other words, code drafters choose to follow different purposes in different contexts. Judges will be faced with the same conflicting choices in interpreting and applying these provisions, as when a judge must decide whether the evidence justifies a jury instruction on an issue.

A skeptic may conclude that the use of such objectives or "purposes" to resolve a difficult choice or to justify a particular code formulation is attractive to legislators and judges because it appears rational but can mask a decisionmaker's true basis for a decision. The suspicion that the traditional purposes of

---

[127] *See* 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 173 n.1 (1984) (insanity); *see also id.* § 171 n.1 (involuntary act defense).

[128] *See, e.g.*, MODEL PENAL CODE § 213.6(1) (1985) (strict liability as to age in some sexual offenses). *See also* IDAHO CODE § 18-6101(1) (1987); MASS. GEN. LAWS ANN. ch. 265, § 23 (West 1986); N.C. GEN. STAT. § 14-27.2(a)(1) (1987); OKLA. STAT. ANN. tit. 21, § 1111(1) (West 1983); S.C. CODE ANN. § 16-3-655 (Law. Co-op. 1976).

[129] *See* 1 P. ROBINSON, CRIMINAL LAW DEFENSES § 102 (a) nn.2 & 9 (provocation, extreme emotional disturbance) (1984); *see also id.* § 102 n.35 (cooling period); 2 P. ROBINSON, CRIMINAL LAW DEFENSES §§ 177 n.1 (duress), 184 nn.1 & 48 (battered-wife) (1984).

punishment are a popular justification precisely because they offer such hidden flexibility, is fueled by the almost universal failure to articulate a guiding framework that gives a principled basis for each choice when purposes conflict.[130] The Model Penal Code, for example, lists the traditional purposes, directs judges to use them in interpreting the provisions of the Code and in fashioning a sentence, then, in commentary, explains that when purposes conflict they should be "justly harmonized."[131] Other codes suggest that the competing interests are to be "balanced," "blended," "accommodated," "taken account of," or "deal[t] with [such that] the public interest will be served."[132] Regardless of whether judges actually have exploited the hidden flexibility of the "purposes" to rationalize arbitrary sentences, a rational and principled system for the distribution of criminal sanctions must prescribe the interrelation between its multiple objectives if the exercise of discretion is to be adequately guided.

IV. Legality and Discretion in Determining Criminal Liability

This Part illustrates how the principles developed in Parts I through III might be translated into workable criminal code sections or jury instructions.[133] The problem of liability rules that are too specific and detailed to allow for needed discretion is discussed in Part II. This Part focuses on the reverse problem, vague standards in criminal codes. To attain the virtues of legality, codes must guide how decisionmakers apply such standards.

The previous discussion suggests the following principles for guiding discretion in abstract or vague code provisions. First, vague provisions should be tolerated only when and to the extent that the injury to legality interests is outweighed by an articulable benefit that flows from the vagueness. The introduction of a normative judgment frequently may provide such a justification. Second, vague language embodying an intuitive

---

[130] Such a principled framework is possible. *See* Robinson, *Hybrid*, *supra* note 126.

[131] Model Penal Code § 1.02 comment 4 (Tent. Draft No. 2, 1954).

[132] *See* Robinson, *Hybrid*, *supra* note 126, at 103.

[133] I have discussed elsewhere how the principles developed in Parts I through III might be translated into a workable sentencing system. *See* Robinson, *21st Century?*, *supra* note 34.

concept is more tolerable than vague language embodying non-intuitive or counter-intuitive concepts. The latter will cause greater injury to legality interests and will, therefore, be harder to justify. Third, the system should expressly state when a normative judgment is called for by the decisionmaker.[134] Fourth, illustrations or descriptions of relevant factors should guide discretion for judgments about when a legal requirement or standard is met, such as single-continuum judgments. Fifth, descriptions of relevant factors and formulae that describe how those factors relate to one another should guide discretion for complex judgments that require decisionmakers to assess and balance two or more abstract concepts, such as multiple-continuum judgments.

Guided by these principles, this Part evaluates current law formulations and constructs a codification proposal for the vague Model Penal Code provisions noted in Part III.[135]

A. *Exclusion of De Minimis Infractions*

At some point along the continuum of seriousness of an offense, society concludes that the harm or evil of an offense is too trivial to be penalized by the criminal law. What is that point? How can it best be described to a judge or jury to illustrate their task, and to minimize unjustified disparities in application?

Existing cases fail to give a clear picture of the point at which a harm is "too trivial." Shoplifting $1.59 in goods,[136] failure to register a shotgun with a barrel of 17 and 9/10 inches rather than 18 inches long,[137] possession of microscopic amounts of a controlled drug,[138] ownership of an unoccupied house without a sewer,[139] and submission of an election expenses report a few days late,[140] were all held to be offenses that were not too trivial. In the last case, however, the trial court held that the conduct was too trivial while the appellate court reversed and remanded

---

[134] *See supra* text accompanying notes 80–83.

[135] *See supra* text accompanying notes 70–77.

[136] Commonwealth v. Campbell, 273 Pa. Super. 407, 417 A.2d 712 (1980).

[137] United States v. Lamb, 294 F. Supp. 419 (E.D. Tenn. 1968).

[138] State v. Brown, 188 N.J. Super. 656, 458 A.2d 165 (N.J. Super. Ct. Law Div. 1983) (possession of .65 gram of cocaine not de minimis).

[139] Commonwealth v. Baker, 71 Pa. D. & C.2d 521 (Somerset County Ct. 1974).

[140] State v. Park, 55 Haw. 610, 525 P.2d 586 (1974).

on the issue of triviality.[141] On the other hand, possession of one ounce of marijuana,[142] possession of microscopic amounts of a controlled drug,[143] and harassment with vulgar language,[144] were all held to be too trivial.

These cases suggest that the same harm may be held trivial in one case and non-trivial in another[145] and that different judges may disagree as to whether the harm in a particular case is trivial or non-trivial,[146] and that a court may disagree with the jury.[147] The present methods of describing the concept of triviality appear to be inadequate and inconclusive.

Indeed, it is not even clear what kinds of factors are relevant to the triviality determination. Certainly objective factors, such as the extent of the harm caused, are central; but is the defendant's personal culpability relevant, as some courts suggest?[148] Do courts limit the kinds of harms that are to be considered?[149]

---

[141] *Id.* at 617, 525 P.2d at 591.

[142] People v. Davis, 55 Misc. 2d 656, 286 N.Y.S.2d 396 (N.Y. Sup. Ct. 1967).

[143] *See* State v. Schofill, 63 Haw. 77, 621 P.2d 364 (1980) (trial judge dismissed indictment where charge involved the attempted but unconsummated sale of a quarter ounce of cocaine; the Supreme Court reversed the dismissal noting that Hawaii's de minimis provision did not provide a basis for the dismissal); State v. Vance, 61 Haw. 291, 602 P.2d 933 (1979) (noting that in some cases, although not the one before the court, possession of an infinitesimal amount of narcotics that is too small to use or sell might call for a dismissal as de minimis).

[144] Commonwealth v. Houck, 233 Pa. Super. 512, 335 A.2d 389 (1975). *See infra* note 154.

[145] *Compare* State v. Brown, 188 N.J. Super. 656, 458 A.2d 165 (N.J. Super. Ct. Law Div. 1983) *with Schofill*, 63 Haw. at 77, 621 P.2d at 364 (1980) *and Vance*, 61 Haw. at 291, 602 P.2d at 933.

[146] *See supra* text accompanying notes 136–44. *See also* State v. Nevens, 197 N.J. Super. 531, 485 A.2d 345 (1984) (defendant was convicted before a judge in Municipal Court of theft for taking two bananas, an orange, an apple, and a pear from open buffet-type restaurant *after* paying for meal; Superior Court, Law Division judge reversed after trial de novo, as de minimis infraction).

[147] *See supra* note 87 (discussing jury nullification because of triviality of conduct where determination of whether nullification had occurred was made according to whether the judge disagreed with the verdict).

[148] *See, e.g.*, State v. Park, 55 Haw. 610, 525 P.2d 586 (1974). In *Park*, the following factors relating to personal culpability were considered relevant to the de minimis issue:

> . . . the background, experience and character of [the defendants] which may indicate whether they knew of, or ought to have known, the requirements of [the statute]; the knowledge on the part of [the defendents] of the consequences to be incurred by them upon violation of the statute; . . . the mitigating circumstances, if any, as to each offender; . . . and any other data which may reveal the nature and degree of culpability in the offense committed by each [defendant].

Id. at 617, 525 P.2d at 591.

[149] People v. Davis, 55 Misc. 2d 656, 286 N.Y.S.2d 396 (1967). In *Davis*, the court dismissed the indictment of a 20-year-old college student who was charged with possession of one ounce of marijuana, stating that "Of determinative significance is the fact that no member of the public has suffered in any way because of the defendant's conduct." *Id.* at 658, 286 N.Y.S.2d at 399.

From who's perspective is the harm evaluated: the defendant's? the juror's? the judge's? the legislature's? the community's?[150] Similarly, one may consider what standard a court will use to compare the triviality of an offense to the harm it causes. One jurisdiction allows a dismissal for trivial conduct when dismissal is "in furtherance of justice."[151] The Code provides slightly more guidance when it describes conduct "too trivial to warrant the condemnation of conviction."[152] Macaulay's Code takes a somewhat different approach; it refers to harm that is "so slight that no person of ordinary sense and temper would complain of such harm."[153] While the definition of "de minimis" may initially seem unworkable—the concept is fundamentally intuitive—certain guidance can be provided that will enhance the accuracy and consistency of such intuitive judgments.

Drawing on the principles developed above, a sample statute might provide:

(1) *De Minimis Infraction Defense Defined.* It is a defense that an actor's conduct caused or threatened a harm or evil that is too trivial to warrant the condemnation of criminal conviction.[154]

(2) *Factors.* In evaluating whether an actor's conduct caused or threatened a harm or evil that is "too trivial to warrant the condemnation of criminal conviction," the decisionmaker should consider, among other things, the following factors:

(a) the nature and degree of tangible harms caused or threatened,

(b) the nature and degree of intangible harms and evils caused or threatened,

---

[150] *See* State v. Vance, 61 Haw. 291, 602 P.2d 933 (1979). In dicta, the court described a drug case where the de minimis defense would be appropriate. Looking to legislative intent indicating that drug offenses were designed to regulate use and sale or transfer for use, the court suggested that possession of a microscopic and unusable amount might constitute a de minimis infraction. *Id.* at 306–08, 602 P.2d 933–44. *Cf.* State v. Brown, 188 N.J. Super. 656, 458 A.2d 165 (1973) (noting de minimis infraction originating from a civil doctrine authorizing the exercise of judicial discretion).

[151] N.Y. CRIM. PROC. LAW §§ 170.40 (motion to dismiss information), 210.40 (motion to dismiss indictment)(McKinney 1982). *See* People v. Rickert, 58 N.Y.2d 122, 446 N.E.2d 419, 459 N.Y.S.2d 734 (1983) (information charging nonsupport dismissed); People v. Tyler, 46 N.Y.2d 264, 385 N.E.2d 1231, 413 N.Y.S.2d 302 (1978) (indictments charging defendant with issuing false certificate and official misconduct dismissed under § 210.40 of the New York Criminal Procedure Law); *Davis*, 55 Misc. 2d at 656, 286 N.Y.S.2d at 396 (indictment charging college student with possession of one ounce of marijuana dismissed).

[152] MODEL PENAL CODE § 2.12(2) (1985).

[153] T. MACAULAY, *supra* note 98, § 73.

[154] This language has been taken from MODEL PENAL CODE § 2.12(2) (1985).

(c) the nature and degree of a disruption of the social order caused or threatened, and

(d) the potential that allowing a defense in this instance would undercut the criminal law's condemnation of related, more serious conduct.

(3) *Illustrations*. The following are examples of instances where a decision-maker might, but is not legally bound, to find that an actor's conduct:

(a) caused or threatened a harm or evil that is "too trivial to warrant the condemnation of criminal conviction":

(i) stealing three pieces of bubble gum,[155]

(ii) littering two spent matches,[156] and

(iii) calling another person "lower than dirt."[157]

(b) caused or threatened a harm or evil that is not "too trivial to warrant the condemnation of criminal conviction;"

(i) unlawfully selling or possessing a small amount of a controlled drug,[158]

(ii) touching another person over their implicit or explicit objection, and

(iii) borrowing another person's car for 20 minutes without permission.

Obviously, the legislature must decide on the precise content of the provision but this form may prove useful. The fact that the existing cases and statutes offer little assistance in formulating the content only confirms that the legislatures have not yet provided adequate guidance. Legislatures must answer the difficult policy questions of what factors are relevant and define standards that best capture the intuitive concept they intend.

---

[155] State v. Smith, 195 N.J. Super. 468, 480 A.2d 236 (N.J. Super. Ct. Law Div. 1984). "In the milieu of bubble gum pilferage the only cases more trivial are those involving two pieces or one. It is difficult to conclude the lawmakers would have intended the dividing line to be drawn at three." *Id.* at 476–77, 480 A.2d at 240.

[156] People v. Feldman, 73 Misc. 2d 824, 342 N.Y.S.2d 956 (N.Y. Crim. Ct. 1973).

[157] Commonwealth v. Houck, 233 Pa. Super. 512, 335 A.2d 389 (1975). In *Houck*, defendant was charged under a harassment by communication statute that prohibited the use of "lewd, lascivious, or indecent words or language" over the telephone. *Id.* at 514, 335 A.2d at 390. The trial court found that the language used by defendant, "lower than dirt" and "morally rotten," was not "lewd, lascivious, or indecent." *Id.* at 515, 335 A.2d at 390. The trial court also found that even if defendant had violated the statute, his conduct was too trivial to warrant a conviction. *Id.* The appellate court affirmed. *Id.* at 518, 335 A.2d at 391.

[158] State v. Schofill, 63 Haw. 77, 621 P.2d 364 (1980); State v. Brown, 188 N.J. Super. 656, 458 A.2d 165 (N.J. Super. Ct. Law Div. 1983).

B. *Aggravation of Reckless Killing to Murder*

While murder typically involves an intentional or knowing killing, some cases of lethal risk-taking are sufficiently egregious to be treated as murder. Common law and older statutes describe this aggravating condition as risk-taking that shows a "depraved mind,"[159] that shows an "abandoned and malignant heart,"[160] or that is "regardless of human life."[161] Modern statutes modeled after Model Penal Code section 210.2(1)(b) make a reckless killing a murder if it manifests "extreme indifference to the value of human life."[162] The commentary makes it clear that the Code seeks to adopt the common law concept; the change in language is designed only to describe that concept more clearly.[163]

Presumably all conscious killings, including those that arise from consciously risking the death of another, reflect something about the actor's valuation of human life. On a continuum of indifference to the value of human life, reckless killings are to be treated as murder if the circumstances suggest that the actor's indifference exceeds a critical point. But a given set of facts and the bare language of "depraved mind . . ."[164] or "extreme indifference . . . ,"[165] give a jury little basis to formulate an "indifference" value for the actor and then determine whether it exceeds the critical point on the continuum.

Similar to the difficulty in determining whether conduct is "too trivial . . ." under the de minimis infraction defense, decisionmakers in past cases have shown considerable disparity in determining when a reckless killing constitutes murder. Courts have found reckless murder when death results from

---

[159] *See, e.g.*, FLA. STAT. § 782.04(2) (1985); MINN. STAT. § 609.195 (1986); N.M. STAT. ANN. § 30-2-1 (1984).

[160] *See, e.g.*, CAL. PENAL CODE § 188 (West 1970); GA. CODE ANN. § 16-5-1(a),(b) (1984); IDAHO CODE § 18-4002 (1979).

[161] *See, e.g.*, MISS. CODE ANN. § 97-3-19(b) (Cum. Supp. 1983); OKLA. STAT. tit. 21, § 701.8(1) (1987); S.D. CODIFIED LAWS ANN. § 22-16-7 (Supp. 1984).

[162] MODEL PENAL CODE § 210.2(1)(b) (1980). *See, e.g.*, ALA. CODE § 13A-6-2(a)(2) (1977); ARIZ. REV. STAT. ANN. § 13-1104(A)(3) (1978); COLO. REV. STAT. § 18-3-102(1)(d) (1986); WASH. REV. CODE ANN. § 9a.32.030(1)(b) (1988).

[163] The commentary states, "Insofar as Subsection [210.2](1)(b) includes within the murder category cases of homicide caused by extreme recklessness, though without purpose to kill, it reflects both the common law and much pre-existing statutory treatment usually cast in terms of conduct evidencing a 'depraved heart regardless of human life' or some similar words." MODEL PENAL CODE § 210.2(1)(b) comment 4, at 22 (1980) (footnotes omitted).

[164] *See supra* note 159 and accompanying text.

[165] *See supra* note 162 and accompanying text.

going through a red light while driving drunk,[166] allowing an intoxicated person to drive while the defendant is asleep in the back seat,[167] grabbing the wheel while another is driving,[168] or passing on the left and hitting an oncoming vehicle while driving drunk.[169] In contrast, a reckless killing has been held not to constitute murder where the defendant drives eighty to one hundred miles per hour and collides with two vehicles as he passes in the lane for oncoming traffic.[170] The court explains that "there was no evidence in the record tending to show that defendant had by the use of any weapon indulged in vicious or brutal conduct which might support a finding of 'an abandoned and malignant heart.'"[171] Other courts have found only manslaughter, not reckless murder, where the defendant put a cord around the victim's neck to help restrain her during forcible rape,[172] or where defendants, as a joke, put straw and hot cinders on a sleeping person who then burned to death.[173] Obviously, these latter courts have a very different notion of what qualifies a reckless killing as murder.

The different perspectives taken by juries and judges reveal further disparities in liability assignment. Juries have found reckless murder where the defendant, while he has the right of way, fails to look a second time for other traffic entering the intersection,[174] or where the defendant helps another addict inject the other's normal dosage of heroin.[175] In both instances, appellate courts disagreed with the lower court's assessment of malice and reversed. Similarly, there could be cases of the reverse: where a jury declines to convict for reckless murder

---

[166] Hamilton v. Commonwealth, 560 S.W.2d 539 (Ky. 1978).
[167] State v. Trott, 190 N.C. 674, 130 S.E. 627 (1925).
[168] Gibson v. State, 476 P.2d 362 (Okla. Ct. App. 1971).
[169] Maxon v. State, 177 Wis. 319, 187 N.W. 753 (1922).
[170] State v. Chalmers, 100 Ariz. 70, 411 P.2d 448 (1966).
[171] *Id.* at 75, 411 P.2d at 452.
[172] State v. Bolsinger, 699 P.2d 1214 (Utah 1985).
[173] Errington and Others' Case, 168 Eng. Rep. 1133 (New Castle Sp. Assizes 1838).
[174] Commonwealth v. Ushka, 130 Pa. Super. 600, 198 A. 465 (1938).
[175] Commonwealth v. Bowden, 456 Pa. 278, 309 A.2d 714 (1973) (Pennsylvania Supreme Court affirmed order in arrest of judgment on the ground that the trial evidence, presented to a judge sitting without a jury, was insufficient as a matter of law to establish malice). *See also* Commonwealth v. Parker, 458 Pa. 381, 327 A.2d 128 (1974). In *Parker*, the court reversed a conviction where the defendant had inserted a needle into the victim's vein, and the victim injected heroin into his own body and died. The court determined that there was no malice although the defendant knew the deceased had been drinking and that a mix of heroin and alcohol could be fatal. *Id.* at 385, 388, 327 A.2d at 130, 131.

but where an appellate court would permit it, if the state could appeal the acquittal.[176]

Without guidance, juries and courts will continue to apply disparate notions of what will aggravate a reckless killing to murder. Decisionmakers need statutory guidance to formulate an "indifference" value for the case and to determine where the critical point lies on the "indifference" continuum. Providing a list of factors that might be relevant, and identifying each as mitigating or aggravating will allow the decisionmaker to evaluate a given offender's "indifference". The determination of the critical point on the continuum may be accomplished by giving illustrations of what constitutes both adequate and inadequate "indifference" for murder. Though the decisionmaker will still have to make a difficult intuitive judgment, that evaluation will reflect community values more accurately and will reach results more consistent with similar cases.

A statute designed to provide such guidance might take the following form:

(1) *Reckless Murder Defined.* An actor who engages in conduct by which he recklessly causes the death of another human being is guilty of murder if the actor's conduct manifests an extreme indifference to the value of human life.[177]

(2) *Factors.* In evaluating whether an actor's conduct "manifests an extreme indifference to the value of human life," the decisionmaker should consider, among other things, the following:

(a) the degree of the risk of death created,[178]

---

[176] *Compare* Commonwealth v. Malone, 354 Pa. 180, 47 A.2d 445 (1946) (defendant and deceased were both teenagers who played a game of "Russian Poker"; although defendant believed that bullet was one chamber to right of firing chamber and would not fire, bullet was discharged; defendant's conviction for depraved-heart murder was affirmed) *with* State v. Cestari, No. A-1450-82T4 (N.J. Super. Ct. App. Div.), *cert. denied*, 97 N.J. 600, 483 A.2d 139 (1984). In *Cestari*, the defendant and the decedent, both retired police officers, posed for photographs at a party while each held a revolver containing one bullet to the other's head. The defendant, who believed the bullet in his gun was a chamber away from the gun's firing chamber on the spent side, cocked his gun. The bullet inside was discharged while he was still pointing the gun at the other officer's head. The jury acquitted the defendant of extreme indifference homicide and convicted him of reckless homicide. The appellate court affirmed the conviction.

[177] This language has been taken from MODEL PENAL CODE § 210.2(1)(b) (1980), which identifies the traditional concept as an indifference to the value of human life. *See supra* note 71 and accompanying text.

[178] State v. Curtis, 195 N.J. Super. 354, 479 A.2d 425 (App. Div.), *cert. denied*, 99 N.J. 212, 491 A.2d 708 (1984). The court ruled that the difference between reckless homicide and extreme indifference homicide consists of the possibility of risk in the former and probability of risk in the latter. The court found the difference to be objective and not to depend on the defendant's state of mind. *Id.* at 364, 469 A.2d at 430.

(b) the degree of the actor's causal responsibility for the death,

(c) the degree of the actor's awareness of that risk, and

(d) the social utility or disutility of the actor's conduct in creating or tolerating the risk.

(3) *Illustrations*. The following situations are examples of where a decisionmaker might, but is not legally bound, to find that a defendant's conduct:

(a) does manifest "an extreme indifference to the value of human life":

(i) the actor drives his car at high speed down a crowded sidewalk,[179]

(ii) the actor shoots another person in the head while playing "Russian roulette,"[180]

(iii) the actor is very drunk and is repeatedly warned not to drive but nonetheless drives, loses control of his car, and kills a pedestrian,[181]

(iv) after several traffic violations and a near-miss accident, the actor continues to drive drunk and hits another car while traveling 84 miles per hour in a 35 miles per hour zone,[182]

(v) the actor shoots at a moving passenger train for sport,[183]

(vi) the actor, a robber, initiates a gun battle in a crowded supermarket, and

(vii) the actor ties a person whom she has injured badly to a tree in an isolated forest in mid-winter;

(b) does not manifest "an extreme indifference to the value of human life":

(i) the actor hits a pedestrian while driving at high speed on a straight and apparently deserted highway in optimal weather conditions,

(ii) the actor shoots an innocent bystander while trying to disarm a mad-man who is on a shooting spree in a shopping mall,

---

[179] People v. Gomez, 65 N.Y.2d 9, 478 N.E.2d 759, 489 N.Y.S.2d 156 (1985).

[180] A less clear variation of this fact pattern appears in Commonwealth v. Malone, 354 Pa. 180, 47 A.2d 445 (1946).

[181] *See infra* note 182.

[182] Fact patterns (iii) and (iv) are a combination of the worst facts of Walsh v. State, 677 P.2d 912 (Alaska 1984), Pears v. State, 672 P.2d 903 (Alaska Ct. App. 1983), People v. Watson, 30 Cal.3d 290, 637 P.2d 279, 179 Cal. Rptr. 43 (1981), and State v. Weltz, 155 Minn. 143, 193 N.W. 42 (1923).

[183] Banks v. State, 85 Tex. Crim. 165, 211 S.W. 217 (1919).

(iii) the actor, who is slightly intoxicated, loses control of his car and strikes another car head-on,

(iv) the actor, shooting at a target range, hits a person who is standing beside the target, and

(v) the actor, driving on bald tires, has a blow-out and kills a pedestrian.

A legislature might well alter the content of this statute to reflect its own view of what factors are relevant and of where the critical point lies on the continuum of "indifference." This form of statute, however, promises more uniform outcomes that more accurately reflect legislative intent.

C. *Definition of Proximate Cause*

Offenses that include a result as an element of the offense, such as "death" in homicide offenses, require that the actor have "caused" the result.[184] A logically necessary—"but for"—cause relationship is insufficient to constitute adequate causation. Consider a defendant who shoots at someone but misses. The target then flees and is struck and killed, six blocks later, by a piano that was accidentally dropped while being hoisted to a fifth-floor apartment. The victim would not have been at the spot under the piano "but for" the defendant's earlier shot, but it seems clear that, for purposes of homicide liability, few persons would think that the defendant should be held liable for having "caused" the victim's death. Such cases of "but for" causation, where the result is too remote from the defendant's conduct, are held legally inadequate to constitute homicide through the doctrine of proximate cause.[185] The assessment of "proximate cause" requires an essentially intuitive judgment. No set formula that properly decides the question has yet been devised.

In exercising its normative judgment, the judge and jury at common law were given little or no guidance beyond being asked to determine whether the defendant's conduct was the

[184] 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 88(a) (1984).

[185] *See* W. LAFAVE & A. SCOTT, CRIMINAL LAW § 3.12 (2d ed. 1986).

"direct cause" or "proximate cause" of the victim's death.[186] The Model Penal Code formulates a slightly more useful standard that focuses upon an intuitive sense of justice: the decisionmaker is asked whether the actor caused a result that is "too remote or accidental in its occurrence to have a [just] bearing on the actor's liability or on the gravity of his offense."[187]

But even the Code standard provides the decisionmaker with little guidance. What factors make a result more or less "remote or accidental"? On the continuum of remoteness, how remote is "too remote"? As one might expect, decisionmakers do not respond consistently to these questions.

Proximate cause has been found where the defendant's accomplice carries a gun during a store robbery and the owner, frightened at the sight of the gun, runs from the store, vomits, falls unconscious, and dies fifteen minutes later.[188] Proximate cause has also been found where the defendant's accomplice strikes the jaw of the deceased who has been suffering from a severe heart condition.[189] By contrast, proximate cause has been held to be lacking where the defendant shakes the victim by his shoulder and briefly chokes him, and the victim, who suffers from high blood pressure, dies of a brain hemorrhage.[190] A finding of proximate cause has been allowed where the victim police officer crashes into a tree while chasing the defendant,[191] and where a policeman accidentally shoots a hostage while chasing the defendant.[192] On the other hand, proximate cause has been held to be lacking where an officer crashes into a bus while chasing the defendant.[193] The defendant's initial act has been

---

[186] *See, e.g.*, Regina v. Martin, 11 Cox's Crim. L. Cas. 136 (1868) (English case). In *Martin*, the defendant struck his victim, left him asleep by a fire in a country by-lane on a cold night, and later put him in a barn and covered him with straw. The victim's body was found in the straw some months afterwards. The issue of causation of the death of the victim was submitted to the jury as follows:

> Did the death result from beatings administered by the prisoner? Did the death result from the exposure on the night in question, and was that exposure the result of criminal neglect on the part of the prisoner? Did the death result from the prisoner leaving the boy in the barn under the straw ill, but not dead? [The judge] told the jury that if they found that death resulted from any one of the above causes, or any two of them, or all three together, they should find the prisoner guilty.

Id. at 137.

[187] MODEL PENAL CODE § 2.03(2)(b), (3)(b) (1985).

[188] Commonwealth v. Tatro, 4 Mass. App. Ct. 295, 346 N.E.2d 724 (1976).

[189] State v. Chavers, 294 So. 2d 489 (La. 1974), *cert. denied*, 419 U.S. 1111 (1975).

[190] Fine v. State, 193 Tenn. 422, 246 S.W.2d 70 (1952).

[191] Commonwealth v. Lang, 285 Pa. Super. 34, 426 A.2d 691 (1981).

[192] Jackson v. State, 286 Md. 430, 408 A.2d 711 (1979).

[193] People v. Scott, 29 Mich. App. 549, 185 N.W.2d 576 (1971).

held the proximate cause where, due to malpractice, a doctor fails to find and treat a stab wound,[194] and where questionable treatment and complications cause death after a head injury.[195] Nevertheless, where the victim of a stab wound to the stomach dies after an operation during which doctors also perform a necessary and medically correct but unrelated hernia operation, the proximate cause requirement has been held not to have been met.[196]

The differences in assessment of proximate cause extend beyond disagreements between trial juries and trial judges. In several of the above cases a lower court has taken one view of proximate cause while an appellate court has taken an opposite view on the same facts.[197]

Greater accuracy and consistency can be achieved by adopting statutory mechanisms to guide decisions without abandoning the inherently normative nature of the proximate cause judgment. A list of the types of factors that are relevant to the proximate cause judgment can direct the decisionmaker to pertinent considerations. Illustrations can also help the decisionmaker identify the critical point on the remoteness-accidental continuum. Such provisions might take the following form:

(1) *Causation Defined.* An actor's conduct is the cause of a result if the result would not have occurred but for the actor's conduct, and if the result is not too remote or accidental to have a just bearing on the actor's liability or on the gravity of his or her offense.[198]

(2) *Factors.* The following factors should be considered in evaluating whether the result is "too remote or accidental in its occurrence to have a just bearing on the actor's liability or on the gravity of his or her offense":

(a) the length of time and distance between the actor's conduct and the result,

(b) the likelihood that the actor's conduct will cause that result,

---

[194] State v. Hills, 124 Ariz. 491, 605 P.2d 893 (1980).

[195] State v. Hall, 129 Ariz. 589, 633 P.2d 398 (1981).

[196] People v. Stewart, 40 N.Y.2d 692, 358 N.E.2d 487, 389 N.Y.S.2d 804 (1976).

[197] In Commonwealth v. Lang, 285 Pa. Super. 34, 426 A.2d 693 (1981), the lower court found no proximate cause. In *Stewart*, 40 N.Y.2d at 692, 358 N.E.2d at 487, the lower court found proximate cause while the New York Court of Appeals reversed, finding no proximate cause.

[198] This language has been taken from MODEL PENAL CODE § 2.03(2)(b), (3)(b) (1985).

(c) the degree to which the manner of occurrence is unexpected,

(d) the independence of an intervening cause,

(e) the volitional nature of an intervening act, and

(f) an intervening actor's comparative causal responsibility.

(3) *Illustrations*. The following are examples of instances where a decisionmaker might, but is not legally bound, to find that a result:

(a) is "too remote or accidental in its occurrence to have a just bearing on an actor's liability or on the gravity of his offense":

(i) the actor comes to the victim's house in the evening to call for help in getting his car pulled from a nearby ditch; the actor's intoxicated state and gruff manner upset the victim who has suffered from hypertension for ten years; the victim's anxiety triggers a cerebral hemorrhage from which she dies three days later,[199]

(ii) the actor throws a rock at the victim's house; when being told of the incident by his son, the victim collapses and dies of fright or some related intense emotional state,[200]

(iii) the actor shoots the victim, inflicting a non-mortal wound; the victim dies from contracting scarlet fever from the attending physician,[201] and

(iv) the actor injures another in a fight and leaves the victim on the victim's front lawn; a sudden tornado snaps a tree which falls on the victim and causes her death.

(b) is not "too remote or accidental in its occurrence to have a just bearing on the actor's liability or on the gravity of his offense":

(i) the actor strikes the victim with his car, causing the victim to remain helpless in the middle of the road, where he is struck and killed by another car,[202]

(ii) the victim accidentally shoots herself when she grabs the gun away from the actor who is pretending to commit suicide in front of her,[203]

---

[199] Graves v. Commonwealth, 273 S.W.2d 380 (Ky. 1954).
[200] Commonwealth v. Colvin, 340 Pa. Super. 278, 489 A.2d 1378 (1985).
[201] Bush v. Commonwealth, 78 Ky. 268 (1880).
[202] People v. Parra, 35 Ill. App. 3d 240, 340 N.E.2d 636 (1975).
[203] State v. Shanahan, 404 A.2d 975 (Me. 1979), *cert. denied*, 444 U.S. 1079 (1980).

(iii) the actor sets fire to a car in a garage which spreads to an attached house and kills two children who could be but are not evacuated by their parents,[204] and

(iv) the actor injures another in a fight and leaves the victim, bleeding and unconscious, in the woods; the victim dies from exposure.

As before, the legislature will want to substitute their own factors and examples. This statute offers an example of a useful form.

Any attempt, such as this, to articulate the types of relevant factors will promote disagreement over fundamental issues. Should subjective issues such as the actor's culpable state of mind toward the result be included? Should the proximate cause determination rest on purely objective criteria?[205] While vague or nonexistent standards avoid confrontation on such issues, the underlying disagreements continue to exist in the less obvious form of inconsistent jury verdicts and unexplained reversals by judges. The value of a statutory list of relevant factors, therefore, lies both in the debate that it fosters on fundamental issues and policy matters and in the guidance that it provides.

D. *Distinction Between Preparation and Attempt*

During the course of conduct from the creation of a criminal plan to its completion, there exists a magic point at which a

---

[204] People v. Nichols, 3 Cal. 3d 150, 474 P.2d 673, 89 Cal. Rptr. 721 (1970), *cert. denied*, 402 U.S. 910 (1971).

[205] The Model Penal Code, for example, treats the remoteness issue as one concerning the defendant's culpability as to causation:

> (2) When purposely or knowingly causing a particular result is an element of an offense, the element is not established . . . unless:
>
> * * *
>
> (b) the actual result . . . is not too remote or accidental in its occurrence. . . .

MODEL PENAL CODE § 2.03(2) (1985); *see also* W. LAFAVE & A. SCOTT, HANDBOOK ON CRIMINAL LAW 252 (1972) (suggesting different standards of remoteness depending upon actor's culpability).

Other statutes, by contrast, treat the issue of remoteness as one that is distinct from the actor's culpability as to the result. *See, e.g.*, ALA. CODE § 13A-2-5 (1978); 18 PA. CONS. STAT. ANN. § 303 (Purdon 1973); TEX. PENAL CODE ANN. tit. 2, § 6.04 (Vernon 1974); P. ROBINSON, FINAL REPORT OF THE RHODE ISLAND CRIMINAL CODE REVISION PROJECT § 11A-2-2 (1985) ("Conduct is the cause of a result if: . . . (2) the result is not too remote. . . .") (on file at Harvard Journal on Legislation) [hereinafter RHODE ISLAND PROJECT].

non-criminal preparation to commit an offense becomes a criminal attempt. The Anglo-American law has recognized a series of different tests to determine this critical point. The common law's proximity tests[206]—how close did the actor come to causing harm?—have given way to the Model Penal Code's "substantial step" test[207]—how far did the actor progress with his or her plan? Most of the tests, however, suffer from a reliance upon vague terms. Common law phrases such as "physical proximity,"[208] "dangerous proximity,"[209] and "the point of probable desistance,"[210] offer no more guidance than the Code's "substantial step" standard. Attempts at more specific formulations, such as the "last step" test,[211] have largely been rejected as generating improper results.[212]

When determining if conduct has come within "dangerous proximity" of completion or whether it has progressed beyond a "substantial step" toward the offense, different decisionmakers, acting without further guidance, are likely to reach different conclusions on similar cases. Under the modern "substantial step" test, for example, most common fact patterns have been resolved in opposing ways by different courts. Hiding in the grass near a gas station with a gun and a black ladies' stocking was held to be a substantial step,[213] while hiding in an alley behind a photo studio wearing gloves and carrying a pry bar and flashlight was held not to be a substantial step.[214] Asking a store owner if he wanted to buy food stamps was held to be a substantial step,[215] but asking a child to commit sodomy was

---

[206] *See, e.g.*, Bell v. State, 118 Ga. App. 291, 163 S.E.2d 323 (1968); People v. Paluch, 78 Ill. App. 2d 356, 222 N.E.2d 508 (1966); People v. Gibson, 94 Cal. App. 2d 468, 210 P.2d 747 (1949).

[207] *See* MODEL PENAL CODE § 5.01(1)(c) (1985).

[208] *See, e.g., Bell*, 118 Ga. App. 291, 163 S.E.2d 323; State v. Boutin, 133 Vt. 531, 346 A.2d 531 (1975).

[209] *See, e.g., Paluch*, 78 Ill. App. 2d 356, 222 N.E.2d 508; People v. Ditchik, 288 N.Y. 95, 41 N.E.2d 905 (1942); People v. Werblow, 241 N.Y. 55, 148 N.E. 786 (1925).

[210] *See, e.g.*, United States v. Stephens, 12 F. 52 (C.C.D. Ore. 1882); People v. Gibson, 94 Cal. App. 2d 468, 210 P.2d 747 (1949); State v. Schwarzbach, 84 N.J.L. 268, 86 A. 423 (Ct. Err. & App. 1913).

[211] *See* Regina v. Eagleton, 6 Cox Crim. L. Cas. 559 (1855) (English case suggesting that attempt may be punished only if the defendant has completed the "last proximate act").

[212] *See* MODEL PENAL CODE § 5.01 comment 6, at 39 (Tent. Draft No. 10, 1960) (noting the universal rejection of the test).

[213] People v. Terrell, 99 Ill. 2d 427, 459 N.E.2d 1337 (1984).

[214] People v. Ray, 3 Ill. App. 3d 517, 278 N.E.2d 170 (1972), *rev'd on other grounds*, 54 Ill. 2d 377, 297 N.E.2d 168 (1973).

[215] People v. White, 84 Ill. App. 3d 1044, 406 N.E.2d 7 (1980).

held not to be a substantial step.[216] The court found a substantial step when the actor unlocked his barbershop and motioned the customer to the chair but had not yet taken any action to begin the haircut (he did not have a license to barber),[217] but it was held not to be a substantial step when actors climbed up a ladder to the roof of a building and were later found hiding with tools within a fenced-in area around the building but had not yet begun to break into the building.[218] A substantial step was found, despite the ambiguity of the conduct, when the actor put merchandise in a bag and walked past a cashier's counter to another department of the same store,[219] but a substantial step was not found, despite the less ambiguous intent, when an actor approached a stranger's house, knocked and rang the doorbell for a long time, shook the door knob, returned to the sidewalk and looked both ways, and walked up the driveway toward the back of the house.[220]

Guidance that will enhance a decisionmaker's accuracy and consistency in such decisions can be provided without undercutting the essentially normative nature of the judgment. A list of factors can tell the decisionmaker the types of facts he or she should consider. A series of illustrations, more importantly, can give the decisionmaker an indication of where, on the preparation-attempt continuum, the actor's conduct becomes subject to criminal sanctions. A provision offering this sort of guidance might be structured as follows:

(1) *Factors Relevant in Determining Whether Conduct is a "Substantial Step."* In evaluating whether an actor's conduct constitutes a "substantial step" toward commission of an offense, a decisionmaker should consider, among other things, the following factors:

---

[216] People v. Spencer, 66 Misc. 2d 658, 322 N.Y.S.2d 266 (N.Y. Crim. Ct. 1971). *See also* State v. Graham, 70 Or. App. 589, 689 P.2d 1315 (1984) (held no substantial step where defendant asked girl if he could "rape" her, then followed her briefly in a car after she refused and walked away); People v. Brown, 75 Ill. App. 3d 503, 394 N.E.2d 63 (1979) (defendant's attempt to talk cronies into robbing goods from enclosure while standing in front of it held not to be a substantial step).

[217] People v. Paluch, 78 Ill. App. 2d 356, 222 N.E.2d 508 (1966). *See also* Berry v. State, 278 Ark. 578, 647 S.W.2d 453 (1983) (held substantial step to rape where defendant forced victim at knifepoint to disrobe and walk into woods but did not touch her); State v. Latraverse, 443 A.2d 890 (R.I. 1982) (defendant seen in car across the street from officer who was to testify before grand jury; had gas, matches, rag, and note that said, "Hi, Sal, know [sic] it's my turn asshole"; held substantial step to intimidate witness).

[218] People v. Peters, 55 Ill. App. 3d 226, 371 N.E.2d 156 (1977).

[219] People v. Falgares, 28 Ill. App. 3d 72, 328 N.E.2d 210 (1975).

[220] R.L.T. v. State, 159 Ga. App. 828, 285 S.E.2d 259 (1981).

(a) the nature and number of the acts the actor already has performed toward the offense;

(b) the extent to which the completed acts indicate an intention to commit the offense;

(c) the nature and number of acts that remain to be performed for the actor to commit the offense; and

(d) the existence of lawful explanations for the actor's conduct.

(2) *Illustrations.* The following are examples of instances where a decisionmaker might, but is not legally bound, to find that an actor's conduct:

(a) does constitute a "substantial step" toward commission of an offense:

(i) lying in wait for, searching for, or following the contemplated victim of the offense,

(ii) enticing or seeking to entice the contemplated victim of the offense to go to the place contemplated for its commission,

(iii) reconnoitering the place contemplated for the commission of the offense,

(iv) unlawful entry of a structure, vehicle, or enclosure in which it is contemplated that the offense will be committed,

(v) possession of materials to be used in the commission of the offense, which are specially designed for such unlawful use or which can serve no lawful purpose of the actor under the circumstances,

(vi) possession, collection, or fabrication of materials to be employed in the commission of the offense, at or near the place contemplated for its commission, where such possession, collection, or fabrication serves no lawful purpose of the actor under the circumstances, and

(vii) soliciting an innocent agent to engage in conduct constituting an element of the offense;[221]

(b) does not, by itself, constitute a "substantial step" toward commission of an offense:

---

[221] This language has been taken from MODEL PENAL CODE § 5.01(2) (1985). *See also* Dennis, *The Law Commission Report on Attempt and Impossibility in Relation to Attempt Conspiracy and Incitement: (1) the Elements of Attempt*, 1980 CRIM. L. REV. 758, 772 (suggesting a modified form of the Model Penal Code list, which would exclude items (iii) and (vii), limit (v) and (vi) by substituting a requirement of using materials in place of the Model Penal Code "possession" requirement, and add a new item—"Preparing, stating or acting a falsehood for the purpose of an offence of fraud or deception").

(i) informing another person of one's intention to commit an offense,

(ii) possessing the materials that could be used for but are not specifically designed for commission of an offense,

(iii) loitering, and

(iv) looking for an opportunity to commit an offense.

Most of the preceding illustrations of what constitutes a "substantial step" were taken from the Model Penal Code provision on attempt.[222] While the Code contains such illustrations, they are of limited value because they provide only legal direction to the judge on what may not, as a matter of law, be held to be insufficient to be a substantial step.[223] The same examples, however, can illustrate for a jury the instances in which it might properly conclude that a substantial step has been taken. Such examples may be the best way to help the jury reach a verdict consistent with the legislature's notion of what constitutes a substantial step and with the result reached by other juries in similar cases.

E. *Balance of Interests in Justification Defenses*

One theory of justification is that an actor's conduct is justified if it prevents a greater harm or evil than that which it causes.[224] The theory underlies all justification defenses,[225] but is most clearly stated in the defense of lesser evils (also called "choice of evils" and, in a somewhat narrower form, "necessity"). The

---

[222] MODEL PENAL CODE § 5.01(2) (1985).

[223] *See supra* note 97 and accompanying text. The language of the Code directs: "Without negativing the sufficiency of other conduct, the following, if strongly corroborative of the actor's criminal purpose, shall not be held insufficient as a matter of law. . . ." MODEL PENAL CODE § 5.01(2) (1985). Thus, the Code merely prohibits the judge from directing a verdict of acquittal; it does not require that one or all of the examples of a substantial step stated in §§ 5.01(2)(a)–(g) be submitted to the jury for guidance. For example, the explanatory note states that "a list of kinds of conduct that corresponds with patterns found in common law cases is also provided, with the requirement that the *issue of guilt* be submitted to the jury if one or more of them occurs. . . ." *Id.* § 5.01(2) explanatory note at 297 (emphasis supplied). The Commentary further states that "if any of the stated circumstances has occurred, a judge has to instruct a jury that it may find a 'substantial step. . . .'" *Id.* § 5.01(2) comment 6(b), at 332. The language of the Code itself does not call for this instruction; it does not address the issue of guidance to the jury.

[224] 1 P. ROBINSON, CRIMINAL LAW DEFENSES § 24 (1984); 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 121 (1984).

[225] 2 P. ROBINSON, CRIMINAL LAW DEFENSES §§ 124 (lesser evils defense), 131 (defensive force defenses), 141 (public authority defenses) (1984).

Model Penal Code, for example, provides a justification defense to otherwise criminal conduct if "the harm or evil sought to be avoided by such conduct is greater than that sought to be prevented by the law defining the offense charged."[226]

The problem with such provisions is not that they use vague terms, such as "too trivial," "extreme indifference," "too remote," or "substantial step," but rather that they require a complex determination that may be beyond the unassisted capacity of the decisionmaker. Such provisions do not call for a single-continuum judgment but, instead, present a multi-dimensional problem. The decisionmaker must identify all significant interests injured by the actor's conduct, assign a value to each interest, and determine the total harm value. He or she must make the same calculation for those interests that the actor's conduct protects or benefits. The decisionmaker must then compare these two values and choose the greater. Thus, unlike the previous provisions, the lesser evils problem is not one of application of a vague, intuitive standard to specific facts. It is, rather, the problem of precisely analyzing and comparing values that cannot be reliably identified or quantified. By placing this unrealistic demand for precision on decisionmakers, the required balancing test creates a potential for generating unintentional and inconsistent results. What are the interests injured or furthered? How are they to be given values? How are these values to be compared? Without guidance, different decisionmakers are likely to answer these questions quite differently.

Decisionmakers regularly come to different conclusions. Disparity exists, for example, in deciding whether medical necessity can justify the unlawful possession of marijuana,[227] whether the need to draw public attention to the dangers of nuclear power plants can justify protestors' otherwise criminal trespass of such plants,[228] or whether the trespass of abortion clinics is justified

---

[226] MODEL PENAL CODE § 3.02(1)(a) (1985).

[227] *See, e.g.*, State v. Tate, 198 N.J. Super. 285, 486 A.2d 1281 (App. Div. 1984) (trial court denied State's motion to strike the defense; the Appellate Division affirmed), *rev'd.*, 102 N.J. 64, 505 A.2d 941 (1986) (with two dissents); State v. Diana, 24 Wash. App. 908, 913, 604 P.2d 1312, 1315 (1979) (remanding conviction so that trial court could hear evidence on medical necessity defense, a defense the defendant had not raised at trial, because "substantial medical and legal developments relating to the medical attributes of marijuana" had been "widely publicized" since the trial).

[228] *Compare* Commonwealth v. Berrigan, 325 Pa. Super. 242, 472 A.2d 1099 (1984) (allowing defendants to make offer of evidence on necessity defense), *rev'd*, 509 Pa. 118, 501 A.2d 226 (1985) (denying defense) *with* State v. Warshow, 138 Vt. 22, 410 A.2d 1000 (1980) (denying the defense).

in order to stop abortions.[229] Similarly, different conclusions have been reached in cases where prisoners have escaped prison to avoid forcible rape by other prisoners.[230]

As suggested previously, by providing guidance for the exercise of intuitive judgment, the accuracy and consistency of such judgments can be increased, while maintaining the normative nature of such judgments. First, the formula for the decision should be stated explicitly. Provisions that merely provide a defense if the actor's conduct is "justified" or "reasonable" are inadequate;[231] the Code formulation quoted above will provide greater accuracy and consistency.[232] The Code's provision can be improved upon, however, with the addition of a list of relevant factors that should be considered. Such a list cannot authoritatively assign an absolute value to every interest, but it can provide an assessment of the relative value of some common and important interests. For example, human life, even that of an unlawful aggressor, should be valued over property, while pride and ego should be valued less than bodily injury.

The complex nature of the justification judgment limits the value of illustrations. Because each case presents a relatively unique combination of facts, a single fact pattern and legal conclusion, without a description of the intermediate steps in the analysis, offers little value. Instead, a factfinder needs specific guidance in identifying and valuing competing interests.

A provision designed to guide the balancing process for the lesser evils defense might take the following form:

(1) *Lesser Evils Justification.* Conduct constituting an offense is justified if any legally-protected interest is unjustifiably threatened, and the actor engages in conduct when and to the extent necessary to protect that interest and thereby avoid a harm or evil greater than that caused by the actor's conduct.[233]

(2) *Factors Relevant in Determining the Harms or Evils at Stake.* In evaluating the interests protected by an actor's conduct and

---

[229] *See* Cleveland v. Municipality of Anchorage, 631 P.2d 1073 (Alaska 1981) (denying defense, and discussing cases that have recognized defense).

[230] *See* People v. Lovercamp, 43 Cal. App. 3d 823, 118 Cal. Rptr. 110 (1975) (permitting defense and discussing conflicting results in both California cases and cases of other jurisdictions).

[231] *See supra* note 108.

[232] *See supra* text accompanying note 224.

[233] This language is taken from Proposed Rhode Island Criminal Code, RHODE ISLAND PROJECT, *supra* note 205, § 11A-3-2.

the interests injured by such conduct, the decisionmaker should consider, among other things, the following factors:

(a) the harm caused to persons,

(b) the harm caused to property, including tangible and intangible property,

(c) the harm caused to social institutions, and

(d) the harm caused by disrupting the social order.

(3) *Special Rules Governing the Determination of Whether the Harms or Evils Avoided are Greater than those Caused.* In evaluating whether the harms or evils avoided are greater than those caused, the decisionmaker shall:

(a) determine the relative value of the interests at stake from the point of view of the community,[234]

(b) defer to judgments of relative value expressed in existing statutes,[235]

(c) value human life, even that of an unlawful aggressor, to be of greater value than property, and

(d) value physical injury to a person as more harmful than injury to a person's pride or ego.

Any attempt to articulate the relevant factors and their comparative values is likely to generate controversy. Is an actor justified in killing an innocent person to save two others?[236] twenty others? Is a killing justified if the innocent victim is a physical (though nonculpable) aggressor? if the victim is not an aggressor but nonetheless is the cause of the threat to others? if the victim could have avoided the threat?[237] Should an actor's mistaken belief that his conduct is objectively justified give him a justification defense?[238] Should an actor's mistaken belief that his conduct is not objectively justified bar his justification de-

---

[234] *See* MODEL PENAL CODE § 3.02 comment 1, at 5 (Tent. Draft No. 8, 1958) (balance cannot be committed to private judgment of the actor); 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 124(d)(1), at 50–52 (1984).

[235] A more narrow form of this rule appears in MODEL PENAL CODE § 3.02(1)(c) (1985).

[236] *See, e.g.*, Regina v. Dudley & Stephens, 14 Q.B.D. 273 (1884–1885) (defendants who were shipwrecked and starving at sea not justified in killing cabin boy for the fluids they needed to save their own lives).

[237] For a discussion of this last series of questions, see 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 124(g)(5) (1984).

[238] *Compare* MODEL PENAL CODE § 3.04(1) (1985) *with* 1 P. ROBINSON, CRIMINAL LAW DEFENSES § 27(e) (1984).

fense?[239] An adequate provision will answer these questions for the decisionmaker, although the drafting of such a provision will necessarily involve debate. A uniform application of the community's values requires that the community first decide what its values are.

F. *Definition of Culpable Risk-Taking*

Many offense definitions require recklessness or negligence, that is, culpable risk-taking or risk-creation. Modern codes distinguish recklessness and negligence as involving awareness and culpable unawareness, respectively,[240] but both levels of culpability require proof that the risk was of such a nature and degree that taking or creating it was blameworthy. The risk may concern the existence of a circumstance—such as whether property is "property of another" in theft, or whether a girl is over sixteen in statutory rape. The risk also may concern the likelihood that the actor's conduct will cause a prohibited result, such as the risk of causing death in homicide offenses or serious bodily harm in aggravated assault offenses.

Whatever the context, a single concept describes how much risk-taking the law considers culpable. Everyone takes and creates risks in daily life. Many risks—such as leaving a roller skate in the living room—are not dangerous enough to be criminally blameworthy. Others, such as driving, are justified by some benefit, while still others, such as tackling a person during a football game, are within the scope of risk-taking that society customarily accepts. The definition of culpable risk-taking defines those risks that are sufficiently dangerous, unjustified, and unacceptable as to be culpable.

Determining whether the taking of a risk is culpable requires several factors to be analyzed and compared. A decisionmaker can assess the substantiality of the risk, including both the probability and the potential harm, in a single evaluation in the

---

[239] *Compare* G. FLETCHER, RETHINKING CRIMINAL LAW 555–60 (1978) (supporting requirement of justificatory purpose and knowledge of justifying circumstances) *with* 2 P. ROBINSON, CRIMINAL LAW DEFENSES § 122 (1984) (requirement is inconsistent with objective nature of justifications); *see generally* 2 P. ROBINSON, CRIMINAL LAW DEFENSES (1984) (discussing conflicting authorities).

[240] *See, e.g.*, MODEL PENAL CODE § 2.02(2)(c), (d) (1985); *see generally* J. KAPLAN & R. WEISBERG, CRIMINAL LAW 187–91, 305–09 (1986) (discussing distinctions between recklessness and negligence); Robinson & Grall, *supra* note 28, at 695–96 (discussing Model Penal Code terms).

same way that the many factors affecting proximate cause may be combined to generate a single judgment of "remoteness." The issue of whether the risk is justified, however, requires a more complex judgment and analysis similar to the justification balancing required for the "lesser evils" defense. Indeed, the ultimate assessment of whether particular risk-taking is culpable in many ways is even more complex. It requires not only an assessment of whether the risk is substantial and unjustified, but also an assessment of how a reasonable person, with certain characteristics of the actor,[241] would have reacted to the risk. The decisionmaker compares the actor's reaction with this hypothetical reasonable person's reaction to reach an ultimate determination. In the language of the Code, taking a risk is culpable if the actor's reaction, consciously disregarding or failing to be aware of the risk, constitutes a "gross deviation"[242] from the standard of conduct that the reasonable person would observe in the actor's situation.

In light of the complexity of this determination, illustrative examples provide limited value; many combinations of any number of variables may produce a particular conclusion. Nonetheless, statutory revisions can enhance the accuracy and consistency of these decisions without abandoning the intuitive nature of the judgment. First, the statute should plainly state the formula for the decisionmaker to follow. A decisionmaker, especially a lay juror, cannot realistically be expected to devise unilaterally a proper method for assessing the culpability of risk-taking. Second, the statute should list the factors that may be relevant to each step of the analysis. The list helps to suggest the relevant range of factors to consider and to organize the decisionmaker's consideration of each factor at the point in the analysis where it becomes relevant.

A statute using such guidance devices might take the following form:

(1) *Culpable Risk Defined.* An actor takes a "culpable risk," as that term is used in the definition of "recklessly" in Section ____

[241] "The risk must be of such a nature and degree that, considering the nature and purpose of the [actor's] conduct and the circumstances known to him, [it] involves a gross deviation from the standard of [conduct/care] that a [law-abiding/reasonable] person would observe in the actor's situation." MODEL PENAL CODE § 2.02(2)(c), (d) (1985).

[242] *Id.*

and "negligently" in Section ____, when he or she takes or creates a substantial and unjustifiable risk that is of such a nature and degree that the actor's disregard of it or failure to perceive it involves a gross deviation from the standard of conduct that a reasonable person would observe in the actor's situation.[243]

(2) *Factors Relevant in Determining Whether a Risk is "Substantial and Unjustifiable" and in Determining the Reasonable Standard of Conduct for a Person in the Actor's Position.* In evaluating whether a risk is "substantial and unjustifiable" and in evaluating what would be required by "the standard of conduct that a reasonable person would observe in the actor's situation," the decisionmaker should consider, among other things, the following factors:

(a) the extent of the harm risked,

(b) the probability that the harm will occur,

(c) the value of any societal benefit that may be gained from taking the risk,

(d) the probability of the benefit from taking the risk,

(e) the extent of the harm that may be avoided by taking the risk, and

(f) the probability of the harm avoided by taking the risk.

(3) *Factors Relevant in Determining Whether an Actor's Risk-Taking is a Gross Deviation from the Reasonable Standard of Conduct.* In evaluating whether an actor's risk-taking involves "a gross deviation" from the standard of conduct that a reasonable person would observe in the actor's situation, the decisionmaker should consider, among other things, the following factors:

(a) the purpose of the actor's conduct,

(b) the circumstances known to the actor, and

(c) any characteristics of the actor or conditions of the situation that may excuse a miscalculation of the relative value of the factors listed in Subsection (2).

As with other provisions trying to guide discretion in assigning liability, an attempt to state the formula and to list and classify the relevant factors may generate a healthy debate over what the legislature deems appropriate. Should the evaluation of a risk depend on the costs and benefits that it presents to society? Would this mean that an actor's creation of a risk to a neighbor's property in order to save her own is non-culpable as long as she

[243] *See id.*

has more property at risk than her neighbor?[244] In judging whether risk-taking constitutes a "gross deviation from the standard of conduct that a reasonable person would observe in the actor's situation," what aspects of the "actor's situation" should be taken into account? emergency conditions? the actor's general inability to deal with emergency conditions as well as the average person can? the actor's inability to cope with the general stress of life (as documented by psychiatric testimony)? the actor's personality characteristic of being consistently careless and irresponsible (as documented by his family and acquaintances)?[245] Debate and resolution of such issues will make the process judges and juries use to make such intuitive judgments about culpable risk-taking more accurate and uniform.

G. *Degree of Impairment Required for Excuse*

The culpability principle forbids criminal liability for offenses committed by an offender who "could not have done otherwise," i.e., had neither the capacity nor a fair opportunity to act otherwise.[246] The circumstances that might relieve an actor of blameworthiness for an offense include duress, insanity, involuntary intoxication, or other disabilities recognized by law.[247] In each instance, the actor argues for an excuse from responsibility for the offense because, due to the special circumstances or characteristics, the actor could not reasonably have been expected to act otherwise. To impose liability in such cases would be unjust.

---

[244] Widely accepted principles of tort law provide a privilege to enter or remain on land of another if it is necessary to prevent serious harm to the actor, his land, or chattels. *See* RESTATEMENT (SECOND) OF TORTS §§ 197, 263 (1965). However, liability attaches for any harm done in the exercise of that privilege. *Id. See, e.g.*, Vincent v. Lake Erie Transportation Co., 109 Minn. 456, 124 N.W. 221 (1910) (defendant shipowner permitted to dock at plaintiff's pier due to severity of a storm but defendant must pay for damage to pier). *See also* Bohlen, *Incomplete Privilege to Inflict Intentional Invasions of Interests of Property and Personality*, 39 HARV. L. REV. 307 (1926); Fletcher, *Fairness and Utility in Tort Theory*, 85 HARV. L. REV. 537, 546 (1972).

[245] *See* MODEL PENAL CODE § 2.02 comment 3, at 126 (Tent. Draft No. 4, 1955) (discussing range of characteristics that should and should not be considered in evaluating whether an actor's conduct is a gross deviation from conduct of a reasonable person in his "situation"; noting that blindness would be an appropriate consideration but that temperament would not be).

[246] H.L.A. HART, PUNISHMENT AND RESPONSIBILITY 153 (1968).

[247] *See* 2 P. ROBINSON, CRIMINAL LAW DEFENSES §§ 161 (excuses generally), 171 (involuntary act defense), 173 (insanity), 174 (subnormality), 175 (immaturity), 176 (intoxication), 177 (duress) (1984).

With respect to duress, the Model Penal Code commentary concludes that

> law is ineffective in the deepest sense . . . if it imposes on the actor who has the misfortune to confront a dilemmatic choice, a standard that his judges are not prepared to affirm that they should and could comply with if their turn to face the problem should arise. Condemnation in such a case . . . is divorced from any moral base and is unjust.[248]

In support of the insanity defense, the Code commentary similarly argues,

> Whether or not elimination of the insanity defense is unconstitutional, . . . rejection of the premise that only those who are responsible should be treated as criminal would constitute abandonment of a deservedly fundamental value in the system of criminal justice and would represent a seriously retrogressive step in the development of the criminal law.[249]

With regard to involuntary intoxication, the Code commentary observes that "[such] intoxication . . . excuses . . . if the resulting incapacitation is as extreme as that which would establish irresponsibility had it resulted from mental disease."[250]

In duress, threats of force or the use of force that "a person of reasonable firmness in [the actor's] situation would have been unable to resist"[251] may excuse an offender. An analogous issue arises in insanity and involuntary intoxication. An actor may be excused if at the time of his conduct he "lacks substantial capacity either to appreciate the criminality [wrongfulness] of his conduct or to conform his conduct to the requirements of the law."[252] The language does not include the explicit reference to a reasonableness or justice standard as in duress but the commentary clarifies that the drafters intended such a standard.[253] Indeed, the first alternative formulation of the Tentative Draft was more explicit on this point. It permitted the excuse when the actor's "capacity either to appreciate . . . or to conform . . . is so substantially impaired that he cannot justly be held re-

---

[248] MODEL PENAL CODE § 2.09(1) comment 2, at 374–75 (1985).
[249] *Id.* § 4.01(1) comment 6, at 186 (footnotes omitted).
[250] *Id.*
[251] *Id.*
[252] *Id.* § 4.01(1) (insanity). *See id.* § 2.08(4) (involuntary intoxication).
[253] *See id.* § 4.01 comment 3, at 172 (quoted *supra* note 77).

sponsible."[254] This language tracks the general "justice" standard adopted in the Code's proximate cause provision.[255]

An important feature of these provisions is that they do not provide an absolute excuse. In some instances of duress, mental illness, or involuntary intoxication, the actor must live up to the requirements of law despite a disability. An actor must resist the coercion or impulse and comply with the law even where, because of the disability, he or she is unsure as to whether his conduct is criminal or wrong. If the actor fails to resist adequately a violation, this failure is blameworthy.[256] The challenge for the decisionmaker is to determine the actor's degree of impairment, the degree of impairment that will excuse, and whether the actor's impairment has reached this point of excuse.

Since impairment from mental illness or involuntary intoxication occurs internally, specific cases are difficult to compare and examine for inconsistent application. The amount of force or threat sufficient to provide an excuse because of duress, however, can be reviewed objectively.

A vague description of the standard for a duress defense, without further guidance, generates disparate results. A threat of harm to the actor or her family has been recognized as a basis for the defense of duress to the offenses of issuing bad checks,[257] escaping from prison,[258] possessing contraband drugs

---

[254] MODEL PENAL CODE § 4.01 comment on alternative formulation of § 4.01(1), at 27 (Tent. Draft No. 4, 1955). This formulation was rejected by the Institute. Those opposing the "justly" formulation argued that it was "unwise to present questions of justice to the jury." They preferred a provision that "in form, at least, confines the [jury's] inquiry to fact." MODEL PENAL CODE § 4.01 comment 4, at 159 (Tent. Draft No. 4, 1955); *But see* United States v. Brawner, 471 F.2d 969, 1033–34 (D.C. Cir. 1972) (Bazelon, J., concurring in part, dissenting in part); *accord* State v. Johnson, 121 R.I. 254, 267–69, 399 A.2d 469, 476–78 (1979) (preferring "justly responsible" standard).

[255] *See* MODEL PENAL CODE § 2.03(2)(b), (3)(b) (1985).

[256] *See, e.g.*, State v. Mriglot, 88 Wash. 2d 573, 564 P.2d 784 (1977), *aff'g* 15 Wash. App. 446, 550 P.2d 17 (1976). In *Mriglot*, the court affirmed the lower court's denial of an instruction that would have excused the defendant because he was "under the influence" or "affected by" liquor or drugs. *Id.* at 574, 564 P.2d at 785. Instead, with respect to general intent offenses, "the level of intoxication must be extreme enough to meet the jurisdiction's insanity test in order to *excuse* the crime." *Id.* at 578, 564 P.2d at 787 (emphasis in original). Because Mriglot was convicted of first-degree forgery, a specific intent offense, the level of intoxication, whether voluntary or involuntary, need only be enough to show that Mriglot lacked the specific intent to commit the offense. *Id.* at 576, 564 P.2d at 786. The lower court had failed to make the distinction between the burdens that a defendant bears when asserting involuntary intoxication as a defense to a specific intent offense and as a defense to a general intent offense. *Id.*

[257] Commonwealth v. Kyslinger, 506 Pa. 132, 484 A.2d 389 (1984) (lower court erred in denying duress instruction where defendant was told that he would be taken against his will to another state if immediate payment was not made).

[258] *See* People v. Trujillo, 41 Colo. App. 223, 586 P.2d 235 (1978). In addition to threats

in prison,[259] and filing a false medical report to perpetrate an insurance fraud.[260] Nonetheless, a threat of harm to the actor or his family has been held insufficient for a duress defense to the offenses of escape from prison,[261] sale of narcotics,[262] and refusal to answer questions at trial.[263] These decisions may reflect the absence of a shared understanding of the degree of coercion necessary to excuse a given criminal harm.

The accuracy and consistency of such a decisionmaker's judgment on the degree of impairment that will excuse can be improved by mechanisms that provide guidance but do not undercut the intuitive nature of the judgment. As with the analysis for justifications and the definition of culpable risk-taking, evaluating an excuse requires a complex judgment. In each instance, the decisionmaker must assess the actor's degree of impairment, the offense, the likelihood that a reasonable person would commit this offense if acting with this impairment, and the actor's blameworthiness in light of the hypothetical reaction of the reasonable person. The assessment involves multi-dimensional analysis. The reasonable person might commit some offenses under certain impairments but not others. The result in a single illustration might be the result of either the degree of impairment, the seriousness of the offense, the amount of resistance that should be expected given a particular degree of impairment and a particular offense, or a combination of these factors. Thus, illustrations may have limited usefulness here.

---

on the actual day of escape, the defendant was threatened two days prior and sexually assaulted five months prior to escape. The appellate court held that the trial court erred in not allowing defendant to offer proof about prior threats and assault where, as here, "they were definite, specific, and imminent." 41 Colo. App. at 225, 586 P.2d at 237. *See also* People v. Luther, 394 Mich. 619, 232 N.W.2d 184 (1975) (instruction that homosexual attacks by other prisoners is not a defense to prison escape was erroneous notwithstanding other instructions that duress was a valid defense).

[259] *See, e.g.*, People v. Maes, 41 Colo. App. 75, 76, 583 P.2d 942, 943 (1978) (defendant entitled to instruction on duress defense where he alleged that other inmates threatened him and his family if he did not "hold their stuff [contraband]").

[260] People v. Toscano, 74 N.J. 421, 378 A.2d 755 (1977). The New Jersey Supreme Court remanded to permit an instruction on duress defense where the defendant twice refused to fill out medical forms but consented when he and his wife were threatened. It rejected the common law requirement of "present, imminent, and pending" threat of harm in favor of the formulation found in the Model Penal Code. *See* MODEL PENAL CODE § 2.09 comment 8 (Tent. Draft No. 10, 1960).

[261] People v. Noble, 18 Mich. App. 300, 170 N.W.2d 916 (1969).

[262] Bailey v. People, 630 P.2d 1062 (Colo. 1981) (mere speculation that injury may occur is not sufficient for an excuse by duress).

[263] People v. Gumbs, 124 Misc. 2d 564, 478 N.Y.S.2d 513 (Sup. Ct. 1984) (witness based duress on defendant's reputation as a killer and on information that defendant was out to kill him (the witness); held, duress not established because no explicit threats were made and no force was used).

It would be useful and feasible, however, to identify the factors that affect the various elements of the decision, and to state a formula that describes the interrelation among these elements. A provision for a duress excuse might take the following form:

(1) *Duress Excuse.* An actor is excused for his or her conduct that constitutes an offense if as a result of being coerced by a threat that a person of reasonable firmness in the actor's situation would not have resisted, the actor is not sufficiently able to control his or her conduct so as to be justly held accountable for it.[264]

(2) *Factors Relevant in Determining Whether a Person of Reasonable Firmness in the Actor's Situation Would Have Resisted the Threat.* In determining whether "a person of reasonable firmness in the actor's situation" would have resisted the threat that coerced the actor, as required by Subsection (1), the decisionmaker should consider, among other things, the following factors:

(a) the extent of the threat, including:

(i) the imminence of the threat,

(ii) the nature of the harm threatened (such as, physical, economic or emotional),

(iii) the extent of the harm threatened,

(iv) the object of the threat (such as, the actor, a relative or a business associate),

(v) the unlawfulness of the threat, and

(vi) the source of the threat (such as, another person or natural forces) and

(b) the extent of harm caused by the actor, including:

(i) the seriousness of the harm caused in relation to the harm threatened, and

(ii) the availability of less harmful but equally effective means of avoiding the threat.[265]

(3) *Factors Relevant in Determining Whether the Actor Should be Justly Held Accountable for his Failure to Resist the Coercion.* In determining whether an actor should be "justly held accountable" for his or her failure to resist the coercion, as required by Subsection (1), the decisionmaker:

(a) should consider, among other things, the following factors:

---

[264] This language is taken from Proposed Rhode Island Criminal Code, RHODE ISLAND PROJECT, *supra* note 205, § 11A-4-10(1), (2) (with modifications).

[265] This language is taken from proposed Rhode Island Criminal Code, RHODE ISLAND PROJECT, *supra* note 205, § 11A-4-10(3).

(i) the circumstances known to the actor,

(ii) the extent of the coercion the actor experiences,

(iii) characteristics of the actor that may exacerbate the effect of the threat, and

(iv) the degree to which the actor's conduct comports with that of "a person of reasonable firmness in the actor's situation" as determined in Subsection (2);[266]

(b) should not consider whether the actor may have a general weakness or susceptibility to coercion.[267]

V. Summary and Conclusion

There seems little justification for our current practice of strict adherence to the legality principle in assigning liability and studied neglect of it in sentencing. "*Nullum crimen sine lege, nulla poena sine lege.*"[268] As the maxim reminds us, the virtues of the legality principle apply to all stages in the process of distributing criminal sanctions. At the same time, normative judgments, which necessarily depend on vague standards, must be made part of the process. Similarly, complex and unusual factors call for decisionmakers to exercise discretion.

Both the liability assignment and sentencing processes can learn much from each other in effectively accommodating the legality concerns with the need for discretion. In sentencing, little discretion is needed for foreseeable cases, involving common factors, that can be embodied in workable concrete rules. Thus, for such cases, the avoidance of unnecessary discretion found in liability assignment can and should be followed in sentencing. Further, simplifying the system by using such mechanisms as non-overlapping components of criminal conduct and general principles of adjustment can reduce the need for sentencing discretion to deal with complex issues. Where sentencing discretion *is* necessary—to account for unforeseeable situations, for issues concerning abstract or normative concepts that cannot be reduced to concrete rules, or for complex issues for

[266] Note that this concept resembles the notion of "gross deviation from the standard of care of a reasonable person in the actor's situation" that is at issue in culpable risk-taking. *See supra* text accompanying notes 240–45.

[267] Note that this concept resembles the notion of exclusion of carelessness as a characteristic relevant to determining the culpability of risk-taking. *See supra* note 245 and accompanying text.

[268] *See supra* note 19.

which workable rules are not yet practical—certain mechanisms can effectively guide the discretion. These mechanisms include, particularly, the adoption of a guiding statement of sentencing objectives and a framework to govern the interrelation of conflicting objectives.

The process of liability assignment can be improved by emulating some of the discretion provided in sentencing. Normative issues should be resolved through the application of broad standards that invite, rather than override, normative judgments. The discretion inherent in such broad standards can be effectively guided by providing decisionmakers with illustrations, descriptions of relevant factors, and, where needed, formulae that describe the interrelation among such factors. Such guided discretion may not adhere to the strict legality standards normally followed in substantive criminal law, but would do much to further the interests underlying the legality principle while maintaining the law's important incorporation of normative standards.

Changes in both sentencing systems and criminal code provisions for assigning liability can enhance both legality and the necessary exercise of discretion in the distribution of criminal sanctions. As a result, the distribution of sanctions can be made more uniform, rational, and just.

NOTE

WAS THE 1966 ADVISORY COMMITTEE RIGHT?: SUGGESTED REVISIONS OF RULE 23 TO ALLOW MORE FREQUENT USE OF CLASS ACTIONS IN MASS TORT LITIGATION

BRUCE H. NIELSON*

The number of lawsuits arising from mass torts has risen noticeably in the past several years. Many of these suits have been brought initially in the form of class actions under Rule 23 of the Federal Rules of Civil Procedure and its equivalent in many states. Unfortunately, most of the plaintiff classes in these actions have been denied certification by trial and appellate courts, often on the grounds that the class action is not an appropriate vehicle for the adjudication of mass tort claims.

In this Note, Mr. Nielson reviews the techniques used by innovative federal district judges in certifying Rule 23 class actions in mass tort cases. He then suggests how these techniques could be codified in a revised Rule 23, to allow its expanded use in the mass tort context.

In the Advisory Committee Note accompanying the 1966 amendments to Rule 23 of the Federal Rules of Civil Procedure, the Advisory Committee stated:

> A "mass accident" resulting in injuries to numerous persons is ordinarily not appropriate for a class action because of the likelihood that significant questions, not only of damages but of liability and defenses of liability, would be present, affecting the individuals in different ways. In these circumstances an action conducted nominally as a class action would degenerate in practice into multiple lawsuits separately tried.[1]

Despite this admonition, several imaginative and innovative federal district court judges have developed techniques to allow class actions in mass tort cases as the number of these cases

---

* Associate, Kirkpatrick & Lockhart, Washington, D.C. B.S. Brigham Young University, 1984; J.D., Harvard Law School, 1987. The author would like to thank Arthur R. Miller for his help in connection with this Note.

[1] FED. R. CIV. P. 23 Advisory Committee's note, subdivision (b)(3), 39 F.R.D. 69, 103 (1966) (citations omitted) [hereinafter Advisory Note].

has increased over the past ten years.[2] These techniques include certifying partial classes, diminishing the importance of the typicality requirement in the context of a 23(c)(4) action, recognizing the practical problems inherent in individual actions, and using subclasses to account for variations in state law.

This Note suggests that these techniques could be codified in a revised Rule 23 to allow for its use in the resolution of mass tort cases. The numerous political issues involved in a revision of the Federal Rules of Civil Procedure and the policy questions surrounding the use of class actions are important subjects that are beyond the scope of this Note.

I. TECHNIQUES ADOPTED BY FEDERAL JUDGES TO ALLOW CLASS ACTIONS IN MASS TORT LITIGATION

A. *Partial Classes and the Predominance and Superiority Requirements*

Many plaintiffs in mass tort litigation who seek class action certification do so under Rule 23(b)(3) of the Federal Rules of

---

[2] *See, e.g.*, In re "Agent Orange" Prod. Liab. Litig., 100 F.R.D. 718, 722 (E.D.N.Y. 1983) (order certifying class action), *mandamus denied*, 725 F.2d 858 (2d Cir.), *cert. denied sub nom.* Diamond Shamrock Chem. Co. v. Ryan, 465 U.S. 1067 (1984). The federal district court judges are not alone in this, as indicated by the Third Circuit's opinion in In re School Asbestos Litig. (School Asbestos II), 789 F.2d 996 (3d Cir.) (Rule 23(b)(3) certification upheld), *cert. denied sub nom.* Celotex Corp. v. School Dist. of Lancaster, 107 S. Ct. 182, *cert. denied sub nom.* Nat'l Gypsum Co. v. School Dist. of Lancaster, 107 S. Ct. 318 (1986). After quoting portions of the Advisory Committee's comments, Judge Weis wrote:

> Although that statement continues to be repeated in case law, there is growing acceptance of the notion that some mass accident situations may be good candidates for class action treatment. An airplane crash, for instance, would present the same liability questions for each passenger, although the damages would depend on individual circumstances. Determination of the liability issues in one suit may represent a substantial savings in time and resources. Even if the action thereafter "degenerates" into a series of individual damage suits, the result nevertheless works an improvement over the situation in which the same separate suits require adjudication on liability using the same evidence over and over again. *See Hernandez v. Motor Vessel Skyward*, 61 F.R.D. 558 (S.D. Fla. 1973).

School Asbestos II, 789 F.2d at 1008. *See also* MANUAL FOR COMPLEX LITIGATION SECOND § 33.24 (1986) ("careful consideration should be given to the propriety of class certification in mass tort litigation").

Civil Procedure.[3] To be maintainable, a class action brought under Rule 23(b)(3) must pose common questions of law or fact that "predominate over any questions affecting only individual members."[4] In addition, the class action must be "superior to other available methods for the fair and efficient adjudication of the controversy."[5] Many federal district court judges, agreeing with the Advisory Committee that the predominance and superiority requirements are not met because significant questions of damages, liability, and defenses would affect individual class members differently,[6] have refused to certify mass tort class actions.[7] Other federal court judges hearing mass tort cases have hurdled the predominance and superiority barriers by certifying partial classes (i.e., subclasses and classes with respect to particular issues).[8]

---

[3] The Federal Rules of Civil Procedure are codified in Title 28 of the United States Code. The text of Rule 23(b)(3) is as follows:

> (b) **Class actions maintainable.** An action may be maintained as a class action if the prerequisites of subdivision (a) are satisfied, and in addition:
>
> * * *
>
> (3) the court finds that the questions of law or fact common to the members of the class predominate over any questions affecting only individual members, and that a class action is superior to other available methods for the fair and efficient adjudication of the controversy. The matters pertinent to the findings include: (A) the interest of members of the class in individually controlling the prosecution or defense of separate actions; (B) the extent and nature of any litigation concerning the controversy already commenced by or against members of the class; (C) the desirability or undesirability of concentrating the litigation of claims in the particular forum; (D) the difficulties likely to be encountered in the management of a class action.

FED. R. CIV. P. 23(b)(3).

[4] FED. R. CIV. P. 23(b)(3).

[5] *Id.*

[6] *See* Advisory Note, *supra* note 1, at 103; *see also* text accompanying note 1.

[7] *See, e.g.*, In re Tetracycline Cases, 107 F.R.D. 719 (W.D. Mo. 1985); Caruso v. Celsius Insulation Resources, Inc., 101 F.R.D. 530 (M.D. Pa. 1984); Sanders v. Tailored Chem. Corp., 570 F. Supp. 1543 (E.D. Pa. 1983); Mertens v. Abbott Labs, 99 F.R.D. 38 (D.N.H. 1983); Feinstein v. Firestone Tire & Rubber Co., 535 F. Supp. 595 (S.D.N.Y. 1982); McElhaney v. Eli Lilly & Co., 93 F.R.D. 875 (D.S.D. 1982); Thompson v. Proctor & Gamble Co., 38 Fed. R. Serv. 2d (Callaghan) 519 (N.D. Cal. 1982); Ryan v. Eli Lilly & Co., 84 F.R.D. 230 (D.S.C. 1979); Mink v. Univ. of Chicago, 27 Fed. R. Serv. 2d (Callaghan) 739 (N.D. Ill. 1979); Causey v. Pan Am World Airways, 66 F.R.D. 392 (E.D. Va. 1975); Yandle v. PPG Indus. Inc., 65 F.R.D. 566 (E.D. Tex. 1974); Boring v. Medusa Portland Cement Co., 63 F.R.D. 78 (M.D. Pa.), *appeal dismissed mem.*, 505 F.2d 729 (3d Cir. 1974); Daye v. Commonwealth of Pa., 344 F. Supp. 1337 (E.D. Pa. 1972), *aff'd*, 483 F.2d 294 (3d Cir. 1973), *cert. denied sub nom.* Meyers v. Commonwealth of Pa., 416 U.S. 946 (1974).

[8] Jenkins v. Raymark Indus., Inc., 782 F.2d 468 (5th Cir.), *reh'g denied*, 785 F.2d 1034 (5th Cir. 1986) (affirming district court's order certifying class action); Walsh v. Ford Motor Co., 106 F.R.D. 378 (D.D.C. 1985) (order certifying class action); In re "Agent Orange" Prod. Liab. Litig., 100 F.R.D. 718 (E.D.N.Y. 1983) (order certifying

Rule 23 (c)(4)[9] contemplates partial class certification, and some federal district court judges have used Rule 23(c)(4) to "sever common issues in such a way as to insure predominance."[10] In his order certifying a class action in *Payton v. Abbott Labs*,[11] for example, Judge Skinner relied upon Rule 23(c)(4)(A) to certify a plaintiff class through which certain specific issues would be resolved.[12] Similarly, Judge Williams in certifying a class action in *In re Northern District of California "Dalkon Shield" IUD Products Liability Litigation*[13] recognized

---

class action); In re N. Dist. of Cal. "Dalkon Shield" IUD Prod. Liab. Litig., 521 F. Supp. 1188 (N.D. Cal. 1981) (order certifying class action), *vacated and remanded*, 693 F.2d 847 (9th Cir. 1982), *cert. denied sub nom.* A.H. Robins Co. v. Abed, 459 U.S. 1171 (1983); Pruitt v. Allied Chem. Corp., 85 F.R.D. 100 (1980) (order certifying class action); Payton v. Abbott Labs, 83 F.R.D. 382 (D. Mass. 1979) (order certifying class action), *vacated*, 100 F.R.D. 336 (D. Mass. 1983). The federal district court judges in these cases appear to have followed the recommendations of some commentators, *see infra* note 10.

[9] The text of Rule 23(c)(4) is as follows:

> **(c) Determination by Order Whether Class Action to be Maintained; Notice; Judgment; Actions Conducted Partially as Class Actions.**
>
> * * *
>
> (4) When appropriate (A) an action may be brought or maintained as a class action with respect to particular issues, or (B) a class may be divided into subclasses and each subclass treated as a class, and the provisions of this rule shall then be construed and applied accordingly.

FED. R. CIV. P. 23(c)(4).

[10] *Payton*, 83 F.R.D. at 391 (citing 3B MOORE'S FEDERAL PRACTICE ¶ 23.45[2] (1978)). *See also* In re School Asbestos Litig. (School Asbestos II), 789 F.2d 996, 1010 (3d Cir. 1986) (cases may exist "in which class resolution of one issue or a small group of issues will so advance the litigation that they may fairly be said to predominate"); *Pruitt*, 85 F.R.D. at 115 ("Rule 23(b)(3), in conjunction with subsections (c)(4) and (d), provides the Court flexibility to tailor class treatment of predominantly common interests where no better method is apparent."). According to Professors Wright, Miller, and Kane:

> Even though a court decides that the common questions do not predominate for purposes of Rule 23(b)(3), the action should not be dismissed if it can proceed on an individual basis. In addition, the court should always consider the possibility of determining particular issues on a representative basis as permitted by Rule 23(c)(4)(A) or reshaping the class pursuant to Rule 23(c)(4)(B) whenever that might prove efficient and economical. The effect may be to make the common issues in the recast class action "predominate" for purposes of Rule 23(b)(3).

7A C. WRIGHT, A. MILLER & M. KANE, FEDERAL PRACTICE AND PROCEDURE § 1778 (2d ed. 1986) (footnote omitted). *See also* S. Williams, *Mass Tort Class Actions: Going, Going, Gone?*, 98 F.R.D. 323, 334 (1983) (suggesting that the Ninth Circuit neither understood nor knew how to treat Judge Williams' order certifying an "issues-only" class action in the *Dalkon Shield* litigation, an order that the Ninth Circuit reversed). *But see* Van Harville v. Johns-Manville Prod. Corp., No. 78-642-H, slip op. (S.D. Ala. Sept. 27, 1979) (court "cannot simply carve off one portion of the case in which common questions do predominate and base class certification thereon").

[11] 83 F.R.D. at 383.

[12] *Id.* at 386–87.

[13] 521 F. Supp. 1188 (N.D. Cal. 1981).

"that certain issues such as causation and damages vary from individual to individual"[14] and therefore certified only "the common liability issues pursuant to Rule 23(c)(4)."[15] Additonally, in *In re "Agent Orange" Product Liability Litigation*,[16] Chief Judge Weinstein ordered that representative claimants be chosen for purposes of resolving causation issues with respect to a number of specific types of injuries.[17]

Following these decisions, subsequent plaintiffs have requested certification of subclasses. For example, the plaintiffs in *Walsh v. Ford Motor Co.*[18] sought certification of five separate subclasses for purposes of litigating various issues arising out of a single core of operative facts.[19] After considering the requirements of Rule 23 with respect to each of the proposed subclasses, Judge Green certified three classes under Rules 23(b)(3) and 23(c)(4) to "pursue claims . . . for alleged breach of written and implied warranty."[20]

The Third and Fifth Circuits have approved certification of partial classes to ensure predominance of common issues.[21] In *Jenkins v. Raymark Industries, Inc.*,[22] for example, the Fifth Circuit affirmed a Rule 23(b)(3) class of asbestos personal injury suits limited to selected common issues, saying that "[t]he purpose of class actions is to conserve 'the resources of both the courts and the parties by permitting an issue potentially affecting every [class member] to be litigated in an economical fashion.'"[23] In *In re School Asbestos Litigation* (*School Asbestos II*),[24] the Third Circuit, in affirming a Rule 23(b)(3) class for school asbestos compensatory damage claims, held that "[t]here may be cases in which class resolution of one issue or a small group of them will so advance the litigation that they may fairly be said to predominate."[25]

---

[14] *Id.* at 1194.

[15] *Id.*

[16] 100 F.R.D. 718 (E.D.N.Y. 1983).

[17] *Id.* at 723.

[18] 106 F.R.D. 378 (D.D.C. 1985).

[19] *Id.* at 383–84. The plaintiffs in *Walsh* proposed six different subclasses for certification, but two of the subclasses proposed were done so in the alternative. *Id.* at 384.

[20] *Id.* at 414. See *infra* notes 30–37 and accompanying text for a discussion of judicial interpretations of the interrelationship between Rules 23(a) and (b) and Rule 23(c)(4).

[21] In re School Asbestos Litg. (School Asbestos II), 789 F.2d 996, 1010 (3d. Cir. 1986); Jenkins v. Raymark Indus., Inc., 782 F.2d 468, 471 (5th Cir. 1986).

[22] 782 F.2d at 468.

[23] *Id.* at 471 (quoting General Tel. Co. of Southwest v. Falcon, 457 U.S. 147, 155 (1982) which quoted Califano v. Yamasaki, 442 U.S. 682, 701 (1979)).

[24] 789 F.2d at 996.

[25] *Id.* at 1010.

In determining the predominance of common issues, some judges have considered an efficiency notion as well as the efficacy of severing issues and creating partial classes.[26] Four years after Judge Skinner in *Payton* certified a plaintiff class under Rule 23(b)(3) to determine particular issues under Rule 23(c)(4)(A),[27] he decertified the class when answers from the Supreme Judicial Court of Massachusetts and other decisions resolved several of the issues that the class had been certified to resolve,[28] thus diminishing the efficiency of further class action. Other judges, including Chief Judge Weinstein in *Agent Orange*, see "litigation economies" as a relevant consideration in "deciding whether common questions predominate."[29]

In interpreting and construing the requirements of Rule 23(a) and (b) with the partial class provisions of Rule 23(c)(4), some federal district court judges have indicated that the importance of the typicality requirement of Rule 23(a) and the predominance requirement of Rule 23(b) is lessened when applying Rule 23(c)(4).[30] For example, Judge Roberts in *In re Tetracycline Cases* reasoned that the analytical process followed in determining predominance with respect to Rule 23(c)(4)(A) is "quite different than in the usual application of [the predominance

[26] *See* Payton v. Abbott Labs, 83 F.R.D. 382, 391–92 (D. Mass. 1979) (plaintiffs represented "that 'over 90% of trial time' in two individual DES cases was devoted to the question of 'whether and when defendants knew or should have known of the dangers of DES exposure'"). In adopting a quantitative measure in *Payton*, Judge Skinner went against the opinions of commentators who oppose the use of such a measure. *See* 3B MOORE'S FEDERAL PRACTICE ¶ 23.45[2] (1978); 7A C. WRIGHT, A. MILLER & M. KANE, *supra* note 10, § 1778. Judge Skinner's approach, together with Chief Judge Weinstein's analysis of litigation economies in *Agent Orange*, *see infra* text accompanying note 29, indicate the real prominence with which notions of efficiency occupy the mind of a judge faced with the prospect of hundreds or thousands of individual lawsuits arising from the same set of operative facts.

[27] *Payton*, 83 F.R.D. at 382, *vacated*, 100 F.R.D. 336 (D. Mass. 1983).

[28] *Payton*, 100 F.R.D. at 336 (motion to decertify granted).

[29] In re "Agent Orange" Prod. Liab. Litig., 100 F.R.D. 718, 722–23 (E.D.N.Y. 1983). *See also* General Tel. Co. of Southwest v. Falcon, 457 U.S. 147, 155 (1982) (purpose of class action is to preserve resources of courts and parties by permitting economical litigation of issues); Califano v. Yamasaki, 442 U.S. 682 (1979) (class action is to preserve resources of courts and parties); Jenkins v. Raymark Indus., Inc., 782 F.2d 468, 471 (5th Cir. 1986) (purpose of class action is to conserve resources of court and parties by permitting issues to be litigated in an economical fashion); Walsh v. Ford Motor Co., 106 F.R.D. 378, 393 (D.D.C. 1985), *vacated*, 807 F.2d 1000, 1007 (D.C. Cir. 1986) (when examining the predominance requirement it is important to remember that the inquiry is meant to determine whether the class action device can achieve economies); Advisory Note, *supra* note 1, at 103 (only where predominance exists can class action device achieve economies).

[30] In re Tetracycline Cases, 107 F.R.D. 719, 727 (W.D. Mo. 1985). *See also* Williams, *supra* note 10, at 334–35 (Ninth Circuit's misunderstanding of Rule 23(c)(4)(A) class concept in *Dalkon Shield* was evidenced by the Circuit Court's preoccupation with "individual issues predominating").

requirement] of Rule 23(b)."[31] According to Judge Roberts, this determination has the effect of lessening the importance of the predominance requirement, which is, for practical purposes, "subsumed to a considerable extent within [the] 'superiority' requirement [of Rule 23(b)(3)]."[32]

Judge Merhige in *Pruitt v. Allied Chem. Corp.* also suggested that Rule 23(c)(4) acts to alter the interpretation of the predominance requirement of Rule 23(b).[33] After finding that "questions affecting only varying subgroups of the proposed class predominate over the questions common to all the class members," Judge Merhige went on to "determine whether the proposed class may be subdivided under (c)(4) so that well-defined subclasses may present common questions concerning [defendant's] liability for the subclasses' damages."[34] The Judge then concluded "that six distinguishable subclasses within plaintiffs' proposed class merit separate class action treatment due to the predominance of common questions among their respective members."[35]

---

[31] *Tetracycline*, 107 F.R.D. at 727.

[32] *Id.*

[33] Pruitt v. Allied Chem. Corp., 85 F.R.D. 100, 111 (1980) (emphasis added):

> A finding that common questions of law or fact do not predominate over the questions affecting only individuals within the class does not, however, end the Court's inquiry. . . . *Rule 23(c)(4) requires the Court to consider employment of these restructuring measures when an apparently unmanageable class action could be converted to a manageable one. See Geraghty v. United States Parole Comm'n*, 579 F.2d 238, 253 (3rd Cir. 1978), *cert. granted*, 440 U.S. 945, 99 S.Ct. 1420, 59 L.Ed.2d 632 (1979) ("A forebearance to consider these options constituted a failure properly to exercise discretion.") This subsection provides the Court with the flexibility and the authority to "treat common things common and to distinguish the distinguishable." *Jenkins v. United Gas Corp.*, 400 F.2d 28, 35 (5th Cir. 1968).

The case Judge Merhige cites for support here, Geraghty v. United States Parole Commission, was vacated and remanded at 445 U.S. 388 (1980). Writing for the majority, Justice Blackmun there held with respect to a judge's obligation to consider creating subclasses:

> Petitioners assert that the Court of Appeals erred in requiring the District Court to consider the possibility of certifying subclasses *sua sponte*. Petitioners strenuously contend that placing the burden of identifying and constructing subclasses on the trial court creates unmanageable difficulties. We feel that the Court of Appeals' decision here does not impose undue burdens on the district courts. . . . On remand, however, it is not the District Court that is to bear the burden of constructing subclasses. That burden is upon the respondent and it is he who is required to submit proposals to the court. The court has no *sua sponte* obligation so to act. With this modification, the Court of Appeals' remand of the case for consideration of subclasses was a proper disposition.

Geraghty, 445 U.S. at 407–08 (emphasis in original). *See supra* note 10 and accompanying text. *See also infra* notes 101–15 and accompanying text.

[34] *Pruitt*, 85 F.R.D. at 111.

[35] *Id.* at 111–12. *See also supra* notes 10–25 and accompanying text.

As for the relationship between the prerequisite of commonality in Rule 23(a)(2) and the Rule 23(c)(4) subclasses created in *Pruitt*, Judge Merhige said:

> These common issues [of defendant's liability] are alone enough to satisfy the Rule 23(a)(2) commonality requirement, "as the rule does not require a complete coincidence of legal claims. It requires only that there be some questions of law and fact in common." *Crockett v. Virginia Folding Box Co.*, 61 F.R.D. 312, 317 (E.D. Va. 1974). *This is not to say that the common questions of law or fact predominate over the individual questions peculiar to various subgroups of the proposed class.*[36]

Judge Pratt in *Agent Orange* indicated that Rule 23(c)(4) could guarantee compliance with Rule 23(a)'s requirement of typicality. Claims that were not typical of the entire class, could be "efficiently managed either on a subclass basis or directly by way of separately determining the issues."[37]

Another matter pertinent to a judge's determination of predominance and superiority is "the interest of members of the class in individually controlling the prosecution or defense of separate actions."[38] Some judges have tempered the interest in individual control with the practical problems individual plaintiffs would face in independent litigation.[39] For example, Judge Williams in *Dalkon Shield* balanced "the importance of each [class] member having her own day in court . . . against the great cost and technical difficulties of discovery and independent litigation in general."[40] Judge Pratt in *Agent Orange* thought such problems inherent in individual actions to be "so great that it is doubtful if a single plaintiff represented by a single attorney pursuing an individual action could ever succeed."[41]

In actions where the alleged tort occurred in more than one state, variations in liability laws can further complicate an already messy Rule 23(b)(3) predominance and superiority analysis. In such cases, federal courts would have to apply varying

---

[36] *Id.* at 105 (emphasis added).

[37] In re "Agent Orange" Prod. Liab. Litig., 506 F. Supp. 762, 787 (E.D.N.Y.) (order certifying class action), *rev'd*, 635 F.2d 987 (1980).

[38] FED. R. CIV. P. 23(b)(3)(A).

[39] In re Northern Dist. of Cal. "Dalkon Shield" IUD Prod. Liab. Litig., 526 F. Supp. 887, 903 (N.D. Cal. 1981); *Agent Orange*, 506 F. Supp. at 790; In re Three Mile Island Litig., 87 F.R.D. 433, 440 (M.D. Pa. 1980).

[40] *Dalkon Shield*, 526 F. Supp. at 903.

[41] *Agent Orange*, 506 F. Supp. at 790.

standards to the parties' conduct,[42] thus raising doubts that common questions of law predominate and that class action is superior.[43] Some federal district court judges have suggested that this problem could be overcome and that multistate and nationwide classes could be certfied by (1) using Rule 23(c)(4) subclasses to account for variances in state law, and (2) requiring plaintiff's discovery and trial briefs to be directed to meet the most stringent test of liability.[44] Similarly, Chief Judge Weinstein in *Agent Orange* dismissed an argument highlighting the problem of varying state laws by saying further subclasses would be created, if necessary.[45]

---

[42] *See* Phillips Petroleum Co. v. Shutts, 105 S. Ct. 2965 (1985) (different state laws may govern the claims of class members residing in different states); Klaxon Co. v. Stentor Elec. Mfg. Co., 313 U.S. 487 (1941) (federal court with diversity jurisdiction must apply state choice of law rules); Erie R.R. Co. v. Tompkins, 304 U.S. 64 (1938) (except in matters governed by the Constitution or by acts of Congress, the law to be applied by federal courts is the law of the state). Choice of law problems may also exist, as between state and federal law in some cases. *See* In re "Agent Orange" Prod. Liab. Litig., 580 F. Supp. 690 (E.D.N.Y. 1984) (order that federal law or national consensus law, not state law, will apply in *Agent Orange*); Pruitt v. Allied Chem. Corp., 85 F.R.D. 100, 111 (1980) (choice between federal maritime tort law and non-maritime tort law).

[43] *See* FED. R. CIV. P. 23(b)(3). *See also* Walsh v. Ford Motor Co., 106 F.R.D. 378, 405 (D.D.C. 1985) (plaintiffs' personal injury class denied because claims for personal injury arise under state law and the court would be compelled to examine fifty-six potential state law variations); Brummett v. Skyline Corp., 38 Fed. R. Serv. 2d (Callaghan) 1443 (W.D. Ky. 1984) (Rule 23(b)(3) class certified for residents of Kentucky only because of problems in applying different states' laws in nationwide formaldehyde class action); Payton v. Abbott Labs, 83 F.R.D. 382, 386 n. 1 (D. Mass. 1979) (plaintiffs' class limited to residents of Massachusetts so as to avoid potential choice of law problem). One commentator has described some of the problems presented by state law variations in multistate class actions:

> [T]here will be a point at which the sheer magnitude of the task of construing the various laws will compel a court not to certify the multistate class or to reduce it to a more manageable number of states. Even short of that point, choice of law may pose major problems. The first is the danger of unwarranted intrusion into another state's legal affairs through a mistaken application of its laws. The court should thus consider its own familiarity with the other state's law, the degree to which the law is unclear or unsettled, and the extent to which it implicates important interests of the other state.

Note, *Multistate Plaintiff Class Actions: Jurisdiction and Certification*, 92 HARV. L. REV. 718, 742 (1979).

[44] In re Asbestos School Litig. (Asbestos School I), 104 F.R.D. 422, 434 (E.D. Pa. 1984) (order certifying class action), *aff'd in part and vacated in part, sub nom.* In re School Asbestos Litig. (School Asbestos II), 789 F.2d 996 (3d Cir.), *cert. denied*, 107 S. Ct. 182, *cert. denied*, 107 S. Ct. 318 (1986). Judge Kelly did not create subclasses to deal with variances in state laws, but he indicated a willingness to do so "as the need arises" and quoted Rule 23(c)(4) in an accompanying footnote. *Id.*

In the part of its opinion affirming Judge Kelly's certification of a Rule 23(b)(3) class, the Third Circuit had "some doubt" about Judge Kelly's handling of the problem of variance in state laws, but the Third Circuit also thought "the effort may nonetheless prove successful." *School Asbestos II*, 789 F.2d at 1010 (footnote omitted).

[45] In re Agent Orange Prod. Liab. Litig., 100 F.R.D. 718, 724 (E.D.N.Y. 1983). *See Agent Orange*, 580 F. Supp. at 690, for Chief Judge Weinstein's interesting and detailed analysis of choice of law problems. After noting the strong federal interest and national

B. *Relaxing the Commonality and Typicality Requirements*

Before being certified as such, a class action must meet the four prerequisites set forth in Rule 23(a). One of the prerequisites, Rule 23(a)(2), provides that class members may sue only if "there are questions of law or fact common to the class."[46] This commonality requirement has been found by some federal court judges to be satisfied by a showing of commonality as to liability.[47] For example, Judge Kelly in *In re Asbestos School Litigation* (*Asbestos School I*) outlined four theories of liability that could be established by common proof and that, though complex, did not vary from class member to class member;[48] these liability theories led the judge to find "[p]laintiffs have demonstrated common questions sufficient to satisfy Rule 23(a)(2)."[49]

The commonality requirement in mass tort cases has also been found to be satisfied by a showing of commonality as to the cause or impact of tortious activity.[50] For example, Judge Rambo in *In re Three Mile Island Litigation* found common issues in both the causes of a mass evacuation following a nuclear accident and the forseeablility of the impact of such evacuation on the local population and economy.[51]

In addition to meeting the commonality standard of Rule 23(a)(2), litigants can bring a class action only if, according to Rule 23(a)(3), "the claims or defenses of the representative parties are typical of the claims or defenses of the class."[52] In *Payton*, this typicality standard was found by Judge Skinner to

---

overtones of the case, and the fact that it was *sui generis*, Chief Judge Weinstein held that, under various conflicts of law rules, federal law or a national consensus law would apply with respect to defendants' liability and the award of punitive damages. The Chief Judge left for another time an analysis of "the nature of the national consensus or federal law." *Agent Orange*, 580 F. Supp. at 713.

[46] FED. R. CIV. P. 23(a)(2).

[47] *Asbestos School I*, 104 F.R.D. at 429 (order certifying class action); In re "Bendectin" Prod. Liab. Litig., 102 F.R.D. 239 (S.D. Ohio 1984) (order certifying class action), *vacated*, 749 F.2d 300 (6th Cir. 1984); Ouellette v. Int'l Paper Co., 86 F.R.D. 476, 479 (D. Vt. 1980) (order certifying class action), *aff'd*, 776 F.2d 55 (2d Cir. 1985), *cert. granted*, 475 U.S. 1081 (1986), *aff'd in part, rev'd in part*, 107 S. Ct. 805 (1987); Coburn v. 4-R Corp., 77 F.R.D. 43, 44 (E.D. Ky. 1977) (order certifying class action), *mandamus sought by some defendants denied sub nom.* Union Light, Heat and Power Co. v. United States Dist. Court, 588 F.2d 543 (6th Cir. 1978), *cert. denied*, 443 U.S. 913 (1979).

[48] *Asbestos School I*, 104 F.R.D. at 429.

[49] *Id.*

[50] *Id.*; *see also* In re Three Mile Island Litig., 87 F.R.D. 433, 439 (M.D. Pa. 1980).

[51] *Three Mile Island*, 87 F.R.D. at 439.

[52] FED. R. CIV. P. 23(a)(3).

be satisfied in part at least because of the plaintiffs' theory of enterprise liability.[53] However, Judge Skinner, recognizing the novelty of the plaintiffs' theory, indicated that the class action could also be maintained by limiting the class or creating defendant subclasses.[54]

The typicality requirement has also been satisfied in part at least because of a juridical relationship among defendants. For example, in his discussion of the typicality requirement in *Dalkon Shield*, Judge Williams recognized "that an important legal relationship justifying class treatment in this case is that each defendant is united in a chain of privity that has allowed them to introduce the Dalkon Shield into the stream of commerce."[55] Among the defendants in *Dalkon Shield* were "the inventors of the Dalkon Shield, the manufacturer and producer of the material and/or end product, and the distributor and/or supplier of the devices."[56]

Judge Williams' holding that typicality was satisfied in *Dalkon Shield* was rejected on appeal.[57] Ninth Circuit Judge Goodwin found the district court's holding to be inconsistent with *La Mar v. H & B Novelty & Loan Co.*, a Ninth Circuit opinion that held "typicality is lacking when the representative plaintiff's cause of action is against a defendant unrelated to the defendants against whom the cause of action of the members of the class lies."[58] However, *La Mar* also recognized two exceptions to this rule, one for cases "in which all injuries are the result of a conspiracy or concerted schemes between defendants," and another for "instances in which all defendants are juridically related in a manner that suggests a single resolution of the dispute would be expeditious."[59] In disagreeing with Judge Williams' holding that *Dalkon Shield* fell within these exceptions, the Ninth Circuit said that some of the plaintiffs did not allege that

---

[53] Payton v. Abbott Labs, 83 F.R.D. 382, 388 (D. Mass. 1979). *See also* Ouellette v. Int'l Paper Co., 86 F.R.D. 476, 480 (D. Vt. 1980) (Rule 23(a)(3)'s typicality requirement should be evaluated in terms of the plaintiffs' claims as to liability).

[54] *Payton*, 83 F.R.D. at 388. *See supra* text accompanying note 37 for a discussion of the interrelationship between the typicality requirement and the creation of partial classes.

[55] In re Northern Dist. of Cal. "Dalkon Shield" IUD Prod. Liab. Litig., 526 F. Supp. 887, 901 (N.D. Cal. 1981) *vacated and remanded*, 693 F.2d 847 (9th Cir. 1982), *cert. denied*, 459 U.S. 1171 (1983).

[56] *Id.*

[57] *Dalkon Shield*, 693 F.2d at 847, *cert. denied*, 459 U.S. at 1171.

[58] La Mar v. H & B Novelty & Loan Co., 489 F.2d 461, 465 (9th Cir. 1973).

[59] *Id.* at 466 (footnotes omitted).

their particular defendants had engaged in a conspiracy.[60] Although the Ninth Circuit did not reject the district court's theory of juridical relationships by name, the Court did refer to the *Dalkon Shield* defendants as including "some common defendants and some separate and uncommon defendants."[61] The Ninth Circuit's opinion implies that juridical relationships alone do not justify class action treatment, nor do they constitute conspiracies so as to create exceptions to the *La Mar* doctrine.[62]

Though the Ninth Circuit found the typicality requirement was not met in *Dalkon Shield*, the court cautioned, "[w]e do not decide or suggest that the typicality requirement of Rule 23(a)(3) may never be met when multiple plaintiffs sue different defendants."[63]

C. *Mandatory Class Actions*

Though many mass tort class actions are certified under Rule 23(b)(3), some federal court judges have used Rule 23(b)(1)[64] to certify mandatory classes in mass tort cases.[65] For example, Judge Wright in *In re Federal Skywalk Cases*[66] certified a Rule 23(b)(1)(A) class action on issues of liability for compensatory damages and liability for punitive damages out of apparent so-

---

[60] *Dalkon Shield*, 693 F.2d at 855.

[61] *Id.*

[62] *Id.*

[63] *Id.*

[64] The text of Rule 23(b)(1) is as follows:

> (b) **Class Actions Maintainable.** An action may be maintained as a class action if the prerequisites of subdivision (a) are met, and in addition:
>
> (1) the prosecution of separate actions by or against individual members of the class would create a risk of
>
> (A) inconsistent or varying adjudications with respect to individual members of the class which would establish incompatible standards of conduct for the party opposing the class, or
>
> (B) adjudications with respect to individual members of the class which would as a practical matter be dispositive of the interests of the other members not parties to the adjudications or substantially impair or impede their ability to protect their interests; . . . .

Fed. R. Civ. P. 23 (b)(1).

[65] In re Asbestos School Litig. (Asbestos School I), 104 F.R.D. 422, 438 (E.D. Pa. 1984); In re "Agent Orange" Prod. Liab. Litig., 100 F.R.D. 718, 728 (E.D.N.Y. 1983); In re Federal Skywalk Cases, 93 F.R.D. 415, 425 (W.D. Mo.), *vacated*, 680 F.2d 1175 (8th Cir.), *cert. denied*, 459 U.S. 988 (1982); *Dalkon Shield*, 526 F. Supp. at 900; Coburn v. 4-R Corp., 77 F.R.D. 43, 46 (E.D. Ky. 1977).

[66] 93 F.R.D. 415 (W.D. Mo. 1982).

licitude for the interests of the defendants.[67] Judge Wright also certified a Rule 23(b)(1)(B) class action on the issue of liability for and amount of punitive damages because of his concern that "[o]nly a single class-wide adjudication . . . can protect the interest of every victim in receiving his or her just share of any punitive damage award."[68]

Judge Wright's certification of Rule 23(b)(1) mandatory classes in *Federal Skywalk* was vacated by the Eighth Circuit, which held the class certification order was a violation of the federal Anti-Injunction Act.[69] Despite this, some commentators argued that "the mandatory class action device is the only procedure presently available to state and federal courts which allows for the equitable and efficient management of mass tort litigation."[70] Against sometimes contradictory authority, some federal district court judges have certified Rule 23(b)(1) mandatory class actions with similar justifications.[71]

---

[67] *Id.* at 423–24. Though no defendant asked the court to certify a Rule 23(b)(1)(A) class action "in order to protect it from successive punitive damage awards or from inconsistent adjudications on the liability issues," the court believed that such certification was "necessary as a ready and fair means of achieving unitary adjudication." *Id.*

[68] *Id.* at 425.

[69] *Federal Skywalk*, 680 F.2d at 1175.

[70] Wright & Colussi, *The Successful Use of the Class Action Device in the Management of the Skywalks Mass Tort Litigation*, 52 UMKC L. REV. 141 (1984). *See also* Note, *Mandatory Class Actions: Litigating Catastrophes Without Creating Litigation Catastrophes*, 19 J. MAR. L. REV. 91 (1985) (Rule 23's goals can be achieved only if mass torts are litigated as mandatory class actions); Note, *Class Certification in Mass Accident Cases Under Rule 23(b)(1)*, 96 HARV. L. REV. 1143 (1983) (proposes increased use of Rule 23(b)(1) as potentially the best procedural approach to mass accident litigation); Note, *Federal Mass Tort Class Actions: A Step Toward Equity and Efficiency*, 47 ALB. L. REV. 1180 (1983) (argues for increased use of Rule 23(b)(1)(B) class actions in mass tort cases).

[71] *See supra* notes 64–65 and accompanying text. In Coburn v. 4-R Corp., 77 F.R.D. 43 (E.D. Ky. 1977), Judge Rubin certified a mandatory class action under Rules 23(b)(1)(A) and 23(b)(1)(B) after discussing decisions in the Fifth and Ninth Circuits that Judge Rubin thought represented "contradictory nonbinding authority" such that the Judge had to "select that course of action that appears at the time to better assure equality of treatment for all litigants." However, the three cases Judge Rubin discussed, Hernandez v. Motor Vessel Skyward, 61 F.R.D. 558 (S.D. Fla. 1973), *aff'd mem.*, 507 F.2d 1278 (5th Cir. 1975), Green v. Occidental Petroleum Co., 541 F.2d 1335 (9th Cir. 1976), and McDonnell Douglas Corp. v. United States Dist. Court for the Cent. Dist. of Cal., 523 F.2d 1083 (9th Cir. 1975), *cert. denied*, 425 U.S. 911 (1976), all appear to stand for the proposition that a Rule 23(b)(1)(B) class action is permitted "only if separate actions '*inescapably* will alter the substance of the rights of others having similar claims.'" *Coburn*, 77 F.R.D. at 46 (emphasis added) (quoting La Mar v. H & B Novelty & Loan Co., 489 F.2d 461, 467 (9th Cir. 1973)).

Hernandez, in which a mandatory class was certified because defendant's amount of liability was statutorily limited, appears to be the kind of limited fund exception that proves the "inescapably alter" rule set forth in *McDonnell*. In any event, Judge Rubin apparently did not subscribe to the "inescapably alter" theory—he certified mandatory classes because he found a "risk of adjudications awarding damages to individual members of the class that might impose inconsistent standards of conduct upon defen-

The use of mandatory class actions in mass tort litigation may not survive the Supreme Court's decision in *Phillips Petroleum Co. v. Shutts*.[72] In *Shutts*, the Court held "that due process requires at a minimum that an absent plaintiff be provided with an opportunity to remove himself from the class by executing and returning an 'opt out' or 'request for exclusion' form to the court."[73] This holding seems to forbid mandatory class actions in mass tort cases, especially in light of the Court's limitation of its holding to "those class actions which seek to bind known plaintiffs concerning claims wholly or predominately for money judgments."[74] However, one commentator, while agreeing that "[t]he largest gray area for opt-out rights lies in the mass tort field," speculates that, in proposed mass tort mandatory class actions, the proscriptive effect of the opt-out right recognized in *Shutts* "likely depends on the particular circumstances involved (particularly how reliable the projected [Rule 23(b)(1)(B)] impairment probability is) and the balancing of the need to protect the class versus the individual interest in preserving the right to litigate one's own claim."[75]

---

dants or might so reduce the fund of money available as to substantially impair or impede the ability of other members of the class to protect their interests." *Coburn*, 77 F.R.D. at 46. *See also* FED. R. CIV. P. 23(b)(1).

[72] 105 S. Ct. 2965 (1985).

[73] *Id.* at 2975.

[74] *Id.* at 2975 n.3.

[75] H. NEWBERG, NEWBERG ON CLASS ACTIONS § 1.13 (1985). Professors Miller and Crump, while noting that "[t]here is no neat and logical means of resolving the question whether mandatory actions survive *Shutts*," propose a "four-factor approach that would narrow use of [mandatory class] actions, yet preserve them for situations whose resolutions are best accomplished through the mandatory class action device":

> The principal factors to be considered include the (1) efficiency and (2) equity concerns that are well illustrated in the pre-*Shutts* mandatory class decisions, as well as (3) the concern about distant forum abuse and (4) the interest in individualized control that seem to underlie *Shutts*. Our theory is that the propriety of mandatory class certification can best be determined by weighing the four enumerated policy factors in the context of each action.

Miller & Crump, *Jurisdiction and Choice of Law in Multistate Class Actions After Phillips Petroleum Co. v. Shutts*, 96 YALE L.J. 1, 52, 55–56, 80 (1986) (footnotes omitted). Applying their theory, Professors Miller and Crump conclude:

> For different reasons, cases such as *Skywalk* and *Agent Orange* also may present sound claims for mandatory certification. In *Skywalk*, the efficiency and equity factors are present and forum abuse is absent. In *Agent Orange*, the efficiency and equity factors arguably are strong enough to overcome the distant forum factor, and the interest in individual control is reduced by the size of the litigation and complexity of the common issues. On the other hand, the dispersed tort action, such as *Dalkon Shield*, presents the weakest claim for mandatory class certification, particularly if the constructive bankruptcy and punitive damage overkill theories are supported only by a generalized risk of inconsistency. Neither efficiency nor equity factors are persuasive in such

In addition to protecting and advancing the interests of defendants and plaintiffs, the certification of a mandatory class action may serve a strategic purpose by making a later Rule 23(b)(3) class certification more probable. For example, after the Eighth Circuit vacated the mandatory class action certification in *Federal Skywalk*, two Rule 23(b)(3) class actions were certified.[76] These two class actions, one in federal court the other in state court, were nearly identical to the mandatory class action that Judge Wright had originally certified.[77]

Similarly, the certification of a Rule 23(b)(1) mandatory class in *Asbestos School I* may have helped ensure the affirmance on appeal of a Rule 23(b)(3) certification. District Judge Kelly certified a Rule 23(b)(1) mandatory class on the issue of punitive damages and a Rule 23(b)(3) class for compensatory damages.[78] On appeal to the Third Circuit, the mandatory class was vacated.[79] However, the Third Circuit affirmed the Rule 23(b)(3) class without detailed analysis,[80] and recognized "growing acceptance of the notion that some mass accident situations may be good candidates for class action treatment" and that "the trend has been for courts to be more receptive to use of the class action in mass tort litigation."[81] Judge Kelly, by certifying a Rule 23(b)(1) class, may have encouraged the Circuit Court to recognize and accept the increasingly widespread use and potential value of the Rule 23(b)(3) class action in mass tort litigation.[82] The Third Circuit was fairly impressed by the district court's novel handling of the case, clearly reluctant to foreclose

---

a case. Furthermore, the factors of distant forum abuse and interest in individual control provide significant counterweights.

Id. at 56.

[76] In re Federal Skywalk Cases, 95 F.R.D. 483 (W.D. Mo. 1982) (order certifying class action); In re Skywalk Cases, No. CV 81-15244 MCF (Cir. Ct. Jackson Co., Mo., Jan. 6, 1983) (order certifying class action and approving settlement).

[77] Wright & Colussi, *supra* note 70, at 142.

[78] In re Asbestos School Litig. (Asbestos School I), 104 F.R.D. 422 (E.D. Pa. 1984).

[79] In re School Asbestos Litig. (School Asbestos II), 789 F.2d 996 (3d Cir. 1986). The Third Circuit's resolution of the case made consideration of possible due process problems under *Shutts* unnecessary; however, the court held "open the possibility of a 23(b)(1)(B) punitive damage class in more appropriate circumstances." *Id.* at 1008.

[80] *See id.* at 1008–11. The Circuit Court was also reluctant to vacate the Rule 23(b)(3) certification because of the "highly unusual nature of asbestos litigation," *id.* at 1011, and because of the innovativeness shown by the district court, *see infra* note 83.

[81] *School Asbestos II*, 789 F.2d at 1009 (citations omitted).

[82] *Cf.* Bentkowski v. Marfuerza Compania Maritima, S.A., 70 F.R.D. 401 (E.D. Pa. 1976) (action seeking damages for food poisoning, originally certified as mandatory class action, redesignated as Rule 23(b)(3) class action after Ninth Circuit opinions suggested mandatory class actions are generally not properly subject to certification in mass tort cases).

the use of the class action device altogether, and certainly less than satisfied with previous attempts at dealing with asbestos mass tort cases.[83]

The *Asbestos School I* case also provides an example of a technique developed by federal district court judges in recent years to manage mass tort class actions brought against defendants with relatively finite resources: certification of a mandatory punitive damages class under Rule 23(b)(1)(B).[84] In *Asbestos School I*, Judge Kelly certified a Rule 23(b)(1)(B) punitive damages class at least in part because of the "substantial possibility that early awards of punitive damages in individual cases will impair or impede the ability of future claimants to obtain punitive damages."[85] With minor variations, this "limited fund" theory has led other federal district court judges to certify Rule 23(b)(1)(B) classes on the issue of punitive damages for mass torts.[86]

In addition to the "limited fund" theory, some federal district court judges have suggested other rationales that favor certification of mandatory punitive damages classes. These include an "overkill" theory, described by Judge Kelly in *Asbestos School I* as the "possibility that subsequent plaintiffs will find themselves in litigation in which defendants 'will have been found [by the courts] to have been sufficiently punished' . . . such that further awards of punitive damages would only result in 'overkill.'"[87] Judge Williams in *Dalkon Shield* asserted that a "Rule 23(b)(1)(B) nationwide class action for punitive damages obviates many of the abuses inherent in multiple punitive dam-

---

[83] *School Asbestos II*, 789 F.2d at 1011. Circuit Judge Weis wrote, "The district court has demonstrated a willingness to attempt to cope with an unprecedented situation in a somewhat novel fashion, and we do not wish to foreclose an approach that might offer some possibility of improvement over the methods employed to date." *Id.*

[84] In re Asbestos School Litig. (Asbestos School I), 104 F.R.D. 422 (E.D. Pa. 1984). For other cases involving certification of Rule 23(b)(1)(B) punitive damages classes, see, e.g., In re Northern Dist. of Cal. "Dalkon Shield" Prod. Liab. Litig., 526 F. Supp. 887 (N.D. Cal. 1981); In re "Agent Orange" Litig., 100 F.R.D. 718 (1983); In re Federal Skywalk Cases, 93 F.R.D. 415 (W.D. Mo. 1982). For commentary on the use of Rule 23(b)(1) punitive damages classes, see Seltzer, *Punitive Damages in Mass Tort Litigation: Addressing the Problems of Fairness, Efficiency and Control*, 52 FORDHAM L. REV. 37 (1983); Note, *Class Actions for Punitive Damages*, 81 MICH. L. REV. 1787 (1983).

[85] *Asbestos School I*, 104 F.R.D. at 437.

[86] *See supra* note 84.

[87] *Asbestos School I*, 104 F.R.D. at 434 (quoting Williams, *supra* note 10, at 333) (citation omitted). *See also* Note, *The Punitive Damage Class Action: A Solution to the Problem of Multiple Punishment*, 1982 U. ILL. L. REV. 153 (class action device is best available method for a court to avoid the problems inherent in the present system of administering punitive damages).

age awards"[88] and went on to identify "unlimited multiple punishment for the same act"[89] as the kind of abuse against which a Rule 23(b)(1)(B) punitive damages class would protect.[90] In certifying a Rule 23(b)(1)(B) class in *Federal Skywalk*, Judge Wright considered the problem of uncertainty over whether state law would preclude multiple awards of punitive damages against a single defendant and saw in the mandatory class a way to protect against this uncertainty and to "resolve the potential conflicts of interest faced by counsel who represent more than one punitive damage claimant."[91]

The federal circuit courts frequently have disagreed with federal district court judges' certifications of Rule 23(b)(1)(B) punitive damages classes.[92] The mandatory punitive damages class in *Dalkon Shield* was decertified on appeal to the Ninth Circuit, which reasoned that precedent prohibited "Rule 23(b)(1)(B) certification of mass tort actions for compensatory or punitive damages unless the record establishes that separate punitive awards inescapably will affect later awards."[93] In *School Asbestos II*, the Third Circuit vacated the lower court's certification of a Rule 23(b)(1)(B) punitive damages class in part because the

---

[88] *Dalkon Shield*, 526 F. Supp. at 899.

[89] *Id.*

[90] *Id.*

[91] In re Federal Skywalk Cases, 93 F.R.D. 415, 425 (W.D. Mo. 1982).

[92] The punitive damages classes certified in each of *Dalkon Shield*, *Asbestos School I*, and *Federal Skywalk* were vacated on appeal, *see supra* notes 8, 44 & 65. Many federal district court judges have rejected Rule 23(b)(1)(B) punitive damages classes in mass tort litigation as well. For example, Judge Roberts writing in *Tetracycline* said:

> Whatever may be the future of class action treatment for punitive damages claims in mass disaster cases, I do not believe a punitive damages claim can adequately be treated as a class action matter in a products liability case postured as this one presently is, at least without grave—and perhaps insurmontable—difficulty.

In re Tetracycline Cases, 107 F.R.D. 719, 735 (W.D. Mo. 1985).

[93] *Dalkon Shield*, 693 F.2d at 851. The precedent upon which Judge Goodwin relied was McDonnell Douglas Corp. v. United States Dist. Court for the Cent. Dist. of Cal., 523 F.2d 1083 (9th Cir. 1975), which barred Rule 23(b)(1)(B) class certification on the issue of compensatory damages. In *Dalkon Shield*, Judge Goodwin extends *McDonnell Douglas*' ban to a 23(b)(1)(B) punitive damages class without any analysis as to why a bar on mandatory *compensatory* damages classes should apply to mandatory *punitive* damages classes. To distinguish between the two types of classes and argue for differential treatment, one might argue that the likelihood is greater that separate punitive damage awards will affect later awards because of a "sufficiently punished" or other similar constitutionally grounded theory. *See Federal Skywalk*, 680 F.2d at 1188–89 (Heaney, dissenting) ("[u]nlimited multiple punishment for the same act determined in a succession of individual lawsuits and bearing no relation to the defendants' culpability or the actual injuries suffered by victims, would violate the sense of 'fundamental fairness' that is essential to constitutional due process"); In re Asbestos School Litig. (Asbestos School I), 104 F.R.D. 422, 434 (E.D. Pa. 1984) (discussion and analysis of the "sufficiently punished" due process argument).

district court "made no factual findings at all as to the potential amount and scope of punitive damages."[94] The Third Circuit in *School Asbestos II* did grapple with the substantive due process arguments in favor of Rule 23(b)(1)(B) punitive damages classes, saying in dictum "these arguments might provide a threshold justification for the exercise of discretion in certifying a nationwide (mandatory) Rule 23(b)(1)(B) class for punitive damages."[95] Given this, one can speculate that the Third Circuit might have affirmed the Rule 23(b)(1)(B) class in *School Asbestos I* had other infirmities in the district court's certification not existed.[96] However, as discussed above, the constitutionality of any mandatory class action for damages, including punitive damages, is at least questionable under *Shutts*.[97]

Just as multistate mass tort class actions may involve varying state laws with respect to liability, the certification of a mandatory punitive damages class in multistate mass tort litigation can also create problems because of varying state punitive damages laws. Chief Judge Weinstein dealt with this problem in *Agent Orange* by reasoning that, because of the nature of punitive damages, "[t]he only jurisdictions concerned with punitive damages are those . . . with whom the defendants have contacts significant for choice of law purposes."[98] Chief Judge Weinstein's reasoning was held by Judge Green in *Walsh* to be "directly applicable to the present case."[99] Judge Green continued, "[b]ecause the purpose of punitive damages is to punish alleged wrongdoings, the jurisdiction with the greatest contacts or the location where the egregious activity took place may be the

---

[94] In re School Asbestos Litig. (School Asbestos II), 789 F.2d 996, 1005 (3d Cir. 1986). Another "critical flaw" in the *Asbestos School I* district court's analysis was that "[t]he class certified [did] not even include all property damage claimants." *Id.*

[95] *Id.*

[96] *See supra* note 94.

[97] *See supra* notes 72–75 and accompanying text.

[98] In re "Agent Orange" Prod. Liab. Litig., 580 F. Supp. 690, 705–06 (E.D.N.Y. 1984). Chief Judge Weinstein went on to list the contacts as: "defendants' place of incorporation, principal place of business, location of the plants that manufactured Agent Orange, and the site of any action taken in furtherance of what plaintiffs refer to as the conspiracy of silence." *Id.* at 706. The Chief Judge had also stated that

> the states of the veterans' domicile do not have an interest in whether or not punitive damages are imposed on the defendants. The legitimate interests of those states are limited to assuring that the plaintiffs are adequately compensated for their injuries and that the proceeds of any award are distributed to the appropriate beneficiaries.

Id. at 705 (citations omitted).

[99] Walsh v. Ford Motor Co., 106 F.R.D. 378, 408 (D.D.C. 1985).

jurisdiction which this Court should look to in determining the standard for punitive damages."[100]

D. *Sua Sponte Class Certification*

Some federal district court judges hearing individual claims that arose from one mass tort or accident have seized the initiative and certified class actions *sua sponte*.[101] However, this approach appears to be useful in mass tort litigation only to the extent the class member parties and their counsel agree with it.

For example, in deciding *Dalkon Shield*, Judge Williams relied upon the district court judge's "obligation to *sua sponte* determine whether an action shall proceed as a class action notwithstanding a motion from either party"[102] and upon the trial court's "equity power"[103] to "certify a class even when no individual plaintiff fashions his complaint seeking such relief."[104] After certifying the class in *Dalkon Shield*, Judge Williams also selected representative parties and designated lead counsel for the class.[105]

The Ninth Circuit pointed out on appeal in *Dalkon Shield* that "the right of litigants to choose their own counsel is a right not lightly to be brushed aside."[106] And, the Ninth Circuit continued, "[e]ven if the class were otherwise acceptable, it would have to be decertified if adequate lead counsel turned out to be unavailable."[107]

The Ninth Circuit recognized another potential problem with the district court's *sua sponte* class certification in *Dalkon Shield*: the district court's designation of representative parties might require a "substantial subdivision of representative subclasses"[108] which would "offer little advantage over the few

---

[100] *Id.*

[101] In re Northern Dist. of Cal. "Dalkon Shield" IUD Prod. Liab. Litig., 521 F. Supp. 1188, 1192 (N.D. Cal. 1981), *vacated and remanded*, 693 F.2d 847 (9th Cir. 1982), *cert. denied sub nom.* A.H. Robins Co. v. Abed, 459 U.S. 1171 (1983); In re Federal Skywalk Cases 93 F.R.D. 415, 419 (W.D. Mo. 1982). *See also supra* note 33 and accompanying text.

[102] *Dalkon Shield*, 521 F. Supp. at 1192 (footnote omitted).

[103] *Id.*

[104] *Id. See also* Senter v. General Motors Corp., 532 F.2d 511, 520 (6th Cir. 1976), *cert. denied*, 429 U.S. 870 (1976) (language of Rule 23(c)(1) is mandatory and district court has duty to certify the class action whether requested to do so or not); Stevenson v. Smith, 73 F.R.D. 79, 81 (D. Del. 1976) (in appropriate circumstances the court may be obligated to act on its own motion to determine the propriety of a class action).

[105] *Dalkon Shield*, 693 F.2d at 851.

[106] *Id.*

[107] *Id.*

[108] *Id.* at 854.

test trials that may produce more settlements than would a lengthy and complicated trial of consolidated cases."[109] Thus, although District Judge Williams presumably thought he had satisfied the 23(a)(3) typicality requirement by selecting "representative parties . . . covering the broadest possible gamut of types of injuries,"[110] the Ninth Circuit disagreed.[111]

If federal district court judges have the power to certify class actions on their own motions, they seemingly would also have the discretion to certify subclasses *sua sponte*.[112] Such discretion was referred to by Judge Williams in *Dalkon Shield*. After discussing the court's duty "to determine on its own motion whether or not a class action may be maintained,"[113] the Judge said, "the court may *sua sponte* certify subclasses during the pendency of an action without being bound by the plaintiff's complaint."[114] As Judge Wright pointed out in *Federal Skywalk*, *sua sponte* certification of subclasses can be useful to ensure commonality and to provide the "distinct treatment" required for resolution of the few issues that are not common to all class members.[115]

E. *Effect of Class Certification on Settlements*

In recent years, several federal judges have explicitly recognized the effect of class certification on the likelihood of prejudgment settlement in mass tort cases and have apparently allowed such recognition to influence their decisions to certify class actions.[116] For example, in certifying a class action in *Agent Orange*, Chief Judge Weinstein wrote:

---

[109] *Id.*

[110] *Dalkon Shield*, 526 F. Supp. at 919–20.

[111] *Dalkon Shield*, 693 F.2d at 855. The Ninth Circuit stated, "To prove liability under a breach of warranty theory, representative plaintiffs must exist for each type of warranty, assurance, or medical advice each plaintiff received. The difficulty of meeting the typicality requirement seems obvious." *Id.*

[112] *See supra* note 33.

[113] *Dalkon Shield*, 526 F. Supp. at 894 (footnote omitted).

[114] *Id. See also supra* note 33.

[115] In re Federal Skywalk Cases, 93 F.R.D. 415, 421 (W.D. Mo.), *vacated*, 680 F.2d 1175 (8th Cir.), *cert. denied*, 459 U.S. 988 (1982). In *Federal Skywalk*, Judge Wright opined that "[t]he issue of liability for punitive damages may require separate treatment of the wrongful death and survival claims." *Id.*

[116] In re "Bendectin" Prod. Liab. Litig., 102 F.R.D. 239, 240 (S.D. Ohio 1984); In re "Agent Orange" Prod. Liab. Litig., 100 F.R.D. 718, 723 (E.D.N.Y. 1983); In re School Asbestos Litig. (School Asbestos II), 789 F.2d 996, 1009 (3d Cir. 1986). *See also Federal Skywalk*, 680 F.2d at 1184–85 (Heaney, dissenting) ("[t]he crucial issue on this appeal is the effect of the district court's class certification order on the settlement process");

> Finally, the court may not ignore the real world of dispute resolution. As already noted, a classwide finding of causation may serve to resolve the claims of individual members, in a way that determinations in individual cases would not, by enhancing the possibility of settlement among the parties and with the federal government.[117]

In upholding the district court's certification of a Rule 23(b)(3) class in *Asbestos School I*, Circuit Judge Weis wrote that "the realities of litigation should not be overlooked in theoretical musings,"[118] that "preliminary maneuverings in litigation today are designed as much, if not more, for settlement purposes than for trial"[119] and that "[s]ettlements of class actions often result in savings for all concerned."[120]

Not only did Chief Judge Rubin consider the impact of class certification on pre-judgment settlement in *In re "Bendectin" Products Liability Litigation*,[121] but he also explicitly used class certification under Rule 23 to assist the parties "in reaching a prompt and equitable disposition of the entire [*Bendectin* litigation] problem."[122] After Chief Judge Rubin determined "that a class action under Rule 23 was not the appropriate way to consolidate these cases for trial,"[123] he recognized that "[t]he resolution of disputes does not necessarily require trial,"[124] and held that "[a] class certification would enable any proposed settlement to be presented to all class members and by them either accepted or rejected."[125]

Transgrud, *Joinder Alternatives in Mass Tort Litigation*, 70 CORNELL L. REV. 779, 782 (1985) (some mass tort cases could be handled under Rule 23(b)(3) for the purpose of structuring pretrial settlements).

[117] *Agent Orange*, 100 F.R.D. at 723. *Agent Orange* was settled on May 7, 1984, approximately five months after Chief Judge Weinstein's class action certification. *See* In re "Agent Orange" Prod. Liab. Litig., 597 F. Supp. 740 (E.D.N.Y 1984) (settlement approved as being fair); In re "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1396 (E.D.N.Y. 1984) (distribution of settlement). A settlement in *Federal Skywalk* was reached on January 10, 1983, a little more than two months after Judge Wright certified a Rule 23(b)(3) class action. In re Federal Skywalk Cases, 95 F.R.D. 483 (W.D. Mo. 1982). The settlement of the federal court litigation was approved by Judge Wright in In re Federal Skywalk Cases, 97 F.R.D. 380 (W.D. Mo. 1983). The *Skywalk* litigation in Missouri state court was settled on January 6, 1983, and the settlement was approved in In re Skywalk Cases, No. CV 81-15244 MCF (Cir. Ct. Jackson Co., Mo., Jan. 6, 1983).

[118] *School Asbestos II*, 789 F.2d at 1009.

[119] *Id.*

[120] *Id.*

[121] 102 F.R.D. 239 (S.D. Ohio 1984).

[122] *Id.* at 240.

[123] *Id.* at 240 n.4. Chief Judge Rubin did not disturb this earlier order when he certified the Rule 23(b)(1) class for settlement purposes. *Id.*

[124] *Id.* at 240 (footnote omitted).

[125] *Id.* at 240.

F. *Meeting the Jurisdictional Amount in Controversy Requirement*

In order to bring a mass tort class action, the members of a proposed plaintiffs' class must each meet the $10,000 jurisdictional amount in controversy requirement of 28 U.S.C. § 1332.[126] Many victims of non-catastrophic multistate mass torts may not have bona fide individual claims of $10,000 and may thus be denied federal court class action treatment on jurisdictional grounds. Under present law, such victims, rather than being able to take advantage of the economies of a single nationwide class action in federal court, would at best be relegated to their respective state courts for relief in state-wide class actions or individual lawsuits. Though the jurisdictional amount in controversy requirement would appear to bar federal mass tort class actions that involve non-catastrophic injuries or indeterminate damage claims, Judge Skinner refused to dismiss the proposed class action in *Payton* for want of the jurisdictional minimum amount in controversy because he could not find "to a legal certainty that the claim of any member of the plaintiff class is less than the jurisdictional amount."[127] Judge Skinner reasoned that he could not dismiss the case because the "[p]laintiffs'

---

[126] Zahn v. Int'l Paper Co., 414 U.S. 291 (1973); Snyder v. Harris, 394 U.S. 332 (1969). Title 28 § 1332 of the United States Code provides in relevant part: "The district courts shall have original jurisdiction of all civil actions where the matter in controversy exceeds the sum or value of $10,000, exclusive of interest and costs. . . ." 28 U.S.C. § 1332 (1982). This amount in controversy requirement for federal court jurisdiction appears to be in tension with one of the factors set forth in Rule 23(b)(3) to be considered in determining the superiority of a class action to other methods of adjudication. Rule 23(b)(3)(A) requires courts to consider "the interest of members of the class in individually controlling the prosecution or defense of separate actions." FED. R. CIV. P. 23(b)(3)(A). Presumably, proposed class members with smaller claims have a lesser interest in individually controlling litigation, In re Tetracycline Cases, 107 F.R.D. 719, 732 (W.D. Mo. 1985); Bentkowski v. Marfuerza Compania Maritima, S.A., 70 F.R.D. 401, 404 (E.D. Pa. 1976). *See also* Walsh v. Ford Motor Co., 106 F.R.D. 378, 405 (D.D.C. 1985) (whenever individual class members claim high value, their interest in controlling the prosecution of those claims is increased); Williams, *supra* note 10, at 329–30; Note, *Mass Accident Class Actions*, 60 CALIF. L. REV. 1615, 1634 (1972) (whenever individual class members claim high value, their interest in controlling the prosecution of those claims is increased and the propriety of a class action is decreased). However, such class members may also be less likely to satisfy the amount in controversy requirement.

[127] Payton v. Abbott Labs, 83 F.R.D. 382, 395 (D. Mass. 1979) (citation omitted). *See also* St. Paul Indem. Co. v. Cab Co., 303 U.S. 283 (1938) (claim can be dismissed for want of the jurisdictional minimum only if judge finds "to a legal certainty" that the claim is less than the jurisdictional amount).

claimed damages were unliquidated and subject to a jury's evaluation of many subjective factors."[128]

II. Suggested Revisions of Rule 23

A. *Predominance and Superiority*

Rule 23 might be revised so as to promote increased use of partial class certifications in mass tort litigation. Many federal district court judges have refused to certify Rule 23(b)(3) class actions in mass tort litigation because they find significant questions in the litigation that affect individual prospective class members in different ways.[129] Unfortunately for parties seeking class action treatment in mass tort cases, many federal judges do not look to Rule 23(c)(4) to determine whether subclasses should be created or whether issue-specific class actions should be certified.[130]

One reason for the judicial inattention to Rule 23(c)(4) partial classes may be the wording of the Rule, which provides for partial class actions "[w]hen appropriate."[131] Though such language properly gives judges discretion in managing cases, it also discourages all but the most innovative and imaginative judges from utilizing the class action device, even where a class action may be more efficient and equitable than individual actions.[132] Even when judges consider Rule 23(c)(4), they may not understand how that rule relates to and is to be construed with the predominance and superiority requirements of Rule 23(b).[133]

---

[128] *Payton*, 83 F.R.D. at 395.

[129] *See, e.g., Tetracycline*, 107 F.R.D. at 733; *see also* cases cited *supra* note 7.

[130] *See* Fed. R. Civ. P. 23(c)(4). Of the cases cited *supra* note 7, in each of which class certification was denied, only *Tetracycline* and *Caruso* contain findings with respect to possible Rule 23(c)(4) partial classes. *Tetracycline*, 107 F.R.D. at 726–27; Caruso v. Celcius Insulation Resources, Inc., 101 F.R.D. 530, 538 (M.D. Pa. 1984).

[131] Fed. R. Civ. P. 23(c)(4). The Rule begins, "(4) When appropriate (A) an action may be brought. . . ."

[132] *See supra* note 7 and accompanying text.

[133] Judge Roberts in *Tetracycline* is one of the few judges who addresses the interplay of the prerequisites of Rule 23(b) and Rule 23(c)(4)(A). He also notes the "absence of judicial commentary with respect to the effect of a Rule 23(c)(4)(A) class certification request upon the traditional analysis followed under Rule 23." *Tetracycline*, 107 F.R.D. at 727. *See* Pruitt v. Allied Chem. Corp., 85 F.R.D. 100, 113 (1980) ("[t]he presence of issues concerning individual plaintiffs' damages and the defenses thereto . . . requires the Court to employ the second device under Rule 23(c)(4) of bifurcating the class actions here certified between liability issues and damage issues"); *Caruso*, 101 F.R.D. at 538 (any common issue class certified under Rule 23(c)(4)(A) must still comply with other applicable subdivisions of Rule 23); *see also supra* notes 30–37 and accompanying text.

In order to promote more judicial consideration of the possibility of partial class certification to overcome problems of predominance and superiority in mass tort litigation, the following language could be added to Rules 23(b)(3) and 23(c)(4), respectively:[134]

> Rule 23(b)(3): . . . **In addition, the court shall consider whether, under subdivision (c)(4) of this rule, partial classes can be certified in which material issues would predominate that otherwise would not predominate with respect to the litigation as a whole. If the court finds that such partial classes can be created and that resolution of the issues for which the partial classes are created will materially advance a disposition of the litigation as a whole, the court shall find that the requirements of predominance and superiority in this subdivision have been met.**[135]

> Rule 23(c)(4): . . . **The court shall create partial classes under either or both parts of this subdivision when so doing would isolate material issues in the litigation the resolution of which would materially advance the disposition of the litigation as a whole, irrespective of whether the litigation as a whole proceeds under this rule.**[136]

In addition to elaborating on the relationship between Rule 23(c)(4) and the requirements of predominance and superiority, a revised Rule 23 could include a quantitative measure of predominance. Such a rule would require judges to consider litigation economies when determining whether the predominance requirement is met in proposed mass tort class actions.[137] Rule 23(b)(3) might be rewritten to contain this quantitative measure of predominance for mass tort litigation based on prior individual litigation experience. Such a measure could be a fifth matter that, under present Rule 23(b)(3), pertains to the court's findings with respect to predominance and superiority (new language in bold face):[138]

> Rule 23(b)(3): . . . The matters pertinent to the findings include: . . . **(E) in an action arising from a nucleus of operative**

---

[134] See Appendix A for the full text of a suggested revised Rule 23. See Appendix B for the full text of the Proposed Advisory Committee Notes to the suggested revisions of Rule 23 presented in this Note.

[135] *See* Appendix B, Proposed Advisory Committe Note to suggested revised Rule 23(b)(3).

[136] *See* Appendix B, Proposed Advisory Committee Note to suggested revised Rule 23(c)(4).

[137] *See supra* notes 26–29 and accompanying text.

[138] *See* FED. R. CIV. P. 23(b)(3).

> **facts with respect to which one or more individual cases have already been tried, the savings in time, money and other judicial and litigant resources that would result from a class action, such savings to be determined by reference to the time consumed and the money and other judicial and litigant resources expended in the prior individual actions.**[139]

A third revision would amend rule 23(b)(3)(A) such that judges would be required to weigh the practical problems inherent in individual actions when considering the individual interest in the litigation. Rule 23(b)(3)(A), which requires courts to consider the extent of individual interest in controlling litigation, often militates against a finding of predominance and superiority in mass tort cases involving severe or catastrophic injuries to claimants because judges tend to equate "interest" with size of potential recovery—as the potential recovery increases, so does a claimant's interest in individual litigation.[140] However, as the size of a claimant's potential recovery increases, often so does the complexity of and barriers to individual litigation.[141] In order to encourage judges to take the latter problems into account and to balance them against economic interests when considering claimants' interests in individual litigation, Rule 23(b)(3)(A) could be rewritten as follows (new language in bold face):

> Rule 23(b)(3): . . . The matters pertinent to the findings include: (A) the interest of members of the class in individually controlling the prosecution or defense of separate actions, **such interest to be determined by reference both to the potential size of individual recoveries and to the countervailing practical problems individual claimants would likely encounter in bringing individual actions because of the complexity, cost, and other problems inherent in the litigation;** . . . .[142]

Fourth, federal judges would perhaps be more likely to certify class actions in multistate mass tort cases involving varying state

---

[139] *See* Appendix B, Proposed Advisory Committee Note to suggested new Rule 23(b)(3)(E).

[140] *See* Sanders v. Tailored Chem. Corp., 570 F. Supp. 1543 (E.D. Pa. 1983) ("requested monetary relief for personal injury and property damage underscores . . . conclusion that class action certification is inappropriate"); Mertens v. Abbott Labs, Inc., 99 F.R.D. 38, 42 (D.N.H. 1983) (court noting that "this is not a situation involving a large number of small claims which would otherwise not be brought" as a factor militating against class certification), Causey v. Pan Am World Airways, Inc., 66 F.R.D. 392, 399 (high financial value placed on wrongful death actions is a factor leading to refusal to certify class). *See also supra* note 126.

[141] *See supra* text accompanying notes 39–41.

[142] *See* Appendix B, Proposed Advisory Committee Note to suggested revised Rule 23(b)(3)(A).

liability laws if Rule 23 permitted judges, in their discretion, to require parties seeking class action treatment to show violation of the most permissive (or compliance with the strictest) of the potentially applicable state standards of conduct. A non-mandatory choice of law rule such as this should help enable litigants in multistate mass tort cases to overcome problems with respect to the predominance of questions of law under Rule 23(b)(3).[143] To this end, Rule 23(b)(3) could be supplemented with the following provision:

> **In litigation in which the alleged wrongful or unlawful conduct occurred in more than one state and different standards of legal conduct apply in the various states in which such alleged conduct occurred, the court may nevertheless preserve the predominance of questions of law for purposes of this subdivision by requiring the parties seeking class certification to show, in the case of plaintiffs, violation of the most permissive relevant state standard of legal conduct, or, in the case of defendants, compliance with the strictest relevant state standard of legal conduct.**[144]

A provision such as this might be problematic under a confluence of the doctrines of *Erie Railroad v. Tompkins*[145] and *Klaxon Co. v. Stentor Electric Manufacturing Co.*,[146] and the requirement in the Rules Enabling Act[147] that the Federal Rules of Civil Procedure "not abridge, enlarge or modify any substantive right."[148] Critics of the proposed revision might argue that parties to a class action in federal court have a substantive right, under *Erie* and *Klaxon*, to have the court apply the choice of law rules of the forum state.[149] This provision, critics might argue, would (1) abridge the rights of parties who would other-

---

[143] *See supra* notes 43–45.

[144] *See* Appendix B, Proposed Advisory Committee Note to suggested revised Rule 23(b)(3).

[145] 304 U.S. 64 (1938).

[146] 313 U.S. 487 (1941).

[147] Rules Enabling Act, 28 U.S.C. § 2072 (1982).

[148] *Id.* As an alternative to the Supreme Court's introducing a discretionary choice of law rule in Rule 23, Congress could presumably enact a federal choice of law standard, to be used in multistate mass tort and other complex class action cases, that would, under the Supremacy Clause of the United States Constitution, U.S. CONST. art. VI, cl. 2, preempt any inconsistent state choice of law rules. *See* Miles v. Illinois Cent. R.R. Co., 315 U.S. 698, 703–04 (1942) (the laws of the United States are "the supreme law of the land, binding on every citizen and every court").

[149] *See Klaxon*, 313 U.S. at 487 (in diversity cases, the federal court must apply state choice of law rules); *Erie*, 304 U.S. at 64 (except in matters governed by the Constitution or by acts of Congress, the law to be applied by federal courts in any case is the law of the state).

wise have a different burden of proof under the standard of conduct applied according to the choice of law rules of the forum state, and (2) thereby also enlarge the rights of some opposing parties.

One could argue in response that *Erie* and *Klaxon* do not create a substantive right for litigants to have a certain state's laws determine the standards of conduct that a federal court with diversity jurisdiction will apply to the litigants or to the parties opposing them. But, even if litigants have such a substantive right, abridgement of that right in a class action may be preferable to the likely outcome if the class action does not go forward: no judicial process at all for many injured plaintiffs.[150] Furthermore, as the United States Supreme Court pointed out in *Hanna v. Plumer*,[151] the rule of *Erie* does not constitute "the appropriate test of the validity and therefore the applicability of a Federal Rule of Civil Procedure."[152] Nevertheless, *Erie*, *Klaxon*, and the Rules Enabling Act's proscription against alteration of substantive rights would probably inform a debate concerning the inclusion in Rule 23 of a discretionary choice of law rule.

B. *The Minimum Commonality Standard*

Though the commonality prerequisite of Rule 23(a)(2) literally requires "questions of law or fact common to the class,"[153] the class action in mass tort litigation can perhaps serve efficiently and effectively to determine conclusively a single important question of law or fact common to the class, such as the question of liability or of the cause or impact of tortious activity.[154] To that end, Rule 23(a)(2) could be revised so as to allow one material, common question of law or fact to satisfy the commonality requirement for class action certification (new language in bold face):

---

[150] As Chief Judge Weinstein wrote before being appointed to the bench, "In class actions necessity makes due process." Weinstein, *Revision of Procedure: Some Problems in Class Actions*, 9 BUFFALO L. REV. 433, 434 (1960).

[151] 380 U.S. 460 (1965).

[152] *Hanna*, 380 U.S. at 469–70.

[153] FED. R. CIV. P. 23(a)(2).

[154] *See supra* notes 46–51 and accompanying text.

> Rule 23(a): . . . (2) there are questions of law or fact common to the class, **or there is at least one material question of law or fact common to the class.**[155]

C. *Typicality*

As indicated in Part I, some federal judges have developed group liability theories to avoid the typicality problems claimants face in mass tort products liability class actions against multiple defendants.[156] However, the "justice" provided claimants under the enterprise liability, concerted scheme, and juridical relationship theories may come at a heavy cost of injustice to defendants who are swept in by such theories and forced to compensate claimants and pay punitive damages for injuries the defendants did not cause.

If defendants in mass tort products liability cases are considered better able than plaintiffs to bear the risk of possible injustice, the typicality requirement of Rule 23(a)(3) could be revised to codify one or more of the enterprise liability, concerted scheme, and juridical relationship theories. For example, all or part of the following Rule 23(a)(i) could be added to supplement Rule 23(a):

> Rule 23(a)[(.1),(.2),(.3)]: **In products liability litigation brought under this rule against more than one defendant, the court shall find the typicality standard of subpart (3) of subdivision (a) to be met with respect to as many defendants as the representative parties bringing the action can show [were engaged, whether dependently or independently, in some aspect of the design, development, production, testing or marketing of products of the type that is the subject of the litigation],**[157] **[and/or] [were engaged in a conspiracy or concerted scheme to design, develop, produce or market the product that is the subject of the litigation],**[158] **[and/or]**
> **[are currently or were formerly juridically related in such a manner that a single resolution of the dispute against them would be expeditious].**[159]

---

[155] *See* Appendix B, Proposed Advisory Committee Note to suggested revised Rule 23(a)(2).

[156] *See supra* notes 55–63 and accompanying text.

[157] *See* Appendix B, Proposed Advisory Committee Note to suggested new Rule 23(a)(.1) [first alternative].

[158] *See* Appendix B, Proposed Advisory Committee Note to suggested new Rule 23(a)(.2) [second alternative].

[159] *See* Appendix B, Proposed Advisory Committee Note to suggested new Rule 23(a)(.3) [third alternative].

The enterprise liability theory alone would not require a showing of a juridical relationship or conspiracy. Thus, the enterprise liability standard would probably be the least stringent standard for plaintiffs to meet. However, the enterprise liability and juridical relationship standards could be combined to impose a heavier burden on plaintiffs. Adding the conspiracy standard to the enterprise liability and juridical relationship standards would impose the heaviest burden on plaintiffs and would perhaps require a preliminary mini-trial just to determine whether the defendants had conspired or acted in concert. The juridical relationship standard alone would presumably be easier for plaintiffs to meet in actions against defendants whose business activities are vertically integrated and harder to meet when defendants are more horizontally positioned as competitors.[160] The least burdensome approach from a plaintiff's perspective would be to list the three standards as alternative ways to satisfy the Rule 23(a)(3) typicality prerequisite.

D. *Mandatory Classes*

Rule 23 might also be revised to require judges to certify mandatory class actions in mass tort litigation when there is a significant threat that plaintiffs who do not participate in a class action will be unable to recover damages in individual actions. Individual lawsuits arising from a catastrophic mass tort present potentially serious problems for defendants and plaintiffs. The former may be declared bankrupt by successive awards of compensatory and punitive damages; the latter may not get to court before the defendant is declared bankrupt.[161] As indicated in Part I, some federal judges in recent years have recognized and attempted to deal with this dilemma by certifying mandatory classes under Rule 23(b)(1) to protect both plaintiffs and defen-

---

[160] *See supra* text accompanying note 55.

[161] Faced with the prospect of more than 16,000 individual asbestos suits, the Johns-Manville Corporation turned to the bankruptcy laws effectively to force a mandatory class action against it in which a liability ceiling for injuries would be imposed and in which all present and future claims against the company would be extinguished. *See* In re Johns-Manville Corp., Nos. 82-B-11656 to 11676 (Bankr. S.D.N.Y. Aug. 26, 1982); H. NEWBERG, *supra* note 75, § 17.05. A.H. Robins Company followed a similar strategy with respect to the individual "Dalkon Shield" damage claims it faced after the Ninth Circuit decertified a Rule 23(b)(1)(B) class in In re N. Dist. of Cal. "Dalkon Shield" IUD Product Liab. Litig., 693 F.2d 847 (9th Cir. 1982). *See also* Diamond, *Robins, In Bankruptcy Filing, Cites Dalkon Shield Claims*, N.Y. Times, August 22, 1985, at A1, col. 2.

dants from being treated unjustly.[162] However, most federal courts have rejected the use of Rule 23(b)(1) mandatory classes in mass tort litigation, usually on the ground that the risks of individual adjudications contemplated by Rule 23(b)(1) (inconsistent standards of conduct for defendants or impairment of subsequent litigants' ability to protect their interests) are not sufficient to justify denying plaintiffs their due process right to prosecute individual actions.[163]

In light of this constitutional argument against mandatory class actions, and especially given the Supreme Court's opinion in *Shutts*,[164] revising Rule 23(b)(1) to require judges to certify mandatory class actions in mass tort litigation may be imprudent. However, the economic realities of mass tort litigation are such that plaintiffs' and perhaps defendants' due process rights may be jeapordized without the use of mandatory classes.[165] Thus, Rule 23(b)(1) might be revised to require judges to certify mandatory class actions in mass tort litigation and other cases only when the threat to plaintiffs' due process rights is outweighed by the threat that some plaintiffs will be unable to recover at all (new language in bold face):

> Rule 23(b): (1) An action may be maintained as a class action, **and an action must be maintained as a class action when the court deems class action treatment necessary to protect the rights of potential claimants,** if the prerequisites of subdivision (a) are satisfied and if the prosecution of separate actions by or against individual members of the class would create a risk of. . . .[166]

E. *Mandatory Punitive Damages Classes*

As indicated in Part I, the federal district court judges who have certified Rule 23(b)(1)(B) mandatory punitive damages

---

[162] *See supra* notes 66–68 and accompanying text.

[163] *E.g.*, In re School Asbestos Litig. (School Asbestos II), 789 F.2d 996, 1002–08 (3d Cir.), *cert. denied sub nom.* Celotex. Corp. v. School Dist. of Lancaster, 107 S. Ct. 182, *cert. denied sub nom.* Nat'l Gypsum Co. v. School Dist. of Lancaster, 107 S. Ct. 318 (1986); In re "Bendectin" Prod. Liab. Litig., 749 F.2d 300, 305–06 (6th Cir. 1984); McDonnell Douglas Corp. v. United States Dist. Court for the Cent. Dist. of Cal., 523 F.2d 1083, 1085–86 (9th Cir. 1975).

[164] Phillips Petroleum Co. v. Shutts, 472 U.S. 797 (1985). See *supra* notes 72–75 and accompanying text for a discussion of *Shutts*.

[165] *Cf. supra* notes 66–68 & 70–71 and accompanying text.

[166] *See* Appendix B, Proposed Advisory Committee Note to suggested revised Rule 23(b)(1).

classes in mass tort litigation have generally done so on the "limited fund" or "sufficiently punished" theories.[167] Although neither theory has been widely accepted by the federal judiciary,[168] federal judges would perhaps be more likely to use one or both of the theories as guidelines when performing the Rule 23(b)(1)(B) analysis in mass tort class actions and would probably, therefore, be more disposed to certify mandatory punitive damages classes, if the two standards were codified as part of Rule 23(b)(1)(B). To that end, Rule 23(b)(1)(B) might be revised to read as follows (new language in bold face):

> Rule 23(b)(1): . . . (B) adjudications with respect to individual members of the class which would as a practical matter be dispositive of the interests of the other members not parties to the adjudications or substantially impair or impede their ability to protect their interests; **in making its determinations under this subdivision (B), the court shall consider (i) evidence of the extent of the defendants' financial resources and (ii) the extent to which independent punitive damages awards to prior individual plaintiffs would cause a court in subsequent litigation to find the defendants had been sufficiently punished so that no further awards of punitive damages would be made.**[169]

Rule 23(b)(1)(B) could also require reference to a single punitive damages standard, given that multistate mass tort class actions may involve a variety of state punitive damages standards.[170] For example, for purposes of applying the "limited fund" and "sufficiently punished" theories, Rule 23(b)(1)(B) could provide that a federal court refer to the standard of the state with which the defendants have had the most significant contacts. Or, when multiple defendants have had their most significant contacts with different states, the court could be directed to apply the relevant state punitive damages standard most favorable to defendants. Thus, the following could be added to the language of Rule 23(b)(1)(B) set forth above:

> **For purposes of the determination in subdivision (ii) of this subpart (B) in multistate litigation, the court shall, as among the relevant applicable state punitive damages standards, refer to the punitive damages standard of the state with which the**

---

[167] *See supra* notes 84–91 and accompanying text.

[168] *See supra* notes 92–97 and accompanying text.

[169] *See* Appendix B, Proposed Advisory Committee Note to suggested revised Rule 23(b)(1)(B).

[170] *See supra* notes 98–100 and accompanying text.

> **defendants have the most significant contacts; in litigation against multiple defendants who do not all have the most significant contacts with the same state, the court shall refer to the relevant state punitive damages standard most favorable to defendants.**[171]

This revision, like the suggested revision to deal with variances in state liability laws presented above, could be attacked as an abridgement (or enlargement) of a substantive right in violation of the Rules Enabling Act.[172] The right abridged (or enlarged) in this case arguably would be a party's right, under *Erie* and *Klaxon*, to have a federal court with diversity jurisdiction apply the choice of law rules of the forum state for determining punitive damages.[173] Again, such an attack might be countered by an argument that this modification of some parties' rights would be minimal compared with the damage done to most plaintiffs' (and perhaps even some defendants') rights if no punitive damages class action can be brought.[174] Furthermore, this provision does not direct courts to apply certain punitive damage standards in determining damage awards. Rather, the courts should refer to the indicated punitive damage standard merely to determine whether the interests of absent parties are at risk of being impaired by individual adjudications via the "limited fund" or "sufficiently punished" theories.[175]

F. Sua Sponte *Class Certification*

Rule 23 could also be revised to require judges to certify a class action *sua sponte* when a large number of cases are pending. The onslaught of mass tort litigation in the past several years has further worsened the situation of already clogged

---

[171] *See* Appendix B, Proposed Advisory Committee Note to suggested revised Rule 23(b)(1)(B).

[172] *See* 28 U.S.C. § 2072 (1982) (rules of civil procedure "shall not abridge, enlarge or modify any substantive right"). *See also supra* notes 144–49 and accompanying text.

[173] *See supra* text accompanying notes 145–49. As with the choice of law problem with respect to varying state liability laws, the *Erie* and *Klaxon* problem here would presumably be avoided if Congress were to enact a choice of law rule for use in determining punitive damages in complex litigation, such as mass tort cases. *See supra* note 149.

[174] *See supra* note 150 and accompanying text for a similar argument with respect to a balancing of rights in the context of varying state liability standards.

[175] In other words, these proposed choices of punitive damage standards are strictly for procedural purposes—they provide a "quick and dirty" test of the likelihood that the outcome described in Rule 23(b)(1)(B) will occur; they are not rules necessarily to be applied in determining the actual punitive damages to be awarded.

federal and state dockets, with the result that many claimants wait several years before recovering damages from tort-feasors.[176] As described in Part I, some federal judges have dealt with dockets burgeoning with cases arising from the same nucleus of operative facts by certifying class actions *sua sponte*.[177] Case law describes both a judge's obligation to determine on his or her own motion whether a class action should be certified, and a judge's power to certify a class *sua sponte*,[178] but Rule 23 is silent on the matter. In fact, the Rule's present language implicitly presumes the initiation of class actions by parties only: "One or more members of a class may sue or be sued as representative parties. . . ."[179]

The revisions of Rule 23(b) suggested above would perhaps reverse the present presumption against *sua sponte* class certification by implying a judge's obligation to certify class actions *sua sponte*. However, more explicit language could require judges to certify class actions *sua sponte* when faced with tens, hundreds, or thousands of cases that arise out of the same nucleus of operative facts.[180] Thus, Rule 23(c) might be supplemented with a Rule 23(c)(i) to encourage judges, at their discretion, to certify class actions *sua sponte*:

> **Rule 23(c)(.1): The court may, on its own motion and regardless whether any party has requested such treatment, certify a class action under this rule if (A) the court has before it a large number of individual cases that arose out of the same nucleus of operative facts, (B) the court finds plaintiffs who will willingly act as representative parties, and (C) the class action meets the prerequisites of subdivision (a) and the requirements of subdivision (b) of this rule. A certification under**

[176] With respect to asbestos victims, a recent newspaper article reported on the activities of "[t]wo West Coast lawyers" who last year prepared "1,000 new asbestos lawsuits with thousands more being readied." Another lawyer has, since September 1986, "filed asbestos claims on behalf of more than 1,500 seamen in federal district court in Cleveland." Another group of attorneys is "looking at the potential for more than 30,000" asbestos related cases on behalf of tire plant workers. Meier and Richards, *Lawyers Lead Hunt for New Groups of Asbestos Victims*, Wall St. J., Feb. 18, 1987, at 1, col. 6. Moreover, according to the Third Circuit, in addition to the "more than 30,000 personal injury claims [already] filed against asbestos manufacturers and producers," an estimated 180,000 asbestos related claims "will be on court dockets by the year 2010." In re School Asbestos Litig. (School Asbestos II), 789 F.2d 996, 1000 (3d Cir. 1986).

[177] *See supra* note 101 and accompanying text.

[178] *See supra* notes 33 & 104.

[179] FED. R. CIV. P. 23(a).

[180] *See* In re N. Dist. of Cal. "Dalkon Shield" IUD Product Liab. Litig., 521 F. Supp. 1188 (N.D. Cal. 1981) (court certifies class action *sua sponte* to adjudicate claims by thousands of women that they were injured by a defective product). *See also supra* note 176 and accompanying text.

> **this subdivision may be conditional, and may be altered or amended before the decision on the merits.**[181]

A revised Rule 23(c) might also authorize *sua sponte* certification of subclasses. As a part of a larger Rule that implies only litigant-initiated class action treatment, Rule 23(c)(4) as written also does not explicitly authorize *sua sponte* certification of partial classes or subclasses.[182] As revised above,[183] Rule 23(c)(4) would encourage the creation of subclasses in mass tort litigation to preserve predominance and superiority for Rule 23(b)(3) purposes when so doing would further the litigation as a whole.[184] However, the revised language does not make explicit a judge's obligation and authority to create partial classes and subclasses *sua sponte*. In order to make this more clear, the following sentence could be added to the revised Rule 23(c)(4):

> **The court may, on its own motion, create partial classes under either or both of the parts of this subdivision if the requirements of this subdivision are met and if the court finds plaintiffs who will willingly act as representative parties for the partial classes created.**[185]

G. *Effect of Class Certification on Settlement*

Even more than in conventional litigation, pretrial or pre-judgment settlement in mass tort cases can help conserve the resources of the parties and the courts.[186] Thus, the prospect of pretrial or pre-judgment settlement should and does play a sig-

---

[181] *See* Appendix B, Proposed Advisory Committee Note to suggested new Rule 23(c)(i).

[182] Rule 23(c)(4)(A) provides that "an action may be brought or maintained as a class action. . . ." FED. R. CIV. P. 23(c)(4)(A). The language of Rule 23(c)(4)(B) arguably contemplates *sua sponte* creation of subclasses: "a class may be divided into subclasses. . . ." FED. R. CIV. P. 23(c)(4)(B). However, this language does not clearly authorize *sua sponte* creation of subclasses.

[183] *See supra* note 136 and accompanying text.

[184] *See supra* notes 112–15 and accompanying text.

[185] *See* Appendix B, Proposed Advisory Committee Note to suggested revised Rule 23(c)(4).

[186] *See supra* text accompanying note 120.

nificant role in mass tort class actions.[187] However Rule 23, not having been drafted as a tool for mass tort litigation, makes no mention of a court's discretion or obligation to consider the effect of a class action certification on settlement. To make Rule 23 more useful and effective in mass tort class actions, the Rule could be revised to encourage judges to consider the practical effect of class action treatment on the prospects for extra-judicial settlement.[188] Thus, Rule 23(b)(3) might be supplemented as follows to include the effect on settlement of a class action certification as an element to be considered in determining whether a class action is "superior to other available methods for the fair and efficient adjudication of the controversy"[189]:

> Rule 23(b)(3): . . . **(F) the effect of class action treatment on the prospects for an equitable pretrial or pre-judgment settlement of the litigation.**[190]

While perhaps not adding much to the Rule 23(b)(3) analysis most federal judges currently perform, this revision, by creating an inherent bias toward a finding of predominance and superiority in proposed Rule 23(b)(3) mass tort class actions, might have an important psychological impact in signaling to the judiciary a reversal of the former bias against the use of Rule 23 in mass tort litigation. In addition, and more generally, such a revision of Rule 23(b)(3) could help signal to the legal profession what may be a growing policy preference for equitable, extra-judicial settlements of mass tort cases.

H. *Amount in Controversy*

In order to make the class action device most effective in non-catastrophic, multistate mass tort cases, the $10,000 jurisdictional amount in controversy requirement of 28 U.S.C.

---

[187] *See supra* note 116 and accompanying text. According to Judge Williams:

> While the [Ninth] Circuit may speculate as to the effort of "a few verdicts" trial judges know the answer: a class adjudication, even on limited issues, which resolves questions of liability and punitive damages will have far greater influence on parties' willingness to settle, and will certainly do a better job of relieving judges of litigation "re-runs" for those who face a series of identical pending cases.

Williams, *supra* note 10, at 328.

[188] *See supra* notes 116–25 and accompanying text.

[189] FED. R. CIV. P. 23(b)(3).

[190] *See* Appendix B, Proposed Advisory Committee Note to suggested new Rule 23(b)(3)(F).

§ 1332 could be waived or revised so as to allow aggregation of claims in mass tort cases. This would also relieve federal judges of having to circumvent the present jurisdictional requirement by resorting to the "legal certainty" doctrine of *St. Paul Indemnity Co. v. Cab Co.*[191] A revision of 28 U.S.C. § 1332 to waive automatically the amount in controversy requirement for multistate mass tort litigation might read as follows (new language in bold face):

> § 1332. (a) The district courts shall have original jurisdiction **of all multistate mass tort litigation and** of all civil actions where the matter in controversy exceeds the sum or value of $10,000. . . .[192]

Alternatively, a revision of 28 U.S.C. § 1332 to allow aggregation of claims might involve the addition of a new subpart to read as follows:

> § 1332 . . . **(e) For purposes of this section, the measure of the amount in controversy in multistate mass tort class actions will be an aggregation of the claims of the individual class members.**

I. *Conclusion*

Finally, but perhaps no less important from a practical psychological standpoint, if Rule 23 is to be revised and made more usable and effective in mass tort litigation, the 1966 Advisory Committee's comment about the inappropriateness of class action treatment for mass tort litigation should be deleted from the notes to the Rule.[193] Many federal judges no longer believe

---

[191] *See supra* notes 127–28 and accompanying text.

[192] *See* 28 U.S.C. § 1332 (1982).

[193] The inhibitive effect of the Advisory Committee's comment is evidenced in the following opinions in which a federal district judge quoted or paraphrased the Advisory Committee's comment for support in refusing to certify a class action: Pearl v. Allied Corp., No. 82-2931, slip op. (E.D. Pa. Aug. 28, 1984); Ryan v. Eli Lilly & Co., 84 F.R.D. 230, 231 (D.S.C. 1979); Mink v. Univ. of Chicago, 27 Fed. R. Serv. 2d (Callaghan) 739, 741 (N.D. Ill. 1979); Marchesi v. Eastern Airlines, Inc., 68 F.R.D. 500 (E.D.N.Y. 1975); Causey v. Pan Am World Airways, Inc., 66 F.R.D. 392, 397 (E.D. Va. 1975); Yandle v. PPG Indus. Inc., 65 F.R.D. 566, 569 (E.D. Tex. 1974); Daye v. Commonwealth of Pa., 344 F. Supp. 1337, 1343 (E.D. Pa. 1972). After quoting the Advisory Committee's comment and citing three factually similar cases that followed the comment and one that did not, Judge Bramwell in *Marchesi* wrote, "Accordingly, since the weight of authority is clear, it is the view of this Court that class action certification should be denied and that all class action allegations in the complaint should be stricken." *Marchesi*, 68 F.R.D. at 501–02. Judge Bramwell made no findings or determinations as to whether the plaintiffs in *Marchesi* met *any* of Rule 23's prerequisites or requirements.

class actions are inappropriate in mass tort litigation, and several judges have shown how Rule 23, even as presently written, can be an appropriate and effective tool in mass tort litigation.[194] What Judge Wright said about the use of the class action device in *Federal Skywalk*[195] is perhaps equally applicable to most mass tort cases that arise in our society and perhaps best exemplifies the spirit of the several innovative and imaginative federal district judges who have found ways to use Rule 23 in mass tort litigation:

> The class action device, as contemplated by Rule 23 of the Federal Rules of Civil Procedure, is the proper tool for the accomplishment of principled, efficient and expeditious adjudication of skywalk claims which threaten to expose defendants to repeated trials of the compensatory and punitive damages issues, which threaten to leave many claimants without a practical chance for redress of punishable acts, and which threaten to congest the state and federal courts of this city for many years. This Court views Rule 23 as the key building block in the federal courts' continuing effort to make the civil procedural system more responsive to the needs of contemporary litigation. The magnitude of the litigation spawned by the collapse of two skywalks challenges this Court to administer these cases with flexibility and imagination.[196]

Deleting the Advisory Committee's comment may itself lead to greater use of Rule 23 in mass tort cases, which in turn should provide the judiciary with more examples of how class actions can be used effectively and appropriately in mass tort litigation.

Id. One commentator had the following to say concerning the present applicability of the Advisory Committee's comment:

> The economies of time, effort, and expense of the class action device cut across categorical tort lines and ought not to be obscured by the narrow application of circumstances or by undue emphasis on traditional interests in one-to-one litigation. "I was an *ex officio* member of the Advisory Committee on Civil Rules when Rule 23 was amended, which came out with an Advisory Committee Note saying that mass torts are inappropriate for class certification. I thought then that was true. I am profoundly convinced now that that is untrue. Unless we can use the class action and devices built on the class action, our judicial system is simply not going to be able to cope with the challenge of the mass repetitive wrong that we see in this case and so many others that have been mentioned this morning and afternoon." Prof. Charles Alan Wright, In Re: School Asbestos Litigation Master File 83-0268 (E.D. Pa.) Class Action Argument, July 30, 1984, Tr 106.

H. NEWBERG, *supra* note 75, § 17.06 (footnote omitted).

[194] *See supra* note 8.

[195] 93 F.R.D. 415 (W.D. Mo.), *vacated*, 680 F.2d 1175 (8th Cir.), *cert. denied*, 459 U.S. 988 (1982).

[196] *Federal Skywalk*, 93 F.R.D. at 420.

APPENDIX A

REVISED RULE 23[197]

Rule 23. Class Actions

(a) **Prerequisites to a Class Action.** One or more members of a class may sue or be sued as representative parties on behalf of all if (1) the class is so numerous that joinder of all members is impracticable, (2) **there are** questions of law or fact common to the class, **or there is at least one material question of law or fact common to the class,** (3) the claims or defenses of the representative parties are typical of the claims or defenses of the class, and (4) the representative parties will fairly and adequately protect the interests of the class.

(a)(.1) Typicality in Products Liability Class Actions. In products liability litigation brought under this rule against more than one defendant, the court shall find the typicality standard of subpart (3) of subdivision (a) to be met with respect to as many defendants as the representative parties bringing the action can show [were engaged, whether dependently or independently, in some aspect of the design, development, production, testing or marketing of products of the type that are the subject of the litigation], [and/or] [were engaged in a conspiracy or concerted scheme to design, develop, produce or market the product that is the subject of the litigation], [and/or] [are currently or were formerly juridically related in such a manner that a single resolution of the dispute against them would be expeditious].

(b) **Class Actions Maintainable; Mandatory Class Actions.** An action may be maintained as a class action, **and an action must be maintained as a class action when the court deems class action treatment necessary to protect the rights of potential claimants, if the prerequisites of subdivision (a) are satisfied, and in addition:**

(1) the prosecution of separate actions by or against individual members of the class would create a risk of

---

[197] Proposed revisions are highlighted in text.

(A) inconsistent or varying adjudications with respect to individual members of the class which would establish incompatible standards of conduct for the party opposing the class, or

(B) adjudications with respect to individual members of the class which would as a practical matter be dispositive of the interests of the other members not parties to the adjudications or substantially impair or impede their ability to protect their interests; **in making its determinations under this subdivision (B), the court shall consider (i) evidence of the extent of the defendants' financial resources and (ii) the extent to which independent punitive damages awards to prior individual plaintiffs would cause a court in subsequent litigation to find the defendants had been sufficiently punished so that no further awards of punitive damages would be made. For purposes of the determination in subdivision (ii) of this subpart (B) in multistate litigation, the court shall, as among the relevant applicable state punitive damages standards, refer to the punitive damages standard of the state with which the defendants have the most significant contacts; in litigation against multiple defendants who do not all have the most significant contacts with the same state, the court shall refer to the relevant state punitive damages standard most favorable to defendants;**

(2) the party opposing the class has acted or refused to act on grounds generally applicable to the class, thereby making appropriate final injunctive relief or corresponding declaratory relief with respect to the class as a whole; or

(3) the court finds that questions of law or fact common to the members of the class predominate over any questions affecting only individual members, and that a class action is superior to other available methods for the fair and efficient adjudication **or settlement** of the controversy. The matters pertinent to the findings include: (A) the interest of members of the class in individually controlling the prosecution or defense of separate actions, **such interest to be determined by reference both to the potential size of individual recoveries and to the countervailing practical problems individual claimants would likely encounter in bringing individual actions because of the complexity, cost, and other problems inherent in the litigation;** (B) the extent and nature of any litigation concerning the controversy already

commenced by or against members of the class; (C) the desirability or undesirability of concentrating the litigation of the claims in the particular forum; (D) the difficulties likely to be encountered in the management of a class action; **(E) in an action arising from a nucleus of operative facts with respect to which one or more individual cases have already been tried, the savings in time, money, and other judicial and litigant resources that would result from a class action, such savings to be determined by reference to the time consumed and the money and other judicial and litigant resources expended in the prior individual actions; (F) the effect of class action treatment on the prospects for an equitable pretrial or pre-judgment settlement of the litigation. In addition, the court shall consider whether, under subdivision (c)(4) of this rule, partial classes can be certified in which material issues would predominate that would otherwise not predominate with respect to the litigation as a whole. If the court finds that such partial classes can be created and that resolution of the issues for which the partial classes are created will materially advance a disposition of the litigation as a whole, the court shall find that the requirements of predominance and superiority in this subdivision have been met. In litigation in which the alleged wrongful or unlawful conduct occurred in more than one state and different standards of legal conduct apply in the various states in which such alleged conduct occurred, the court may nevertheless preserve the predominance of questions of law for purposes of this subdivision by requiring the parties seeking class certification to show, in the case of plaintiffs, violation of the most permissive relevant state standard of legal conduct, or, in the case of defendants, compliance with the strictest relevant state standard of legal conduct.**

(c) [Sua Sponte Class Certification; Court] Determination by Order Whether Class Action to be Maintained; Notice; Judgment; Actions Conducted Partially as Class Actions.

(.1) The court may, on its own motion and regardless of whether any party has requested such treatment, certify a class action under this rule if (A) the court has before it a large number of individual cases that arose out of the same nucleus of operative facts, (B) the court finds plaintiffs who will willingly act as representative parties, and (C) the class action meets the prerequisites of subdivision (a) and the requirements of subdivision (b) of this

rule. A certification under this subdivision may be conditional and may be altered or amended before the decision on the merits.

(1) As soon as practicable after the commencement of an action brought as a class action, the court shall determine by order whether it is to be so maintained. An order under this subdivision may be conditional and may be altered or amended before the decision on the merits.

(2) In any class action maintained under subdivision (b)(3), the court shall direct to the members of the class the best notice practicable under the circumstances, including individual notice to all members who can be identified through reasonable effort. The notice shall advise each member that (A) the court will exclude him from the class if he so requests by a specified date; (B) the judgment, whether favorable or not, will include all members who do not request exclusion; and (C) any member who does not request exclusion may, if he desires, enter an appearance through his counsel.

(3) The judgment in an action maintained as a class action under subdivision (b)(1) or (b)(2), whether or not favorable to the class, shall include and describe those whom the court finds to be members of the class. The judgment in an action maintained under subdivision (b)(3), whether or not favorable to the class, shall include and specify or describe those to whom the notice provided in subdivision (c)(2) was directed, and who have not requested exclusion, and whom the court finds to be members of the class.

(4) When appropriate (A) an action may be brought or maintained as a class action with respect to particular issues, or (B) a class may be divided into subclasses and each subclass treated as a class, and the provisions of this rule shall then be construed and applied accordingly. **The court shall create partial classes under either or both parts of this subdivision when so doing would isolate material issues in the litigation the resolution of which would materially advance the disposition of the litigation as a whole, irrespective of whether the litigation as a whole proceeds under this rule. The court may, on its own motion, create partial classes under either or both of the parts of this subdivision if the requirements of this subdivision are met and if the court finds**

plaintiffs who will willingly act as representative parties for the partial classes created.

(d) **Orders in Conduct of Actions.** In the conduct of actions to which this rule applies, the court may make appropriate orders: (1) determining the course of proceedings or prescribing measures to prevent undue repetition or complication in the presentation of evidence or argument; (2) requiring, for the protection of the members of the class or otherwise for the fair conduct of the action, that notice be given in such a manner as the court may direct to some or all of the members of any step in the action, or of the proposed extent of the judgment, or of the opportunity of members to signify whether they consider the representation fair and adequate, to intervene and present claims or defenses, or otherwise to come into the action; (3) imposing conditions on the representative parties or on intervenors; (4) requiring that the pleadings be amended to eliminate therefrom allegations as to representation of absent persons, and that the action proceed accordingly; (5) dealing with similar procedural matters. The orders may be combined with an order under Rule 16 and may be altered or amended as may be desirable from time to time.

(e) **Dismissal or Compromise.** A class action shall not be dismissed or compromised without the approval of the court, and notice of the proposed dismissal or compromise shall be given to all members of the class in such manner as the court directs.

APPENDIX B

PROPOSED ADVISORY COMMITTEE NOTE TO REVISED RULE 23

In its Note to the 1966 Amendments to Rule 23, the Advisory Committee stated "a 'mass accident' resulting in injuries to numerous persons is ordinarily not appropriate for a class action. . . ."[198] In the more than two decades since the enactment of the 1966 Amendments, mass tort litigation has become fairly

[198] Advisory Note, *supra* note 1.

common, and several federal judges have shown how the class action device can and should be used effectively and efficiently in such litigation. The Amendments described in this Note are an attempt to codify and, to some extent, legitimate some of the techniques adopted by federal judges to make Rule 23 more useful in mass tort cases and other complex litigation. The 1966 Advisory Committee's comment about the inapplicability of the class action device to mass tort litigation is hereby deleted. A class action is appropriate for the adjudication or settlement of at least some of the questions of law or fact in many mass tort and mass accident cases.[199]

Subdivision (a)(1). (No revisions proposed).

Subdivision (a)(2). This subdivision has been supplemented to make clear that common questions of law or fact are not always a prerequisite to a class action: a single, material question of law or fact common to the class is sufficient to meet the subdivision (a)(2) commonality requirement. This revision of subdivision (a)(2) is intended to make class certification more likely in mass tort cases and other complex litigation in which the class action device would be useful for the efficient and conclusive determination of at least one material question of law or fact such as the question of liability for or the cause or impact of wrongful or unlawful activity.[200]

Subdivision (a)(.1) [first alternative]. This subdivision is added to allow plaintiffs in products liability litigation brought under this Rule against more than one defendant to meet the typicality standard of subdivision (a)(3) of this Rule by showing that the defendants were in some way involved in bringing the product in question to the market. This subdivision codifies the so-called "enterprise liability" theory previously adopted by some courts.[201]

---

[199] *See In re School Asbestos Litig.*, 789 F.2d 996 (3d Cir.), *cert. denied sub nom. Celotex Corp. v. School Dist. of Lancaster*, 107 S. Ct. 182 (1986), *cert. denied sub nom. National Gypsum Co. v. School Dist. of Lancaster*, 107 S. Ct. 318 (1986); MANUAL FOR COMPLEX LITIGATION, SECOND § 33.24 (1986).

[200] *See In re School Asbestos Litig.*, 789 F.2d 996 (3d Cir.), *cert. denied*, 107 S. Ct. 182 (1986), *cert. denied*, 107 S. Ct. 318 (1986).

[201] *See Payton v. Abbott Labs*, 83 F.R.D. 382 (D. Mass. 1979), *vacated*, 100 F.R.D. 336 (D. Mass. 1983).

Subdivision (a)(.2) [second alternative]. This subdivision is added to Rule 23 to allow plaintiffs in products liability litigation brought under this Rule against more than one defendant to meet the typicality standard of subdivision (a)(3) of this Rule by showing that the defendants conspired in some part to bring, or acted in concert in bringing, the product in question to the market. This subdivision codifies an exception to the doctrine espoused by some courts that Rule 23(a)(3) "typicality is lacking when the representative plaintiff's cause of action is against a defendant unrelated to the defendants against whom the cause of action of the members of the class lies."[202] Though a mere allegation by plaintiffs that defendants engaged in a conspiracy may in some cases satisfy this subdivision,[203] courts should exercise discretion in every case to determine whether to require of plaintiffs some additional proof of conspiracy or concerted scheme for purposes of this subdivision.

Subdivision (a)(.3) [third alternative]. This subdivision is added to Rule 23 to allow plaintiffs in products liability litigation brought under this Rule against more than one defendant to meet the typicality standard of subdivision (a)(3) of this Rule by showing that the defendants were or are juridically related. This subdivision codifies an exception to the doctrine espoused by some courts that Rule 23(a)(3) "typicality is lacking when the representative plaintiff's cause of action is against a defendant unrelated to the defendants against whom the cause of action of the members of the class lies."[204] In general, plaintiffs' showing under this subdivision should be easier to make when defendants are vertically integrated businesses and more difficult to make when defendants are competing businesses.

Subdivision (b)(1). This subdivision is supplemented to require courts to certify a class action under subdivision (b)(1)(A) or (b)(1)(B) of this Rule when the court deems the class action device necessary to protect the rights of potential claimants. For example, if the court hearing a motion for a Rule 23(b)(1) class action finds that the prosecution of separate actions would almost certainly lead to the outcomes described in subdivision

---

[202] *La Mar v. H & B Novelty & Loan Co.*, 489 F.2d 461 (9th Cir. 1973).

[203] *Cf. In re N. Dist. of Cal. "Dalkon Shield" IUD Product Liab. Litig.*, 693 F.2d 847 (9th Cir. 1982), *cert. denied sub nom. A.H. Robins Co. v. Abed*, 459 U.S. 1171 (1983).

[204] *La Mar v. H & B Novelty & Loan Co.*, 489 F.2d 461 (9th Cir. 1973).

(b)(1)(A) or (b)(1)(B), the court must certify a class action under either subdivision (b)(1)(A) or (b)(1)(B), whichever applies. In determining whether class action treatment under this subdivision is necessary to protect the rights of potential claimants, the court should weigh the threat to such claimants' due process rights from not being able to opt out of a class action under Rule 23(b)(1),[205] against the risk such claimants face of not being able to recover at all from defendants if a class action is not certified. This latter risk may arise when independent damages awards bankrupt defendants before other claimants can recover damages,[206] or when independent punitive damages awards make later punitive damages awards less likely under state laws or general notions of fairness and sufficient punishment.[207] As revised, this subdivision should increase the use of Rule 23(b)(1) class actions in mass tort litigation.[208]

Subdivision (b)(1)(B). This revised subdivision directs courts to apply the "limited fund" and "sufficiently punished" theories developed by some courts to determine whether a risk exists under this subdivision that absent parties will not be able to protect their interests in recovering damages as a result of earlier, independent awards of compensatory and punitive damages.[209] The operation of this subdivision and the revised subdivision (b)(1) should lead to greater use of the class action device in mass tort litigation.[210]

---

[205] *See Phillips Petroleum Co. v. Shutts*, 472 U.S. 797 (1985).

[206] *See In re Federal Skywalk Cases*, 93 F.R.D. 415 (W.D. Mo.), *vacated on other grounds*, 680 F.2d 1175 (8th Cir.), *cert. denied sub nom. Rau v. Stover*, 459 U.S. 988 (1982); *Coburn v. 4-R Corp.*, 77 F.R.D. 43 (E.D. Ky. 1977), *mandamus sought by some defendants denied sub nom. Union Light, Heat and Power Co. v. United States Dist. Court*, 588 F.2d 543 (6th Cir. 1978), *cert. denied*, 443 U.S. 913 (1979).

[207] *See In re "Agent Orange" Product Liab. Litig.*, 100 F.R.D. 718 (E.D.N.Y. 1983), *mandamus denied*, 725 F.2d 858 (2d Cir. 1984), *cert. denied sub nom. Diamond Shamrock Chem. Co. v. Ryan*, 465 U.S. 1067 (1984).

[208] *Cf.* Miller & Crump, *Jurisdiction and Choice of Law in Multistate Class Actions After Phillips Petroleum Co. v. Shutts*, 96 YALE L.J. 1 (1986); *but cf. Shutts*, 472 U.S. 797 (1985).

[209] *See In re "Agent Orange" Product Liab. Litig.*, 100 F.R.D. 718 (E.D.N.Y. 1983), *mandamus denied*, 725 F.2d 858 (2d Cir.), *cert. denied sub nom. Diamond Shamrock Chem. Co. v. Ryan*, 465 U.S. 1067 (1984); *In re Federal Skywalk Cases*, 93 F.R.D. 415 (W.D. Mo.), *vacated on other grounds*, 680 F.2d 1175 (8th Cir.), *cert. denied sub nom. Rau v. Stover*, 459 U.S. 988 (1982); *Coburn v. 4-R Corp.*, 77 F.R.D. 43 (E.D. Ky. 1977), *mandamus sought by some defendants denied sub nom. Union Light, Heat and Power Co. v. United States Dist. Court*, 588 F.2d 543 (6th Cir. 1978), *cert. denied*, 443 U.S. 913 (1979).

[210] *See* Rule 23(b)(1).

Subdivision (b)(1)(B) has been further supplemented to provide courts possible punitive damage standards to which courts may refer in making the determination under subdivision (b)(1)(B)(ii) of this Rule with respect to the likelihood that independent punitive damages awards will lead courts in subsequent litigation to conclude that the defendants had been sufficiently punished such that no further punitive damage awards would be made. The state standards to which this subdivision refers the court, that of the state with which defendants have had most significant contacts or the relevant state standard most favorable to defendants, are intended to be used by the court only as a measure for purposes of subdivision (b)(1)(B)(ii) of this Rule, not necessarily as the measure for the actual amount of punitive damages to be awarded after adjudication of claims.[211]

Subdivision (b)(3). In order to clarify the interrelationship between the predominance and superiority requirements of this subdivision and the provision for the creation of partial classes in subdivision (c)(4) of this Rule, this revised subdivision requires the court to consider whether subdivision (c)(4) partial classes can be created to preserve the predominance of material issues in the litigation.[212] If the use of such partial classes would advance the litigation as a whole, the court is to find the predominance and superiority requirements of this subdivision to have been met. This revision should make the Rule 23(b)(3) class action a more useful device in complex litigation, especially in mass tort cases, in which the resolution of a few class-wide issues, or even single class-wide issues such as causation and liability, would save time and resources for litigants and the courts while determining (1) the rights of multiple claimants—many of whom would otherwise not have practical access to the courts to seek redress in individual actions—to recovery for damages, and (2) the rights of defendants to have claims dismissed.

---

[211] *But see Walsh v. Ford Motor Co.*, 106 F.R.D. 378 (D.D.C. 1985); *In re "Agent Orange" Product Liab. Litig.*, 580 F. Supp. 690 (E.D.N.Y. 1980), *mandamus denied*, 725 F.2d 858 (2d Cir.), *cert. denied sub nom. Diamond Shamrock Chem. Co. v. Ryan*, 465 U.S. 1067 (1984), for examples of courts that looked to the states with which defendants had most significant contacts in order to determine the properly applicable punitive damages standard.

[212] *See In re Tetracycline Cases*, 107 F.R.D. 719 (W.D. Mo. 1985); 7A C. WRIGHT, A. MILLER & M. KANE, FEDERAL PRACTICE AND PROCEDURE § 1778 (1986).

Provision allowing the court to require parties seeking class action treatment to show violation of the most permissive or strictest state standard of conduct: this provision is intended to give discretion to courts hearing complex multistate cases, especially mass tort litigation, to preserve the predominance of questions of law in such cases, thus helping preserve the superiority of the class action device in such cases.[213] Though this supplement to subdivision (b)(3) may perhaps appear to abridge, in violation of the Rules Enabling Act, 28 U.S.C. § 2072 (1982), the substantive rights of some litigants to have a relatively more or less stringent state liability standard applied to their conduct or the conduct of an opposing party, we are of the opinion that, even assuming litigants have such a substantive right,[214] this provision, by making the class action device more available in multistate complex litigation, such as mass tort cases, will on the whole allow more claimants to exercise their rights to recover damages for wrongful or unlawful activity. With respect to defendants who seek class action certification, this provision effectively requires a trade off: a showing of compliance with the strictest relevant state standard is required in exchange for allowing defendants to have multiple similar claims against them adjudicated simultaneously.

Subdivision (b)(3)(A) is revised to make explicit the court's obligation to balance the claimants' financial interest in individual litigation against the real problems individual claimants would face in attempting to prosecute individual actions. In performing this balancing, the court should make special note of the complexity of the litigation and the costs individual litigants are likely to face. The assumption underlying this provision, which may not be true in every case, is that the more complex and costly individual actions would be, the less likely claimants will be willing or able to bring the actions individually. This assumption is likely to be true in mass tort cases involving personal injuries; thus, this revised subdivision should make class certification more likely in such cases.[215]

---

[213] *Cf. In re Asbestos School Litig.*, 104 F.R.D. 422 (E.D. Pa. 1984), *aff'd in part and vacated in part*, 789 F.2d 996 (3d Cir.), *cert. denied*, 107 S. Ct. 182 (1986), *cert. denied*, 107 S. Ct. 318 (1986).

[214] *See generally Day & Zimmerman, Inc. v. Challoner*, 423 U.S. 3 (1975); *Klaxon Co. v. Stentor Elec. Mfg. Co.*, 313 U.S. 487 (1941).

[215] *See In re "Agent Orange" Product Liab. Litig.*, 506 F. Supp. 762 (E.D.N.Y. 1980).

Subdivision (b)(3)(E) introduces a fifth matter pertinent to the court's findings of predominance and superiority under this Rule. The court is to consider the time and resources expended in the prosecution of individual suits arising out of the same set of facts in order to determine the extent to which a class action might conserve time and resources for future litigants and the courts.[216] This subdivision should help tip the balance in the predominance and superiority analysis in favor of class action certification in mass tort actions.[217]

Subdivision 23(b)(3)(F). This subdivision is added as an additional factor for courts to consider when determining whether the predominance and superiority requirements of subdivision (b)(3) of this Rule are met. Because class certification will usually result in an increased likelihood of pretrial settlement, this subdivision should generally work in favor of a finding of predominance and superiority. This bias toward class certification is intended to help make the class action device a more effective tool in the efficient and fair resolution of mass tort cases and other complex litigation.[218]

Subdivision (c)(.1). This subdivision has been added to make explicit the court's discretion and power to certify a class action on the court's own motion if the necessary prerequisites under this subdivision and other subdivisions of this Rule are met. This subdivision does not obligate courts to certify class actions *sua sponte*.[219] This subdivision should lead to greater use of the class action device in mass tort litigation, but only in cases in

---

[216] *See General Telephone Co. of Southwest v. Falcon*, 457 U.S. 147 (1982); *Califano v. Yamasaki*, 442 U.S. 682 (1972).

[217] *See Jenkins v. Raymark Indus., Inc.*, 782 F.2d 468 (5th Cir), *reh'g denied*, 785 F.2d 1034 (5th Cir. 1986); *Walsh v. Ford Motor Co.*, 106 F.R.D. 378 (D.D.C. 1985); *In re "Agent Orange" Product Liab. Litig.*, 100 F.R.D. 718 (E.D.N.Y. 1983), *mandamus denied*, 725 F.2d 858 (2d Cir.), *cert. denied sub nom. Diamond Shamrock Chem. Co. v. Ryan*, 465 U.S. 1067 (1984).

[218] *See generally In re School Asbestos Litig.*, 789 F.2d 996 (3d Cir.), *cert. denied*, 107 S. Ct. 182 (1986), *cert. denied*, 107 S. Ct. 318 (1986); *In re "Agent Orange" Product Liab. Litig.*, 100 F.R.D. 718 (E.D.N.Y. 1983), *mandamus denied*, 725 F.2d 858 (2d Cir.), *cert. denied sub nom. Diamond Shamrock Chem. Co. v. Ryan*, 465 U.S. 1067 (1984); Transgrud, *Joinder Alternatives in Mass Tort Litigation*, 70 CORNELL L.REV. 779 (1985).

[219] *See generally Geraghty v. United States Parole Comm'n*, 579 F.2d 238 (3d Cir. 1978), *vacated and remanded*, 445 U.S. 388 (1980); *but cf. Senter v. General Motors Corp.*, 532 F.2d 511 (6th Cir.), *cert. denied*, 429 U.S. 870 (1976); *Pruitt v. Allied Chem. Corp.*, 85 F.R.D. 100 (1980); *Stevenson v. Smith*, 73 F.R.D. 79 (D. Del. 1976).

which representative parties and lead counsel can be found who will adequately and willingly represent the class.[220]

Subdivision (c)(4). As revised, this subdivision requires courts to create partial classes when so doing would allow courts to isolate and make determinations with respect to issues whose resolution would advance the litigation as a whole. This subdivision is a counterpart to the revised subdivision (b)(3) of this Rule, the purpose of both revisions being to make class actions under subdivision (b)(3) more useful and effective in complex litigation, especially mass tort litigation.[221]

Subdivision (c)(4), as supplemented with a sentence about the court's power and discretion to create partial classes *sua sponte*, corresponds with the new subdivision (c)(.1) of this Rule.[222] One effect of this explication of the court's power to certify subclasses on its own motion should be to ensure commonality under subdivision (a)(2) of this Rule, thus making class action treatment more available and useful in mass tort cases and other complex litigation brought under this Rule.[223]

---

[220] *See generally In re N. Dist. of Calif. "Dalkon Shield" IUD Product Liab. Litig.*, 521 F. Supp. 1188 (N.D. Cal. 1981), *vacated and remanded*, 693 F.2d 847 (9th Cir. 1982), *cert. denied sub nom. A.H. Robins Co. v. Abed*, 459 U.S. 1171 (1983); *In re Federal Skywalk Cases*, 93 F.R.D. 415 (W.D. Mo.), *vacated on other grounds*, 680 F.2d 1175 (8th Cir.), *cert. denied sub nom. Rau v. Stover*, 459 U.S. 988 (1982).

[221] *See* Rule 23(b)(3).

[222] *See* Rule 23(c)(.1).

[223] *See In re Federal Skywalk Cases*, 93 F.R.D. 415 (W.D. Mo. 1982), *vacated on other grounds*, 680 F.2d 1175 (8th Cir.), *cert. denied sub nom. Rau v. Stover*, 459 U.S. 988 (1982).

POLICY ESSAY

THE PREVENTION OF INSIDER TRADING: A PROPOSAL FOR REVISING SECTION 16 OF THE SECURITIES EXCHANGE ACT OF 1934

S.S. SAMUELSON*

Passed in the heat and fervor of reform generated by Congressional hearings into stock market abuses, Section 16 of the Securities Exchange Act of 1934[1] was designed to prevent the unfair use of corporate information by insiders: officers, directors, and major shareholders.[2] Section 16 takes a two-pronged approach: (1) insiders are required to report trades in securities of their own company by the tenth day of the following month, and (2) insiders must disgorge to the corporation short-swing profits (i.e., profits made from any purchase and sale, or sale and purchase, of company securities in a six-month period).[3]

---

* Assistant Professor, Management Policy, Boston University. A.B., Harvard University, 1974; J.D., Harvard Law School, 1977.

[1] Securities Exchange Act of 1934, § 16, 48 Stat. 881, 896 (codified as amended at 15 U.S.C. § 78p (1982)) [hereinafter Securities Exchange Act or Act].

[2] Section 16 is the only provision of the federal securities laws explicitly designed to regulate insider trading. The other provisions of the federal securities laws that are used to combat insider trading (*see* Securities Exchange Act of 1934, § 10(b), Rule 10b-5, 48 Stat. 881, 891) are general anti-fraud provisions. Liability under these provisions is not limited to corporate insiders, nor even to those who actually trade.

[3] Section 16 of the Securities Exchange Act provides in part that:

> (a) Every person who is directly or indirectly the beneficial owner of more than 10 per centum of any class of any equity security (other than an exempted security) which is registered pursuant to section 12 of this title, or who is a director or an officer of the issuer of such security, shall file, at the time of the registration of such security on a national securities exchange or by the effective date of a registration statement filed pursuant to section 12(g) of this title, or within ten days after he becomes such beneficial owner, director, or officer, a statement with the Commission (and, if such security is registered on a national securities exchange, also with the exchange) of the amount of all equity securities of such issuer of which he is the beneficial owner, and within ten days after the close of each calendar month thereafter, if there has been a change in such ownership during such month, shall file with the Commission (and if such security is registered on a national securities exchange, shall also file with the exchange) a statement indicating his ownership at the close of the calendar month and such changes in his ownership as have occurred during such calendar month.
>
> (b) For the purpose of preventing the unfair use of information which may have been obtained by such beneficial owner, director, or officer by reason of his relationship to the issuer, any profit realized by him from any purchase and

The Act also prohibits "short sales"[4] and "sales against the box"[5] by corporate insiders.

Section 16 has not, however, accomplished Congressional intent. A half century of experience with Section 16 and recent analysis by economists of insider trading reveal that the Act has not prevented insiders from profiting at the expense of other shareholders and the general public.[6] The evidence also suggests that corporate officials not only earn a higher return than the market when trading stock of their own companies, but that insiders frequently trade while in possession of secret, material information.[7] Indeed, since the statute only regulates round-trip trading, it does not prohibit single trades based on secret information. At the same time, Section 16 imposes a heavy burden on corporate insiders by preventing many legitimate trades (i.e., those made without inside information). The reporting requirements often snag not the guilty but the unwary.[8]

Such failure suggests that although the general purpose of the statute is laudable,[9] Section 16 itself is seriously flawed. Two commentators conclude that, "[j]udging solely from the facts stated in the opinions in the decided cases, the function of Section 16(b) would appear to be to impose unjust liability upon

---

> sale, or any sale and purchase, of any equity security of such issuer (other than an exempted security) within any period of less than six months, unless such security was acquired in good faith in connection with a debt previously contracted, shall inure to and be recoverable by the issuer, irrespective of any intention on the part of such beneficial owner, director, or officer in entering into such transaction of holding the security purchased or of not repurchasing the security sold for a period exceeding six months.
>
> (c) It shall be unlawful for any such beneficial owner, director, or officer, directly or indirectly, to sell any equity security of such issuer (other than an exempted security), if the person selling the security or his principal (1) does not own the security sold, or (2) if owning the security, does not deliver it against such sale within twenty days thereafter, or does not within five days after such sale deposit it in the mails or other usual channels of transportation. . . .

15 U.S.C. § 78p (1982). *Cf.* Securities and Exchange Act of 1934, § 10(b), 48 Stat. 881, 891 (section 16 is a per se rule, unlike section 10(b), where such factors as scienter and reliance are critical).

[4] A "short sale" is a contract for the sale of stock which the seller does not own, but rather borrows against the future sale.

[5] A "sale against the box" is a type of short sale in which the insider sells borrowed stock identical to stock that he holds.

[6] *See infra* notes 37–42 and accompanying text.

[7] *Id.*

[8] American Law Institute, FEDERAL SECURITIES CODE, § 1714 comment at 751 (1980).

[9] Some commentators argue that insider trading ought not to be regulated at all. *See infra* note 73.

entirely innocent persons."[10] It is therefore time to take a serious look at either revising or repealing the statute.

Part I of this Essay examines more closely the purposes of Section 16. Part II discusses evidence that the Section fails to achieve these goals. Part III proposes revisions of the Section that would more effectively carry out Congressional intent. Part IV, finally, concludes and summarizes the findings of this Essay.

I. SECTION 16: CONGRESSIONAL INTENT

In the 1930's, insider speculation and manipulation of stock was generally not prohibited by federal or state law. Most state courts had held that officers and directors of a corporation engaging in market trading of corporate stock owed no fiduciary duty to shareholders with respect to their stock market trading, and were not subject to liability for trading on inside information.[11] During hearings before Congress in 1933 and 1934,[12] many witnesses testified about stock market abuses that were believed to have caused a lack of investor confidence and ultimately contributed to the 1929 stock market crash. As revealed at the hearings and in the resulting House and Senate Reports,[13] Congress' primary concern in passing Section 16 was to protect the interests of public market participants against company insiders who were able to take advantage of privileged information to manipulate stock prices.[14] The counsel for the Senate Banking

[10] R. JENNINGS & H. MARSH, SECURITIES REGULATION 1173 (4th ed. 1977).

[11] Committee on Federal Regulation of Securities, *Report of the Task Force on Regulation of Insider Trading, Part II: Reform of Section 16*, 42 BUS. LAW. 1087, 1090 (1987) [hereinafter *Report of the Task Force, Part II*].

[12] *Stock Exchange Practices: Hearings before the Senate Banking and Currency Committee on S. Res. 84 (72d Congress) and S. Res. 56 and S. Res. 97 (73d Congress)*, 73d Cong., 2d Sess. (1934), *reprinted in* LEGISLATIVE HISTORY OF THE SECURITIES ACT OF 1933 AND SECURITIES EXCHANGE ACT OF 1934 item no. 22 (compiled by J.S. Ellenberger and Ellen P. Mahar (1973)) [hereinafter *Senate Hearings*]; *Stock Exchange Regulation: Hearings before the House Interstate and Foreign Commerce Committee on H.R. 7852 and H.R. 8720*, 73d Cong., 2d Sess. item no. 23 (1934) [hereinafter *House Hearings*].

[13] SENATE BANKING AND CURRENCY COMMITTEE, STOCK EXCHANGE PRACTICES, S. REP. NO. 1455, 73d Cong., 2d Sess. item no. 21 (1934) [hereinafter SENATE REPORT I]; S. REP. NO. 792 to accompany S. 3420, 73d Cong., 2d Sess. item 17 (1934) [hereinafter SENATE REPORT II]; H. REP. NO. 1383 to accompany H.R. 9323, 73d Cong., 2d Sess. item 18 (1934) [hereinafter HOUSE REPORT].

[14] Rubin & Feldman, *Statutory Inhibitions Upon Unfair Use of Corporate Information by Insiders*, 95 U. PA. L. REV. 468, 468 (1947).

and Currency Committee compared the stock market to a poker game with marked cards.[15]

The first step necessary to protect the interests of the public was thought to be the prevention of trading by corporate officials and beneficial owners on specific secret information.[16] The drafters of the Securities Exchange Act recognized that it was futile for various provisions of the Act to require companies to furnish shareholders with extensive information about the corporation[17] if insiders could undermine the value of that information by trading on it before the investor even received it.[18] Congress also sought to regulate the use of "softer" information considered relevant to investment decisionmaking but which insiders are not required to disclose, such as: the likely success of new products, the real risks of disclosed contingencies, and the competence of the chief executive officer.[19] By requiring disgorgement of short-swing trading profits, Congress hoped to eliminate the incentives to misuse inside information.

Witnesses at the Congressional hearings had testified about insiders who deliberately manipulated the market price of a stock for the purpose of making a quick profit.[20] An insider could buy a large block of stock, announce a substantial dividend, and then divest before the dividend was reduced. Likewise, pools of insiders were able to buy large blocks of stock (sometimes even using corporate funds), to drive up the price

---

[15] *Senate Hearings*, *supra* note 12, item 15; *see also* Letter from the Counsel for the Senate Committee on Banking and Currency under S. Res. 84 to the Senate Committee on Banking and Currency (Feb. 18, 1933) [hereinafter Banking and Currency Committee Letter].

[16] *See Senate Hearings*, *supra* note 12, at 7741:

> The theory behind the beneficial owner concept was that ownership of 5 percent of the stock would practically constitute him an insider, and by virtue of that position he could acquire confidential information which he might use for his own enrichment by trading in the open market, against the interests of the general body of the stockholders.

(statement of Ferdinand Pecora, Counsel to the Senate Committee on Banking and Currency); *House Hearings*, *supra* note 12, at 132 ("That [Section 16(a)] is to prevent the insider from taking advantage of information to sell or buy shares ahead of the release of information to the public about the company") (statement of Thomas G. Corcoran, Counsel with the Reconstruction Finance Corporation); *Senate Report II*, *supra* note 13, at 9 ("The bill [the Securities Exchange Act of 1934] further aims to protect the interests of the public by preventing directors, officers, and principal stockholders of a corporation, the stock of which is traded in exchanges, from speculating in the stock on the basis of information not available to others.").

[17] *See* Securities Exchange Act of 1934, §§ 10, 14, 48 Stat. 881, 891, 895.

[18] *See* Rubin & Feldman, *supra* note 14, at 469.

[19] *Report of the Task Force, Part II*, *supra* note 11, at 1092.

[20] *See id.* at 1093–94; SENATE REPORT II, *supra* note 13; Banking and Currency Committee Letter, *supra* note 15, at 7–32.

of the shares before unloading the securities at a profit. Section 16(b) removes the temptation for such round-trip trading by eliminating the profit from this kind of manipulation. By removing the incentive to focus on short-term price fluctuations at the expense of the long-term financial health of the company, Congress hoped to encourage insiders to uphold their managerial and fiduciary responsibilities.[21]

It was essential to prevent insiders from trading on secret information and manipulating the market, but that was only the first step in restoring investor confidence. The legislative record indicates that the drafters also felt it was necessary to ensure that insiders were held to a higher standard than pertains in an arm's length transaction. Congress expected insiders to act as fiduciaries and agents. According to the House Report, "[a] renewal of investors' confidence in the exchange markets can be effected only by a clearer recognition upon the part of the corporate managers of companies whose securities are publicly held of their responsibilities as trustees for their corporations."[22]

At common law, a fiduciary is under a duty to his beneficiary to act solely in the beneficiary's interest as a trustee. This duty includes an obligation to deal fairly with the beneficiary and to communicate to him all material facts.[23] Thus, the insider/trustee is required to disclose to the shareholders of his corporation any information that might be relevant to them when making a

---

[21] Banking and Currency Committee Letter, *supra* note 15, at 32; *see also* SENATE REPORT I, *supra* note 13, at 68.

[22] HOUSE REPORT, *supra* note 13, at 13; *see also* SENATE REPORT I, *supra* note 13, at 55. Likewise, Professor Victor Brudney has stated that:

> The obligations of the corporation and its insiders to disclose nonpublic corporate information in dealing with security holders may appropriately be seen as an extension of the arrangements protecting beneficiaries against overreaching by fiduciaries, which the common law was haltingly fashioning to constrain corporate insiders when the federal securities legislation was enacted.

Brudney, *Insiders, Outsiders and Informational Advantages Under the Federal Securities Laws*, 93 HARV. L. REV. 322, 326 (1979). Thus, in one of the earliest cases brought under Section 16, the court stated:

> we must suppose that the statute was intended to be thorough-going, to squeeze all possible profits out of stock transactions, and thus to establish a standard so high as to prevent any conflict between the selfish interest of a fiduciary officer, director, or stockholder and the faithful performance of his duty.

Smolowe v. Delendo Corp., 136 F.2d 231, 239 (2d Cir.), *cert. denied*, 320 U.S. 751 (1943). On the issue of agency, see *House Hearings*, *supra* note 12, at 43 ("[It] is simply an application of an old principle of the law that if you are an agent and you profit by inside information concerning the affairs of your principal, your profits go to your principal.") (statement of Thomas G. Corcoran).

[23] RESTATEMENT (SECOND) OF TRUSTS § 170 (1959).

decision about the purchase or sale of corporate securities.[24] An insider also, through control of the corporation, creates an impression of corporate affairs which he has an affirmative duty to correct if it becomes inaccurate at any time.[25] Most importantly, a trustee must account to the beneficiary for any personal profit made, even if that profit is not the result of a breach of trust.[26]

An agent at common law has similar responsibilities—including loyalty and accountability for profits arising out of his employment.[27] An agent may not use confidential information to the disadvantage of the principal and, in addition, must account for any profits made by the use of such information, even if its use has not harmed the principal.[28] So, for example, under common law if a corporation decides to locate a facility at a certain site, an officer who thereafter buys land nearby must turn over any profit he makes from the sale of the land, even though the purchase does not adversely affect his company.[29] To apply this principle to the purchase and sale of stock by a corporate official, that officer must turn over to the corporation any profits earned from trading stock if such profit is based on the informational advantage gained from his position as an insider.[30]

Congress viewed full and prompt disclosure of detailed corporate information, including stock trading by insiders, as a necessary obligation of corporate fiduciaries.[31] Under Section 16, not only are insiders required to report all trades of their company's stock within ten days after the end of each month,[32] but also they are forbidden from making sales against the box.[33] As the Senate Report states, "[t]he vice of this practice [sales

---

[24] *See* Brudney, *supra* note 22, at 344 ("Denying insiders the informational advantage could rest merely upon the narrow premise that they are not entitled to it in dealing with corporate stockholders because they acquired the information as agents or fiduciaries of the stockholders in the course of pursuing the latter's business.").

[25] *Id.* at 345.

[26] RESTATEMENT (SECOND) OF TRUSTS § 203 (1959).

[27] *Id.* §§ 387, 388.

[28] *Id.* § 388.

[29] *Id.*

[30] As Professor Brudney has pointed out, the insider's fiduciary responsibility is not affected by the fact that he has sold rather than bought stock: "[h]e is no more justified in resorting to concealment in order to induce a person to become his beneficiary than to cease to be his beneficiary." Brudney, *supra* note 22, at 344 & n.74.

[31] "[T]he most potent weapon against the abuse of inside information is full and prompt publicity." HOUSE REPORT, *supra* note 13, at 13.

[32] 5 U.S.C. § 78p(a) (1982). *See supra* note 3.

[33] 5 U.S.C. § 78p(c) (1982). *See supra* note 5.

against the box] . . . is that it is usually done in another name or in such a manner as to prevent the public from knowing that an officer or director in a certain corporation is disposing of his stock."[34]

The task Congress set for itself in 1934 was enormous: not only to prevent insiders from taking advantage of others by manipulating stock prices, but also to require insiders to act affirmatively in the best interests of their stockholders. Although the tools chosen to effect these goals—disclosure of trades and disgorgement of short-swing profits—are not adequate to the task, some commentators have suggested that Section 16 is valuable for the moral suasion it brings to bear on the business community.[35] In 1934, as now, members of Congress recognized the extreme difficulty of regulating market manipulation by insiders. Congress, however, did not intend merely to engage in moral suasion. As Congressman Wolverton stated during the House hearings, "I am interested in the practical side of it and not the theoretical side. I am not in favor of just putting words into a bill without giving serious consideration to their effectiveness."[36]

II. SECTION 16: THE RESULTS

Over the last decade economists have studied insider trading, not out of concern over legal issues, but as part of an examination of the efficiency of the stock market. Their research has consistently found that insiders make greater returns trading the stock of their own companies than other shareholders of the company do.[37] In order to evaluate the effectiveness of Section

---

[34] Banking and Currency Committee Letter, *supra* note 15, at 7.

[35] Although the reporter of the Model Federal Securities Code acknowledges that much legitimate criticism has been leveled at Section 16, "[t]he Code proceeds on the theory, however, that 16(b) [of the Securities Exchange Act] has a symbolic significance that must be, and deserves to be, recognized." FEDERAL SECURITIES CODE § 1714 comment at 751 (1980). *See also* H. MANNE, INSIDER TRADING AND THF STOCK MARKET 28 (1966).

[36] *House Hearings*, *supra* note 12, at 136.

[37] Joseph Finnerty found in his early study of insider trading that "insiders earn above average returns when they buy securities of their respective corporations." Finnerty, *Insiders and Market Efficiency*, 31 J. FIN. 1141, 1146 (1976).

For a discussion of economic research on insider trading, see Baesel & Stein, *The Value of Information: Inferences from the Profitability of Insider Trading*, 14 J. FIN. & QUANTITATIVE ANALYSIS 553, 554 (1979).

For other research on insider trading, see Penman, *Insider Trading and the Dissemination of Firms' Forecast Information*, 55 J. BUS. 479 (1982); Elliott, Morse & Rich-

16 in achieving Congressional goals, it is important to understand at the outset how insiders profit from trading activities.[38]

A. *The Use and Misuse of Corporate Information*

Despite Section 16 and other federal anti-fraud provisions,[39] corporate insiders still take unfair advantage of specific and secret information. Anecdotal evidence includes, for example, the 1982 sale by Warner Communications insiders of $7 million worth of stock within the three months before it became publicly known that the company had dramatically decreased its fourth-quarter earnings projections.[40] Such anecdotes simply confirm more rigorous economic findings. Penman determined that corporate insiders time their trades to benefit from announcements of their firms' earnings prospects.[41] Elliott, Morse, and Richardson examined insider trading surrounding the public announcements of annual earnings, large dividend changes, bond rating changes, mergers, and bankruptcies. They found that "insider trading is generally consistent with insiders' using private information in a profitable manner."[42]

Although Section 16 may have had some deterrent effect, the evidence shows that trading by insiders on secret information is still routine.

---

ardson, *The Association Between Insider Trading and Information Announcements*, 15 RAND J. ECON. 521 (1984); Givoly & Palmon, *Insider Trading and the Exploitation of Inside Information: Some Empirical Evidence*, 58 J. BUS. 69 (1985).

Lastly, Norman Fosback, editor of *Insiders*, a newsletter devoted to insider trading, states, "Historical studies dating back to the early 1960's show rather uniformly that stocks puchased by several insiders tend to outperform the markets 2-1 over the year following the insiders' trades." Fosback, *Alarming Habits of the Insider*, Boston Globe, Apr. 18, 1983, at 15, col. 2.

[38] See Appendix for a technical discussion of the economic research.

[39] *See* Securities Exchange Act of 1934, §§ 10, 14(c), 48 Stat. 881, 891, 895.

[40] Fosback, *supra* note 37, at 15, col. 3.

[41] Penman, *supra* note 37.

[42] Elliott, Morse & Richardson, *supra* note 37, at 535.

It could be argued, perhaps, that since SEC v. Texas Gulf Sulphur Co., 446 F.2d 1301 (2d Cir.), *cert. denied*, 404 U.S. 1005 (1971), Rule 10b-5 has been the first line of defense against insiders who trade on specific information. *See, e.g.*, H. MANNE, WALL STREET IN TRANSITION 58 (1974). However, a study comparing the volume and profitability of insider trading before and after *Texas Gulf Sulphur* concluded that the case had not had any impact on trading by company insiders. Jaffe, *The Effect of Regulation Changes on Insider Trading*, 5 BELL J. ECON. 93, 105 (Spring 1974). It would appear, therefore, that the burden of preventing insider trading by corporate officials is not being adequately carried by Rule 10b-5.

B. *Insiders As Leading Indicators of Company Prospects*

To discourage insider trading, Section 16 requires company officials to report all trades in the stock of their own company. Although Congress expected that investors would find these reports relevant in making their own decisions,[43] it rejected concerns of the investment community that the reports would "create far more harm than the most vicious form of tip."[44] Ironically, the investment community was right. Even if insiders do not know specific information about earnings or product development, for example, they are believed to be privy to "softer" information that they are not required to disclose but may be relevant in making investment decisions.[45] Insiders profit from making the required disclosures because the market mimics their transactions.

Analysts and investors assiduously track stock trades by corporate officials, which are published by the Commission each month in the Official Summary of Security Transactions and Holdings.[46] As a result, the wisdom of such trades becomes a self-fulfilling prophecy. If the market reads an insider sale as a

---

[43] The reporting provision was intended "to give investors an idea of the purchases and sales by insiders which may in turn indicate private opinion as to prospects of the company." HOUSE REPORT, *supra* note 13, at 24.

[44] *House Hearings*, *supra* note 12, at 488 (statement of Mr. Hancock).

[45] *See supra* text accompanying note 19; *see also* Fosback, *supra* note 37 point cite: "They [insiders] have greater insight into their companies' day-to-day affairs than any Wall Street analyst, and often command a better advantage of their companies' long-term promise."

[46] For evidence that the investment community follows trades by insiders with great interest, see Jaffe, *Special Information and Insider Trading*, 47 J. BUS. 410 (1974); Wu, *An Economist Looks at Section 16 of the Securities Exchange Act of 1934*, 68 COLUM. L. REV. 260, 268 (1968) ("It is certainly reasonable to assume that insider trading affects public expectation. People in the market who realize that insiders have advantages in trading try to follow their market behavior."); Sandler, *Sperry Insider Selling Sends a Bearish Signal About the Stock to Some Money Managers*, Wall St. J., Jan. 20, 1986, at 35, col. 1 ("Some money managers find it strange that Sperry's top brass were selling Sperry shares while declaring their faith in the computer company's future"); Rundle, *Insiders' Surge to Buy Regional Bank Stocks Spurs New Round of Rosy Forecasts for Group*, Wall St. J., Dec. 23, 1985, at 37, col. 1 ("A recent surge in purchases of regional bank stocks by corporate insiders is helping to convince some people that the marathon bull market in this group still has further to run"); McFadden, *Companies the Bosses Bet On*, FORTUNE, Jul. 21, 1986, at 112 ("No praise for a company is more impressive than the decision of an insider to lay out his own money for the stock," followed by a list of companies in which insiders had been trading); Marcial, *Taking a Cue from Company Insiders*, BUS. WK., May 12, 1986, at 88; Koretz, *Tracking Corporate Insiders to Get in on a Good Thing*, BUS. WK., May 19, 1986, at 34; Boland, *The Insiders Turn Surprisingly Bullish*, N.Y. Times, Jul. 12, 1987, § 3, at 10, col. 1 ("The insider buying suggests that these companies' shares are undervalued. Their judgment may be wrong, but analysts who follow insiders believe that the people who are the closest to corporate trends enjoy an edge in deciding when to buy or sell.").

prediction by someone in the know that his company's prospects are dim, then outsiders will sell as well, causing the stock price to decline. The same phenomenon works in reverse with purchases by insiders. Indeed, one study concluded that much of observed insider trading profit is generated by the wave of copycat transactions.[47] Not surprisingly, one economist has suggested that insiders can profit by what he calls "gamesmanship":

> A shrewd insider without information can capitalize on the market's belief in the special knowledge of all insiders by buying shares in his company at any time. Outsiders, learning of this transaction, should bid up the stock price, allowing the insider to sell what he bought at the now higher price. The reverse can occur for selling.[48]

Thus, without violating the provisions of Section 16 (or Rule 10b-5 and common law prohibitions against fraud), insiders have an important advantage over outsiders when it comes to trading stock.

The disclosure requirements of Section 16 have, therefore, not had the effect intended by Congress. Instead of preventing insiders from profiting from their trading activities, it has increased the probability that they will. Such profit is earned in violation of an insider's fiduciary obligations. As described above, an insider has a duty to communicate all material facts to his stockholders, and, indeed, to correct any misimpression about corporate affairs. He is not permitted to profit from confidential information, even if he does not harm his principal in the process. Yet insiders routinely profit from disclosure of their own trades.

C. *Enforcement*

Enforcement of Section 16 has not, by and large, been successful. The Section was designed to be largely self-enforcing, but compliance on the part of company officials has been, at best, inconsistent. In one review of the Official Summary of Security Transactions and Holdings, there were more late filings

[47] Givoly & Palmon, *supra* note 37, at 69.
[48] Jaffe, *Special Information and Insider Trading*, *supra* note 46, at 414.

than timely ones.[49] It appears, as well, that many insiders fail to report their stock transactions at all.[50] According to Securities and Exchange Commissioner Edward Fleischman, "the SEC has found a 'frightening percentage' of seriously delinquent filings."[51]

The only remedy available to the Commission in the event of a Section 16(a) violation is an injunctive action.[52] The Commission cannot levy a fine or bring administrative proceedings. Moreover, it is unlikely that the Commission would invoke the full force of the federal government in an action against someone who has simply failed to file a form on time.

The problem of non-filers is harder to ascertain, and is thus even more severe. The Commission does examine annual reports on Form 10-K for evidence of changes in stock holdings by large shareholders. Although the Commission will not disclose how many annual reports are verified in this manner, if one considers that only major shareholders must disclose ownership on the Form 10-K, that the Commission has only 2,200 employees, and that 7,000 annual reports and 150,000 Section 16 forms are filed every year, one must conclude that the likelihood of catching a non-filer is low.[53]

The Commission's enforcement authority is limited to the reporting requirements of Section 16(a). The right to bring an action to recover short-swing profits under Section 16(b) lies with the corporation and its shareholders.[54] Not surprisingly, however, corporations are often reluctant to bring suit against one of their own officers, directors or major shareholders.

---

[49] Some of the filings may have been marked "late" simply because they were incorrectly filled out. In any event, according to a Securities and Exchange Commission Branch Chief, the percentage of late transactions is unacceptably high. Letter from Daniel Hirsch to S.S. Samuelson (Mar. 3, 1988) (discussing late transactions) (on file at Harvard Journal on Legislation).

[50] Woods, *Insiders Make Mockery of SEC*, Boston Bus. J., Sept. 8, 1986, at 1, col. 2.

[51] *SEC May Require Companies to Certify Executives' Disclosures*, 19 SEC. REG. & L. REP. (BNA) 1247 (Aug. 14, 1987).

[52] The Department of Justice, however, may take action for criminal violations of the federal securities laws.

[53] To compound the Commission's enforcement difficulties, a discrepancy in stock holdings reported on the Form 10-K does not necessarily mean a violation of Section 16 has occurred, since under Rule 16a-9 any transaction in an amount less than $10,000 need not be reported. *See* 17 C.F.R. § 240.16a-9 (1986).

[54] Cook & Feldman, *Insider Trading Under the Securities Exchange Act*, 66 HARV. L. REV. 385, 410 (1953).

Therefore, enforcement has largely been left to shareholders.[55] Although the notion of "private attorneys general" is well-established in other areas of law, such as antitrust, it is typically an addition to, rather than a replacement for, federal enforcement authority. Unless the Commission is granted greater enforcement authority (and the resources to implement it) or unless Section 16(b) is revised to be genuinely self-enforcing, the Act will remain ineffective in combating insider trading.

To sum up, Section 16 has not achieved the goals set for it by Congress. The economic evidence suggests that insiders are still trading on secret information, that they routinely violate common law standards of fiduciary obligation, and that enforcement of Section 16 has been ineffective.

III. A Proposal for Revising Section 16

The problem of regulating insider trading is particularly intricate. The three goals — (1) to stop insiders from trading on secret information, (2) to prevent insiders from profiting when the market follows their trades, and (3) to improve enforcement—have proven thus far to be inconsistent. If, for example, the only goal were to prevent insiders from profiting when the market follows their trades, the solution would be to keep secret all information filed with the Commission under Section 16. This "solution" would, however, make enforcement much more difficult since individual stockholders would be precluded from bringing enforcement actions on their own. All enforcement would, by necessity, be left to the Commission with the result that insider trading on specific secret information would be less effectively deterred than under the current version of Section 16.

In addition, as Professor Brudney has pointed out, there is an important advantage to publicizing corporate information. The faster information relating to the assets or expected perfor-

[55] Consider, for example, the cases brought under Section 16 by Mr. Blau: Blau v. Lehman, 368 U.S. 403 (1962); Blau v. Lamb, 363 F.2d 507 (2d Cir. 1966), *cert. denied*, 385 U.S. 1002 (1967); Blau v. Max Factor & Co., 342 F.2d 304 (9th Cir.), *cert. denied*, 382 U.S. 892 (1965); Blau v. Mission Corp., 212 F.2d 77 (2d Cir.), *cert. denied*, 347 U.S. 1016 (1954); Blau v. Ogsbury, 210 F.2d 426 (2d Cir. 1954); Blau v. Oppenheim, 250 F. Supp. 881 (S.D.N.Y. 1966); Blau v. Allen, 171 F. Supp. 669 (S.D.N.Y. 1959); Blau v. Albert, 157 F. Supp. 816 (S.D.N.Y. 1957); Blau v. Hodgkinson, 100 F. Supp. 361 (S.D.N.Y. 1951); Blau v. Martin, 8 Misc. 2d 541, 67 N.Y.S.2d 662 (N.Y. Sup. Ct. 1957).

mance of an enterprise is found, appraised, and acted upon, the more efficiently the market will function.[56] One could imagine, too, that if information about insider trades were not public, rumors would abound and, to the extent investors found the information valuable, a black market could develop. As a result, the "average" investor, who would not have access to such information or whose moral standards would prohibit him from using it if he did have access, would be at an even greater disadvantage. Another goal should thus be added to our list of objectives: (4) to make information about inside trades available to stockholders and the general public while, at the same time, prohibiting insiders from profiting from this availability.

A ban on equity ownership by company officers and directors would prevent insiders from profiting when the stock market followed their lead and also from trading on secret information. However, it has generally been recognized that equity ownership constitutes a powerful, if not the most powerful, performance incentive for those contemplating either the risky and back-breaking role of entrepreneur or the less risky but equally demanding job of corporate executive. Investors *want* management to be directly rewarded by the enterprise's financial success. If equity ownership by insiders were prohibited, management compensation could be paid at least partly in some form of Stock Appreciation Rights,[57] but then the executive would bear no downside risk (as he does now when he invests a substantial portion of his assets in the company). Nor would he have a choice about what proportion of his personal assets to invest in his company. In addition, under such a proposal, shareholders owning ten percent of a company's stock would escape regulation.

Consider an alternative proposal that would require insiders to report sales or purchases of stock, not after the event, but at least ninety days prior to any trade. The director, officer, or ten percent shareholder would file a form with the Commission reporting the date (at least ninety days in advance) on which he wished to buy or sell a certain number of shares, and the price

---

[56] Brudney, *supra* note 22, at 341. *See also* Carlton & Fischel, *The Regulation of Insider Trading*, 35 STAN. L. REV. 857 (1983). Carlton and Fischel argue that stockholders may value insider trading because it gives the company an additional method of communicating and controlling information that flows to the stock market.

[57] Stock Appreciation Rights are plans under which the manager does not actually own stock, but his compensation is based on the appreciation in stock price.

range in which he was willing to trade (e.g., any price above or below a certain price per share).[58] The Commission would release such information immediately and, at the same time, return to the insider an acknowledgment that the required form had been filed. On the date specified in the form, the insider would be required to buy or sell the specified number of shares, provided that the stock was trading within the range specified.[59] A broker would not be permitted to make trades for anyone known by him to be an insider without first receiving a copy of the Commission's acknowledgment. Any insider who traded in violation of this provision would be required to disgorge to his company three times his profit.[60]

The requirement that an insider announce a trade at least ninety days in advance would greatly reduce his opportunity to trade on secret information. Financial data, approvals of patents, mineral discoveries, litigation advances or setbacks, bond rating changes, or even takeover attempts are rarely kept from the market for ninety days, even by insiders. In addition, if an insider knew of or suspected such an event, his public announcement that he intended to trade would signal the market, thereby reducing his expected profit.[61]

---

[58] The shareholder could specify a range so that if, for example, the stock price had declined dramatically in the ninety-day interim, he would not be forced to sell.

[59] If the insider were not required to trade, he could report that he intended to trade every day and then execute only those he wanted to, thereby dramatically limiting the value of advance insider trading information.

[60] The Insider Trading Sanctions Act of 1984 provides a precedent for treble penalties for insider trading. Pub. L. No. 376, 98 Stat. 1264 (1984).

In the event that an insider traded without having first filed the requisite form with the Commission, he would also be required to disgorge three times the profit realized. If he made a loss on the trade, there could be a minimum penalty of, say, $10,000 for having failed to comply with the law. In the event that he filed the requisite form but did not actually purchase or sell on the date that he had specified, he would be required to purchase or sell that number of shares on the date when the Commission issued an injunction so ordering. If, as a result, he made a greater loss than he would have on the date originally specified, that amount, plus a $10,000 fine would be his loss. If he realized a greater gain than he would have on the date originally specified, he would disgorge to his company three times the difference between the two prices.

[61] This proposal would not apply to involuntary trades—those over which the insider had no control—such as the compulsory sale of stock to a successful white knight after a losing merger battle. One questions whether such transactions ought to be regulated by insider trading statutes at all. They are hardly the kind of manipulative, unfair activities about which Congress was concerned when enacting the Securities Exchange Act.

For a discussion of other complications that have arisen in applying Section 16 to so-called "unorthodox" transactions, see *Report of the Task Force, Part II, supra* note 11, at 1112–23. These "unorthodox" cases include situations where ten percent shareholders convert privately held common stock into publicly held common stock, convert preferred stock into common stock, or trade between wholly-owned entities. All of the

This proposal would work as follows. If, for example, a company's stock was trading at $25 per share, and an insider suspected a tender offer was likely, he could file a notice stating that in ninety days he intends to buy 1,000 shares at any price up to $25. He would not specify a higher price for two reasons: if the tender offer was not subsequently made he would not want to buy at a price higher than $25; or if the information was already public by the date of his trade and the price of the stock had increased, he would not be able to sell later for a profit. If he announced that he intended to buy and then promptly sell the stock, he would be revealing that he was trading on inside information. If news about the tender offer did become public before the specified trade date and the stock price rose higher than $25, the insider would neither be permitted nor required to purchase stock.

If the possibility that a tender offer would be made had not seeped into the marketplace by the trade date but was still a possibility, the insider would get to buy at $25 and might still profit when the offer became public, but the announcement of his intent to buy stock would have signaled the market as to his expectations. This signal might cause the market price to rise above his specified $25 per share purchase price, in which case he would not be able to purchase the stock. If the tender offer had been aborted, and the stock price drifted lower than $25, the insider would be required to buy 1,000 shares at market price (since he had specified that he would buy stock at any price up to $25). The net effect of this proposal would be to discourage insiders from timing trades in order to benefit from secret information. A more profitable strategy for an insider would thus be to make investment decisions based on his long-term evaluation of the company's prospects rather than hoping to time trades to benefit from secret information.

This proposal would not just discourage trading on secret information; it would also reduce insiders' profits from their role as leading indicators. The market would, after all, have ninety days to respond to the news of an intended trade. Is ninety days enough? Too much? It is seemingly as arbitrary a period as the six-month period in the current Section 16.[62] However, eco-

hair-splitting analysis applied in such cases could be avoided under the abovementioned proposal as long as ninety days notice is given.

[62] *See* H. MANNE, INSIDER TRADING AND THE STOCK MARKET, *supra* note 35, at 162: "[w]e should notice just how arbitrary this provision can be in operation. First, the six-month duration bears little relation to anything connected with insider trading."

nomic evidence indicates that the market generally responds to insider trades within six months, with the largest response being in the first month.[63] Thus, if the only concern were to prevent insiders from profiting when the market followed their trades, thirty days might be satisfactory. However, thirty days would not sufficiently deter trading on specific secret information. Well-placed insiders could, for example, often know about an unusual dividend announcement or unexpected earnings more than thirty days before such information became public. To require insiders to determine trades six months in advance is too onerous; ninety days gives the market sufficient protection without unduly burdening company officials.

One could argue that it is not fair to require insiders to announce purchases and sales in advance but then wait until after the market has responded before being permitted to trade. However, such a requirement would be consistent with the rules governing fiduciary relationships.[64] It is, after all, the fiduciary's obligation to inform beneficiaries of all material facts. The market has indicated that information about insider trading is material.[65] Under the proposal advanced in this Essay, outside investors would not be placed at a disadvantage by having to trade on less information than insiders. Congress, after all, intended when enacting Section 16 that insiders not profit at the expense of their investors. Directors, officers, and large shareholders would be put in a position similar to the market. This proposal thus approaches the attainment of that oft-sought but yet-to-be-acheived goal: the level playing field.[66] Furthermore, the securities markets would, as a result, be not only fairer but also more efficient. Information about insider trades would get into the market faster—before the trades rather than forty days or longer afterwards as is now the case.[67]

Undoubtedly, there will always be some slippage in the enforcement of insider trading regulations. The vast volume of trading virtually insures it. However, increasing the penalty to

---

[63] Finnerty, *supra* note 37, at 1146.

[64] *See supra* notes 23–30 and accompanying text. Professor Brudney points out that, "[a] fiduciary duty can be violated in some circumstances merely by the fiduciary engaging in transactions without the informed consent of the beneficiaries. . . ." Brudney, *supra* note 22, at 351.

[65] *See supra* note 46.

[66] Some precedent for advance notice of sales can be found in Rule 144 which requires advance notice to the Commission on some securities sales. *See* 7 C.F.R. § 230.144 (1986).

[67] *See supra* note 3.

three times the profit realized would discourage violations. In addition, requiring brokers to be in possession of a Commission acknowledgment before trading would have a deterrent effect. For those who were not seeking to violate the law, compliance would be routine: upon receipt from the Commission of an acknowledgment that the requisite form had been filed, the insider would simply send the acknowledgment to his broker with instructions to carry out the specified trades. In opening a new brokerage account, among other information requested, the broker would ask the new customer to list the companies for which he was an insider.

Under the current system, insiders who violate the law by failing to file the requisite forms when due, are not, in the vast majority of cases, subjected to any sanctions. Indeed, virtually the only instance in which an executive is sanctioned (either by the Commission or by shareholders) is when he has voluntarily reported his own misdeed by filing Form 4.[68] Under the proposal advanced in this Essay, however, an insider who failed to make the requisite filing would receive the most appropriate sanction of all. He would be unable to trade the stock. Brokers who repeatedly failed to comply with these regulations would lose their licenses, a risk hardly commensurate with the potential gain to them.

The Commission could continue to monitor insider trades much as it now does, but at least part of the enforcement burden would be shifted to brokers, who are in a much better position to know when the law is violated. Indeed, brokers are already responsible for monitoring other federal securities regulations, such as margin regulations, for example.

Although it might be difficult for a broker to determine whether his customer is an insider under Section 16, it would be relatively easy for him to ascertain whether or not his client was a director, employee, or ten percent shareholder. If the client argued that he was an employee of the company but not a statutory insider, it would behoove the broker to obtain such an opinion in writing from company counsel.

Although insiders can not be expected to be enthusiastic about a statute that would encumber their ability to trade stock of

---

[68] Form 4 is filed pursuant to 17 C.F.R. § 240.16a-1(a) (1986). The purpose of this filing is to determine and disclose the holdings of officers, directors, and beneficial owners of registered companies.

their own company, a close examination of this proposal reveals at least one benefit to insiders. Under Section 16 as currently written, the exercise of an option is considered to be a purchase of a security.[69] As a result, if an insider sells stock within six months of the exercise of an option, he must disgorge profits on the transaction.[70] Under the proposal advanced here, he could announce ninety days in advance that he intended to exercise his option and sell part or all of the stock immediately. An insider without sufficient funds to exercise an option would, under this proposal, be able to raise cash by selling immediately part of the stock received upon exercise.

Despite the advantages of this proposal, the result of its enactment should be to reduce the profit of insiders trading stock of their own companies. The appropriate response to this problem (if it is indeed a problem) is to find other, more equitable and efficient methods of compensation. After all, it would make more sense to base an officer or director's compensation on his contribution to the company than on his skill in manipulating the market. Congress recognized fifty years ago that the trading of company securities on inside information was not good for companies or the American financial markets.[71] Likewise, to the extent that it is desirable to encourage venture capitalists and entrepreneurs,[72] the return to the investor is more appropriately based on long-term trends in the stock rather than on his ability to fine tune his purchases and sales.[73]

---

[69] 17 C.F.R. § 16a-6 (1986).

[70] Under 17 C.F.R. § 240.16b-6 (1986), the profits to be disgorged by the insider are limited to the difference between the proceeds of the sale and the lowest market price of any security of the same class within six months before or after the date of sale. The effect of this regulation is that the insider is liable only for appreciation occurring during the short-swing period and not for appreciation up to the date of exercise.

[71] *See supra* text accompanying note 21.

[72] Commentators who argue against the prohibition of insider trading tend to base their arguments, in part, on the belief that such prohibitions result in a less hospitable business environment. *See, e.g.*, Carlton & Fischel, *supra* note 56, at 870–71, 892.

[73] It has been assumed throughout this Essay that trading on inside information ought to be eliminated to the extent possible. That was, after all, the assumption of Congress when enacting Section 16. It is not, however, an assumption shared by all. For a discussion of the evils of insider trading, see Committee on Federal Regulation of Securities, *Report of the Task Force on Regulation of Insider Trading, Part I: Regulation Under the Antifraud Provisions of the Securities Exchange Act of 1934*, 41 BUS. LAW. 218, 228–29 (1985); L. LOSS, THE FUNDAMENTALS OF SECURITIES REGULATION 541–47 (1988); Scott, *Insider Trading: Rule 10b-5, Disclosure, and Corporate Privacy*, 9 J. LEGAL STUD. 801 (1980); Rubin & Feldman, *supra* note 14, at 500–01.

For a discussion of why insider trading ought not to be regulated, see H. MANNE, *supra* note 35, at viii; Carlton & Fischel, *supra* note 56; Wu, *supra* note 46; Dooley, *Enforcement of Insider Trading Restrictions*, 66 VA. L. REV. 1 (1980).

IV. CONCLUSION

In 1934, Congress, stunned by the economic disasters that had befallen the United States, passed Section 16 of the Securities Exchange Act of 1934. The intent of the statute was clear and specific: to prevent insiders from trading on inside information and to require corporate officers to act as fiduciaries for their corporations. However, while imposing a real burden of compliance on insiders, Section 16 has not been effective in obtaining Congressional goals.

The framework Congress created in 1934 for regulating insider trading was larger and more robust than the shrunken and damaged securities market that had survived the 1929 Crash. Now, however, in an era of sophisticated communications and securities exchanges on which the sun never sets, the financial markets have outgrown the framework created by Congress. The time has come to create a new, enlarged system rather than just to patch up the old one.

The proposal introduced in this Essay has broad implications for the regulation of futures contracts, options and tender offers, as well as other areas of securities law. An exhaustive discussion of these issues is beyond the scope of this Essay. Rather, the purpose of this introduction is to consider the nature of the fiduciary duty insiders owe shareholders, and to question whether current law provides an adequate framework for insuring that insiders discharge their responsibilities properly. The proposal considers a new approach toward achieving the elusive goal of a "level playing field" in the securities markets.

APPENDIX

Economic Studies of Insider Trading: Excess Returns and Timing

In the last fifteen years, a growing number of economic studies have examined insider trading addressing two principal questions: (1) Do insiders earn excess returns in their security transactions? and (2) Is the type and timing of insider trading related to public announcements concerning the performance of the company whose security is traded? An affirmative answer to

both of these questions constitutes evidence of profitable insider trading based on private information.

An answer to the first question requires care in identifying "excess" returns. The main foundation of modern finance theory holds that a security's expected return increases linearly with its risk (properly defined).[74] Loosely speaking, higher returns can be obtained only by assuming greater risks. The earliest studies documented high returns from insider trading but did not account for the associated risks. To begin with, the capital asset pricing model (CAPM) has been used to compute excess returns. In its most frequently used form, the CAPM predicts a security's expected return (r^*) according to

$$r^* = r_f + \beta(r_m - r_f),$$

where r_f is the return on a risk free asset, r_m is the return on the market portfolio, and β is the measure of correlation between the return of the security and the return of the market. The greater the security's "risk" (β), the greater its expected return. A statistical analysis of the security's past return behavior is used to estimate β. Then the above equation is used to predict the security's normal return r^*. In a given time period, the security is found to earn an excess return if its actual return exceeds r^*. Finally, the past record of returns obtained by insider trading is analyzed to determine whether statistically the returns are significantly different from normal returns.

Without exception, the principal studies have found that insider trades generate statistically (and economically) significant excess returns after accounting for security risks.[75] Stocks purchased by insiders earned higher returns than would be expected given their associated risks. Stocks sold by insiders subsequently earned lower returns than expected. (In the latter case, the insider who originally sold the security earns an excess return by repurchasing it after the fall in price.) For example, Jaffe[76] and Givoly and Palmon[77] find average excess returns over eight month periods following insider trades of five percent and

---

[74] *See, e.g.*, B. MALKIEL, A RANDOM WALK DOWN WALL STREET (4th ed. 1985).

[75] *See* Elliott, Morse & Richardson, *supra* note 37; Finnerty, *supra* note 37; Jaffe, *supra* note 46.

[76] Jaffe, *Special Information and Insider Trading*, *supra* note 46, at 421.

[77] Givoly & Palmon, *supra* note 37, at 86.

eight percent respectively. Finnerty,[78] Penman,[79] and Elliott, Morse, and Richardson[80] also find significant excess returns for insider purchases and sales.

The second question is whether superior private information is the basis for insiders' profitable trades. An answer can be obtained by examining the timing of insider trading relative to public announcements of information (earnings forecasts and the like) pertinent to company performance. Insider security purchases in the period prior to announcements of "good" news and insider sales prior to "bad" news would constitute evidence of trading based on privately held information.

On the timing issue, recent economic studies have detected some partial evidence of insider trading based on private information. Penman examined insider trading prior to company announcements of earnings forecasts and found evidence of trades in the expected direction. Good news announcements (those associated with the greatest appreciation in stock price) were preceded by insider purchases of shares; bad news announcements were preceded by net selling.[81] Elliott, Morse, and Richardson[82] examined insider trading surrounding announcements of annual earnings, dividend changes, bond rating changes, mergers, and bankruptcies for the period 1975–1979 and found results weakly consistent with some use of private information. Givoly and Palmon,[83] reviewing insider trading and public announcements during 1973–1976, also found significant excess returns to insider trading during eventless (no news) periods. They concluded that outsiders follow the footsteps of insiders (make subsequent trades in the same direction) affording the latter excess returns even without an informational advantage.[84]

---

[78] Finnerty, *supra* note 37, at 1147.
[79] Penman, *supra* note 37, at 502.
[80] Elliott, Morse & Richardson, *supra* note 37.
[81] Penman, *supra* note 37, at 488.
[82] Elliott, Morse & Richardson, *supra* note 37.
[83] Givoly & Palmon, *supra* note 37, at 83.
[84] *Id.* at 85.

SYMPOSIUM

IS THE GRAMM-RUDMAN-HOLLINGS ACT AN EXERCISE IN LEGISLATIVE FUTILITY?

Throughout the past decade, the large and growing federal budget deficit has emerged as one of the most important economic problems facing the United States. Economists may disagree on the ultimate significance of an unbalanced budget, but the deficit's profound impact on national economic policy and the way budget decisions are made is unquestionable.

With recent deficits at times exceeding $200 billion, the public has become increasingly apprehensive about the government's ability to control the federal deficit. Congress itself acknowledged that it could not reduce the shortfall within the existing budget framework when it passed the Gramm-Rudman-Hollings Act in 1985. Designed to achieve a balanced budget over a period of several years, the Act features an automatic sequestration procedure whereby revenue and spending cuts are imposed automatically if Congress fails to establish a budget that meets the targeted schedule. The cuts are intended to occur in a budget-neutral and politically-neutral way, but they are specifically designed to be only an undesirable last resort.

The reporting process established by the Act for implementing sequestration cuts was declared by the Supreme Court in 1986 to be a violation of the principle of separation of powers. To rectify this defect, Congress in 1987 rewrote several of the Act's provisions in the Reaffirmation Act. The Reaffirmation Act revises the reporting process and redefines the deficit targets.

Gramm-Rudman-Hollings has been hailed by some observers as a landmark initiative and an effective tool for budget control. Others have criticized it as an empty gesture, easily sidestepped and offering little to address the underlying causes of budget increases. Is the Act an exercise in legislative futility? The effectiveness of Gramm-Rudman-Hollings and the Reaffirmation Act in reducing the deficit, and the political process behind their passage, are the subjects of this Symposium.

In his Essay, Professor John Ellwood addresses the question by examining the incentive structure that underlies Congressional decision-making. He observes the motivations behind the deficit-reduction process and sees a legislature motivated by the often conflicting goals of maximizing prospects for re-election,

enhancing institutional power, and furthering substantive policy objectives. The resulting dynamic has obstructed budget reduction efforts in the past, and has threatened to undermine Gramm-Rudman-Hollings. The Act that emerged from this process, he argues, did not alter it. Instead, the Act merely tries to limit the effects of an incentive structure that it leaves largely unchanged.

Senator Pete Domenici (R-N.M.), on the other hand, argues that the Gramm-Rudman-Hollings Act has succeeded in increasing the perceived importance of the federal budget deficit both in Congress and among the public. The Act therefore is not an act of legislative futility. Although he is sensitive to the limitations in its process-oriented approach, he points out that the nominal deficit has fallen over time and asserts that there have been great changes in the way that Congress approaches budget reform.

Representative Thomas Downey (D-N.Y.) disagrees. The Act has not affected the way that Congress addresses the deficit, he asserts, and it will have little lasting impact. Congress has continued to evade its own budget deadlines, enacting enormous spending bills and engaging in accounting manipulation to comply only technically with Gramm-Rudman-Hollings' requirements. Congress therefore has not been forced to make the difficult policy choices that the Act is supposed to force, and the legislature has instead "fudged, ignored, and then eased the Act's deficit cutting requirements."

Senator Robert Kasten (R-Wisc.) believes that Gramm-Rudman-Hollings is more than just a "fiscal fig leaf." He asserts that the Act has succeeded in imposing a measure of fiscal discipline on Congress that was absent before its passage. The Act has slowed the growth of federal spending and increased political accountability for deficit-increasing measures. Senator Kasten concedes that the process is still subject to manipulation and that the 1987 budget compromise failed to reduce spending significantly; the process still needs improvement. Nevertheless, he says, the Gramm-Rudman-Hollings Act can provide an effective framework for achieving deficit reduction in coming years.

The belief that the Act has forced Congress to make the hard political choices needed for successful deficit reduction is shared by Senator Dan Quayle (R-Ind.). He asserts that the Act's effectiveness derives from the fact that its automatic sequester requirement creates a political scenario worse than any that responsible budget planning would create. Furthermore, the

Act's automatic budget-cutting provisions make it easier for legislators to withstand criticism from those opposing needed spending reductions or tax increases, thus making hard choices easier.

Former Pennsylvania Governor Richard Thornburgh sees an analogue of today's budget deficit debate in the efforts during the 1830's to reduce a worrisome government *surplus* over a period of years. Deficit reduction has since proved far more difficult to achieve, and Mr. Thornburgh does not believe that Gramm-Rudman-Hollings can accomplish it. Instead, he argues, a balanced budget amendment to the United States Constitution is needed. In his Essay, he responds to five common objections to the proposal, concluding that such an amendment would be an effective and flexible solution.

Finally, Professor Randall Strahan addresses Gramm-Rudman-Hollings within a larger context, examining the broader implications of the Act's procedural mandates. Professor Strahan takes the position that in passing the Act, Congress has resorted to semi-autonomous procedures to establish taxing and spending levels because the political leadership has failied to define acceptably the proper role of government in society. He traces what he sees as the decline in the 1980's of a liberal public philosophy calling for an interventionist federal government participating actively in the management of the economy. He then observes that the economic conservatism of the Reagan era has failed to replace this view of the government's proper role in the public mind. As a result, deficit politics—perhaps the most significant Congressional economic enterprise of the era—has been aimless and ineffective. A crippling weakness in the Gramm-Rudman-Hollings Act, he concludes, is its failure to synthesize a grand consensus for deficit reduction.

The Gramm-Rudman-Hollings Act is still of recent vintage, and its ultimate success or failure remains to be seen. The authors of this Symposium approach the issues it raises from diverse backgrounds in government and academe. The views expressed are their own, and it is hoped that their insights will contribute to a better understanding of the budgetary process and the operation of the federal government.

—H. Bradley Southern
Policy Essay Editor

THE GRAMM-RUDMAN-HOLLINGS BUDGET PROCESS: AN ACT IN LEGISLATIVE FUTILITY?

PETE V. DOMENICI*

The simple answer to the "futility" of the Gramm-Rudman-Hollings Act[1] is: No, the Act was not a futile exercise at all. Gramm-Rudman-Hollings' record is a good one; not perfect, but good.

Declaring the success or failure of any legislation, especially major budget legislation, is of course very difficult. Times and budgets change rapidly. The objective measures of success are subject to rapid recalibration. And, realistically, the two years during which Gramm-Rudman-Hollings has been in effect is an exceedingly brief period in which to evaluate the impact of such major policy legislation.

Nonetheless, we can see immediately that Gramm-Rudman-Hollings was not a futile act by simply observing that the result its authors were looking for—namely a reduction in the federal deficit—was accomplished. Technicians will argue over "baselines and economic assumptions," but Congressional Budget Office's deficit projection numbers demonstrate that the current outlook for deficit reduction is significantly better than it was when Gramm-Rudman-Hollings was initially debated in the fall of 1985:

---

* Member, United States Senate (R-N.M.). B.S., University of New Mexico, 1954; LL.B, Denver University, 1958. Senator Domenici is a member of the Senate Appropriations Committee and is the former chairman and current member of the Senate Budget Committee.

[1] The Balanced Budget and Emergency Deficit Control Act of 1985, Pub. L. No. 99-177, 99 Stat. 1037 [hereinafter Gramm-Rudman-Hollings], *amended by* The Balanced Budget and Emergency Deficit Control Reaffirmation Act of 1987, Pub. L. No. 100-119, 101 Stat. 754 [hereinafter Reaffirmation Act].

TABLE

CONGRESSIONAL BUDGET OFFICE DEFICIT PROJECTIONS
FROM AUGUST 1985 & FEBRUARY 1988
FOR FISCAL YEARS 1986 THROUGH 1991
(Including off-budget entities)
(in billions of dollars)

	1986	1987	1988	1989	1990	1991
August, 1985[2]	212	229	243	264	285	na
February, 1988[3]	221[4]	150	157	176	167	158
GRH 1985 Target[5]	172	144	108	72	36	0
GRH 1987 Target[6]	na	na	144[7]	136	100	64[8]

Whether the federal deficit is measured in terms of projected deficits or other measurements, such as overall budget growth, it is clear that the period 1985 to 1987 produced a significant slowdown in the growth of federal spending. Between 1980 and 1984, federal spending increased at an annual rate of 9.6 percent.[9] Between 1985 and 1987, the rate of spending growth fell to 3 percent annually.[10] And in fiscal year 1987, the first full year of Gramm-Rudman-Hollings, federal spending grew by less than 1.4 percent.[11]

[2] CONGRESSIONAL BUDGET OFFICE, REPORT TO SENATE AND HOUSE COMMITTEES ON THE BUDGET: THE ECONOMIC AND BUDGET OUTLOOK & UPDATE xxi (Aug. 1985).

[3] CONGRESSIONAL BUDGET OFFICE, REPORT TO SENATE AND HOUSE COMMITTEES ON THE BUDGET: THE ECONOMIC AND BUDGET OUTLOOK—FISCAL YEARS 1989–1993 51 (1988).

[4] Actual deficit for fiscal year 1986. The maximum deficit reduction permitted in fiscal year 1986 from the across-the-board cuts was $11.7 billion, half in defense outlays and half in nondefense outlays. The $11.7 billion in cuts was agreed upon in the Omnibus Budget Reconciliation Act of 1986, Pub. L. No 99-509, 100 Stat. 1874.

[5] These are the deficit targets established by the Gramm-Rudmann-Hollings Act. *See supra* note 1.

[6] These adjusted targets were established in the Reaffirmation Act. *See supra* note 1.

[7] The actual deficit target in fiscal years 1988 and 1989 is defined from a deficit reduction baseline. Thus, the actual target figure will differ from the absolute figures shown in this table. For fiscal year 1988 a minimum of $23 billion in net deficit reduction was required by the Reaffirmation Act. For fiscal year 1989, the Reaffirmation Act requires $36 billion in deficit reduction, or the amount necessary to reach the fixed target of $136 billion.

[8] The fixed target for fiscal year 1992 is $28 billion with balance to be reached in fiscal year 1993 according to the Reaffirmation Act.

[9] Senate Budget Committee, Worksheet: Growth of Federal Spending (prepared by Committee staff using Congressional Budget Office data) (on file with the Minority Staff of the Senate Budget Committee).

[10] *Id.*

[11] *Id.* This was the case even after Gramm-Rudman-Hollings' sequestration procedure was found constitutionally invalid in Bowsher v. Synar, 106 S. Ct. 3181 (1986). The

The Act's record of success is no modest accomplishment. Senator Gramm pointed out some of the reasons for this success on the Act's first anniversary:

> What the Balanced Budget and Emergency Deficit Control Act did was to set out targets. It forced the President to submit budgets. It forced us to deal with those targets, and it set out some binding constraints. It did not yield a solution. But rather than the whole process of the Senate being biased against controlling spending, Gramm-Rudman-Hollings sought to bias the process in favor of controlling spending. . . . I believe that we have tilted the process toward controlling spending and fiscal responsibility.[12]

Still today, the adequacy of the Gramm-Rudman-Hollings solution to the federal deficit problem is questioned. The requirements of Gramm-Rudman-Hollings were designed to control the budget *process* to achieve deficit reduction. While I have been a strong advocate for budget process reform, I never believed that process could replace hard legislative decisions. Now, as in the past, both elected officials and the public make the budget *process*, and thus Gramm-Rudman-Hollings, the culprit for Congressional failures in demonstrating fiscal responsibility in budgetary decisions.

Additionally, a plurality in Congress presently considers Gramm-Rudman-Hollings to be inadequate. A recent survey of both House and Senate members discovered mixed feelings regarding the impact of Gramm-Rudman-Hollings on the legislative and budget process. Overall, 46.5 percent of the Members surveyed believed that Gramm-Rudman-Hollings had not improved the process, 34.2 percent believed that it had, 4.4 percent said that the effect was mixed, while 11.4 percent of the Members said that Gramm-Rudman-Hollings had little effect on the process, and 3.5 percent had no opinion.[13]

Act's sequestration procedure required the Comptroller General to make automatic across-the-board budget cuts in the event that Congress failed to provide a budget meeting the deficit reduction targets required by the Act. The Supreme Court held that in giving a congressional officer, the Comptroller General, a budgetary "veto power," the Act's sequestration procedure violated the Constitutional requirement of separation of powers. The Court held that such a budgetary veto power was allowable only for the executive branch.

[12] 132 CONG. REC. S16940 (daily ed. Oct. 17, 1986) (statement of Sen. Gramm (R-Tex.)).

[13] CENTER FOR RESPONSIVE POLITICS, CONGRESSIONAL OPERATIONS: CONGRESS SPEAKS—A SURVEY OF THE 100TH CONGRESS 26–27 (1988) (on file at the Harvard Journal on Legislation).

Another major problem with the Act was revealed when its sequestration procedure was declared unconstitutional by the Supreme Court.[14] The sequestration process was restored as one of the amendments under the Reaffirmation Act in a form that should be constitutionally acceptable. Currently, under the Reafirmation Act, automatic sequestration authority has been moved to the Office of Management and Budget, an agency located within the executive branch.

Finally, some of the deficit reduction success shown in the above table cannot be ascribed entirely to the direct action of Gramm-Rudman-Hollings. For example, a one-time military payday shift into the fiscal year 1988 budget "saved" $2 billion in the fiscal year 1987 budget, and physical and loan portfolio asset sales contributed an additional $8 billion toward the 1987 deficit reduction. Absent Gramm-Rudman-Hollings, however, Congress may not have made these asset sales and payment shifts.

Do these difficulties make Gramm-Rudman-Hollings an act of legislative futility? Again, I argue they do not. Understanding Gramm-Rudman-Hollings' successes and failures, however, requires an appreciation of how and why the Act was created.

Gramm-Rudman-Hollings is a diverse and complex piece of legislation. The original Gramm-Rudman-Hollings legislation included a number of budget process changes that had been recommended by the House Task Force on the Budget Process.[15] These modifications included providing a single budget resolution instead of two; establishing a new budget process timetable; extending the annual budget act to cover credit authority; and including in the budget a variety of programs at the time not computed in the basic budget.

The best-known features of Gramm-Rudman-Hollings are the specific deficit targets and the mechanisms, such as the sequester, established to attain the targets. The first mechanism for meeting the deficit targets shifted the responsibility for budget cuts directly to the Congress. Following specified procedures, Congress would manually trim the budget to meet deficit targets calculated by the Congressional Budget Office.

---

[14] *See supra* note 11.

[15] This task force, commonly called the Beilenson Task Force for its chair, Representative Tony Beilenson (D-Cal.), was created by the Committee on the Rules in the 98th Congress to suggest budget reform.

The second technique involved automatic, across-the-board cuts through what was called a sequester. As mentioned above, this process was later found to be unconstitutional.[16] The third and most recent mechanism is the reassertion of the sequestration process in a form we believe to be constitutional. This newest mechanism also realistically shifts overall budget-setting policy back to the executive branch.[17]

How has Congress reached this position of ensuring both a high level of Congressional involvement in the budget process and ultimate executive branch sequestration power? Congress reached it following a twisted and convoluted route which brought the budget to executive branch control prior to 1974, to legislative branch control thereafter, and then, with the Reafirmation Act in 1987, back to executive branch control. An understanding of the Act's history indicates why Gramm-Rudman-Hollings was not an act of legislative futility.

In 1974, Congress passed the Congressional Budget and Impoundment Act ("1974 Act") to strengthen the role of Congress in fiscal policy.[18] This Act established the Congressional budget committees and the process through which Congress was to develop budget proposals. However, Congress regularly missed budget deadlines and a massive budget deficit continued to grow. The 1974 Act appeared inadequate to force Congress to make the necessary hard choices required to reduce the red ink. The 1985 Gramm-Rudman-Hollings Act grew out of Congress' considerable frustration with the budget process created by the 1974 legislation.

Congress' general feeling of frustration was exacerbated by the need to set a new federal debt ceiling—one of horrific proportions—namely *$2.079 trillion.*[19] This new ceiling pushed the borrowing authority of the government above $2 trillion for the first time from an already incredible $1.8 trillion.[20] However, perhaps the greatest impetus for creation of the Gramm-Rudman-Hollings Act came from the fact that Congress was forced

---

[16] *See supra* note 11.

[17] The amendment was made in The Reaffirmation Act, *supra* note 1. Automatic sequestration power was shifted to the Office of Management and Budget, an agency located within the executive branch.

[18] The Congressional Budget and Impoundment Control Act of 1974, Pub. L. No. 93-344, 88 Stat. 297.

[19] *See supra* note 3 at 51.

[20] *Id.*

to breach the $2 trillion barrier at a time when no fundamental changes were being considered to bring the deficit down.

In general Congress has been institutionally incapable of the discipline needed to make budget choices. This feeling was only fueled by the frustration in 1985 of a year-long debate on the fiscal year 1986 budget resolution. The fiscal year 1986 budget package had been based on the assumption that the deficit would decline to $172 billion for fiscal year 1986. By autumn of 1985, however, it was clear that Congress was unable to make the decisions to enact legislation that would make that figure a reality.[21]

It was in this atmosphere of congressional frustration with the old, failed budget process and Congress' inability to control the federal deficit that a majority of Members voted for a mechanical tool to force hard congressional choices.

The 1974 Act, which Gramm-Rudman-Hollings replaced, did not attempt to promote any particular fiscal policy or set of spending priorities:

> From the outset, the main purpose of budget reform was to establish fiscal responsibility in Congress. However in the course of developing the Budget Act [1974 Act], Congress transformed the concept of responsibility from a substantive to a procedural test. . . . The first operative paragraph of the Joint Study Committee's report opened with an indictment of recurring and growing budget deficits. The JCS bill, would have imposed strict controls on legislative spending action, *but these controls were relaxed in favor of a process permitting Congress to act according to its preferences. . . . Congress could take whatever tax or spending action it deemed appropriate, provided that it acted in accordance with the procedure laid out in the Budget Act* [1974 Act].[22]

In contrast, Gramm-Rudman-Hollings was designed specifically to establish a particular procedure within the budget process to force a particular outcome: a balanced budget by fiscal year 1991.

---

[21] In the Senate that year Majority Leader Senator Bob Dole (R-Kan.) and I had struggled to pass a major deficit reduction package, S. Con. Res. 32. On May 10, 1986 our budget package finally passed the Senate on a tie vote, which was broken by the Vice President, only to be later defeated in the House of Representatives.

With hindsight, fiscal year 1986 turned out to produce the worst one-year deficit we ever recorded—$221 billion. That was the total figure after $11.7 billion was removed by the one and only Gramm-Rudman-Hollings sequester.

[22] A. SCHICK, CONGRESS AND MONEY: BUDGETING, SPENDING, AND TAXING 77–78 (1980) (emphasis supplied).

The Reaffirmation Act in 1987 restored the sequester and pushed the target dates for a balanced federal budget from the original goal of 1991 to 1993.[23] However, the Reaffirmation Act was designed not only to produce a specified level of deficit reduction for the 1988 fiscal year—$23 billion—but also effectively to produce it according to the general blueprint of the Democrats' spring budget proposal.[24]

The second full year of Gramm-Rudman-Hollings, during which the fiscal year 1989 budget has been prepared, was decidedly less successful than the first year. The Democratically-controlled 100th Congress found itself confronting major deficit reduction requirements in order to achieve the fiscal year 1988 deficit goal of $108 billion.[25] Using the deficit estimates of the Congressional Budget Office, a deficit reduction of slightly over $61 billion would have been required.

Perhaps one of the great ironies of public policy is that we will never know whether such cuts would have taken place in the fiscal year 1988 budget—$11.5 billion in defense (a 10.4 percent reduction) and $11.5 billion in nondefense (an 8.7 percent reduction) with Gramm-Rudman-Hollings as the sole motivation. It was the October 19, 1987 stock market crash that forced the Administration and Congress to come together at the budget table. The public urgency created in the securities markets, *not the threat of the across-the-board cuts of Gramm-Rudman-Hollings*, forced this issue.

After thirteen years of close involvement with the budget process, it appears that Congress has substantially changed the thrust of this process. At the same time, the congressional role has changed from neutral bystander to active participation for spending restraint. What is the outlook? I believe that the Congress and the President are once again on a glidepath toward a balanced budget. But possibly the most important effect of Gramm-Rudman-Hollings has been that it has elevated the importance of the deficit in the public consciousness.

---

[23] *See supra* note 8.

[24] I did not support the Reaffirmation Act. I supported Gramm-Rudman-Hollings, but the Reaffirmation Act went much further in attempting to force a Democratic budget onto a reluctant Administration using significant and dangerously destabilizing cuts in our national security as a threatening alternative.

[25] This is the figure set by Gramm-Rudman-Hollings. This figure and the adjusted figure of $144 billion under the 1987 Reaffirmation Act are included in the table *supra* page 2.

As its authors planned, Gramm-Rudman-Hollings has made the Congress, the President, and the public begin to focus more on holding down spending and on setting priorities.

Now that Gramm-Rudman-Hollings is with us, its powerful sequestration process revived by the Reaffirmation Act, we will never escape its shadow. This result hardly suggests legislative futility.

THE FUTILITY OF GRAMM-RUDMAN-HOLLINGS

THOMAS J. DOWNEY*

If someone were to come to you, after devoting an incredible amount of time and energy to solving a pressing concern, and comment that his or her solution is a "bad idea whose time has come," you would naturally be suspicious.[1] Struck by the ambivalence of the comments, you would question both the credibility of the solution and the commitment of your acquaintance to carry it out.

This is exactly what happened when Congress enacted The Balanced Budget and Emergency Deficit Control Act of 1985 ("Gramm-Rudman-Hollings").[2] During its final consideration late one night in December 1985, congressman after congressman, whether opposed to the bill or not, spoke with misgivings about it. It was characterized as having "many faults,"[3] as a "sad reflection on [Congress],"[4] and as a "mindless, unfeeling, unthinking, bloodless formula."[5]

Since the passage of the original Gramm-Rudman-Hollings, the Supreme Court acted to correct one of the "faults," holding the sequestration process created by the Act unconstitutional.[6] Congress has in effect recognized many other faults, making fundamental amendments to the deficit reduction targets and deadlines set by the original Gramm-Rudman-Hollings with last

* Member, United States House of Representatives (D-N.Y.). B.A., Cornell University School of Industrial and Labor Relations, 1970; St. John's University Law School, 1972–74; J.D., American University Law School, 1979. Representative Downey is the acting Chairman of the Public Assistance and Unemployment Compensation Subcommittee of the House Ways and Means Committee.

[1] Statement of Senator Phil Gramm (R-Tex.), CONGRESSIONAL QUARTERLY, CONGRESSIONAL QUARTERLY ALMANAC Vol. 41 at 459 (1985).

[2] The Balanced Budget and Emergency Deficit Control Act of 1985, Pub. L. No. 99-177, 99 Stat. 1037 [hereinafter Gramm-Rudman-Hollings].

[3] 131 CONG. REC. H11877 (daily ed. Dec. 11, 1985) (statement of Rep. Rostenkowski (D-Ill.)).

[4] 131 CONG. REC. H11881 (daily ed. Dec. 11, 1985) (statement of Rep. Daub (R-Neb.)).

[5] 131 CONG. REC. H11885 (daily ed. Dec. 11, 1985) (statement of Rep. Hyde (R-Ill.)).

[6] The sequestration process in Gramm-Rudman-Hollings required the Comptroller General of the United States to make automatic across-the-board budget cuts if Congress did not provide a budget meeting the deficit reduction targets required by the Act. In Bowsher v. Synar, 106 S. Ct. 3181 (1986), the Supreme Court held that in giving a Congressional officer, the Comptroller General, a budgetary "veto power," the Gramm-Rudman-Hollings sequestration process violated the separation of powers. The Court held that such a budgetary "veto power" was allowable only for the Executive branch.

summer's Balanced Budget and Emergency Deficit Control Reaffirmation Act of 1987 ("Reaffirmation Act").[7]

As members of a body that acts by consensus, it is not unusual for representatives and senators to voice reservations about legislation they pass. However, because Gramm-Rudman-Hollings affects the operation of the Congress itself, the equivocation which surrounded its passage was foreshadowing of the Act's futility.

Gramm-Rudman-Hollings was designed to lower and eliminate the Federal budget deficit. However, to the representatives and senators who participated in the massive conference committee that originally shaped Gramm-Rudman-Hollings, and to those who amended it with the September 1987 Reaffirmation Act, the real question of Gramm-Rudman-Hollings was *how* to go about accomplishing the deficit-elimination goal.

The specter of Gramm-Rudman-Hollings' automatic cuts, split evenly between defense and non-defense programs, was supposed to drive Congress and the President to make the hard choices on spending programs and taxes necessary to reduce the deficit. According to Senator Gramm himself, the direct goal of the Gramm-Rudman-Hollings legislation was to change the budget debate.[8] The sequestration order implementing budget cuts was intended to be a last resort.[9]

Thus, in evaluating the effectiveness of Gramm-Rudman-Hollings, one must not only look at whether the Act simply has led to a reduction in the Federal budget deficit, but also how it has affected the way Congress addresses itself to the budget. Because 1986 was the only year in which an automatic sequestration order was in effect, the work in 1986 in preparing a fiscal 1987 budget is instructive in assessing Gramm-Rudman-Hollings.

---

[7] Balanced Budget and Emergency Deficit Control Reaffirmation Act of 1987, Pub. L. No. 100-119, 101 Stat. 754. These amendments to the original Gramm-Rudman-Hollings Act were part of the September 1987 House of Representatives Joint Resolution 324, an act to raise the national debt ceiling. H.R. Res. 324, 100th Cong., 2d Sess. (1987). The Reaffirmation Act extended the time that Congress has to balance the budget by two years. It also raised the amount of deficit allowable in each year's budget. Finally, it restored an automatic sequestration process to the law in a constitutionally acceptable way. Under the Reaffirmation Act, the Executive branch Office of Management and Budget (OMB) is responsible for implementing automatic cuts.

[8] Rauch, *Is it Really Working?*, NATIONAL JOURNAL, Jan. 31, 1987, at 244, 245.

[9] STAFF OF HOUSE BUDGET COMM., 99TH CONG., 1ST SESS., A SUMMARY: THE BALANCED BUDGET AND EMERGENCY DEFICIT CONTROL ACT OF 1986 3 (Comm. Print 1986).

Gramm-Rudman-Hollings attempted to change Congress's approach to the budget process with the imposition of both strict budget deadlines and a budgeting structure to force Congress to focus on longer term planning. However, the history of Congress and self-imposed deadlines does not provide much hope that Gramm-Rudman-Hollings will have any lasting effect on the way we do business. Since 1977, the House and Senate Budget Committees have missed the deadline for reporting the first budget resolution forty-one percent of the time.[10] In addition, Congress has missed the date by which it must adopt a budget resolution eleven out of the last twelve years,[11] and the date by which it must finish its reconciliation bill six out of the last eight years.[12]

True to form and ignoring Gramm-Rudman-Hollings, in preparing the fiscal 1988 budget Congress missed the earlier June deadline mandated by Gramm-Rudman-Hollings by which the House was to report its deficit reduction bill. Rather, the date to report the bill was extended for two months until the end of September.

What about other elements of the budgeting procedure effected by Gramm-Rudman-Hollings? So far, congressional compliance is on the same failing course. In both 1986 and 1987 Congress did not even consider the reports prepared by the Congressional Budget Office and the Office of Management and Budget which specify cuts needed to meet the Gramm-Rudman-Hollings targets for those years. These reports and their consideration by both the House and Senate constitute the back-up mechanism to Gramm-Rudmann-Hollings' automatic cuts and were included in anticipation of a ruling that the automatic deficit reduction mechanism was unconstitutional.

Moreover, procedural rules included in Gramm-Rudman-Hollings designed to make legislation involving new and large expenditures more difficult to enact have not prevented the 100th Congress from considering new spending programs. A multibillion reauthorization of the Clean Water Act[13] was the first item to be considered by both the House and the Senate last

---

[10] Memorandum to Honorable Thomas Downey from Sandy Streeter Concerning Budget Deadlines 2 (Aug. 28, 1987) (on file at the Harvard Journal on Legislation) [hereinafter Downey Memorandum].

[11] CONGRESSIONAL RESEARCH SERVICE, SELECTED TABLES ON THE FEDERAL BUDGET REGARDING CHRONOLOGIES OF CERTAIN ACTIONS, AGGREGATE BUDGET LEVELS, AND OTHER INFORMATION 2 (Aug. 20, 1987) (on file at the Harvard Journal on Legislation).

[12] Downey Memorandum, *supra* note 10, at 3.

[13] Water Quality Act of 1987, Pub. L. No. 100-4, 101 Stat. 7.

year. It passed each chamber by a wide margin.[14] Welfare reform and catastrophic health insurance bills, both costing approximately $5 billion, are also working their way through the legislative process.[15]

It could be that Congress is treating the Act's procedural requirements as it treats any other budget rule—cautiously. But more likely, the uneasiness expressed by Members during the initial Gramm-Rudman-Hollings debate continues to run high within the Congress. Simply put, it is difficult for Members to swallow Gramm-Rudman-Hollings' force feeding of budget and tax decisions. A majority of the House and Senate may have voted in support of Gramm-Rudman-Hollings. But this same majority also realizes that it is fantasy to assume that all public policy decisions can be reduced solely to the question of deficit reduction.

As for Gramm-Rudman-Hollings' actual deficit reduction effects, the news is not much better. In 1986, Congress proclaimed that it had "met" the Gramm-Rudman-Hollings deficit target by reducing the deficit to $151 billion. Under the Act, the actual deficit target for fiscal 1987 was $144 billion. However, because of the $10 billion "margin of error" allowable under the Act (which was supposed to compensate for errors in economic forecasting), Congress was able to claim credit for meeting the deficit target and complying with the law. The $151 billion figure was the result of a $11.7 billion sequester that took place in March[16] and the enactment of a budget reconciliation bill that cut another $11.7 billion.[17]

Despite the appearance of near success, the truth is that the hard choices about which programs to cut or how to raise taxes were not made. Instead, Congress relied heavily on smoke and

---

[14] *See* 133 CONG. REC. H214 (daily ed. Jan. 8, 1987); 133 CONG. REC. S1003 (daily ed. Jan. 21, 1987).

[15] The Family Welfare Reform Act, H.R. 1720, 100th Cong., 2d Sess. (1987) has passed the House and is currently being marked up in a Senate committee. The Catastrophic Protective Health Act of 1987, H.R. 2470, 100th Cong., 2d Sess., has been passed by both the House and Senate and is in conference.

[16] After the sequestration order was found unconstitutional in Bowsher v. Synar, 106 S. Ct. 3181 (1986), the budget cuts which resulted from the sequestration order were reaffirmed by votes in both the House and Senate. *See* H.R.J. Res. 672, Pub. L. No. 99-366, 100 Stat. 773 (1986).

[17] Omnibus Budget Reconciliation Act of 1986, Pub. L. No. 99-509, 100 Stat. 1874 "Reconciliation" is a legislative mechanism through which tax and spending bills are adjusted to conform with the figures set in the congressional budget resolution. The mechanics of the budget reconciliation process is set forth in the Congressional Budget and Impoundment Control Act of 1974, Pub. L. No. 93-344, 88 Stat. 297.

mirrors to remain within the Gramm-Rudman-Hollings deficit reduction targets. First, in constructing the deficit reduction reconciliation bill, Congress relied heavily on additional revenue that the Tax Reform Act of 1986 was supposed to produce. As it turned out, the first-year revenue increase under the tax reform bill made up twelve percent of the 1987 deficit "cuts."[18] Another twelve percent of the 1987 deficit "cuts" came from one-time sell-offs of federal assets,[19] such as Conrail, and from accounting gimmicks such as moving a military payday from the last day of fiscal 1987 to the first day of fiscal 1988.[20] Gramm-Rudman-Hollings actually forced the Congress *away* from the long-term planning that its proponents so desperately desired.

Congress openly ignored Gramm-Rudman-Hollings' $108 billion target for fiscal 1988. According to the budget agencies it would have required about $60 billion in savings to reach this goal. But the fiscal 1988 Budget Resolution only called for about $37 billion in savings to be achieved by about $19 billion in program cuts and $19 billion in new taxes. Nevertheless, while the Gramm-Rudman-Hollings target might not be met with this resolution, one might argue that in deciding to raise a significant amount of new revenue, Congress had met the challenge implicit in Gramm-Rudman-Hollings.

Wrong. Instead of facing the original Gramm-Rudman-Hollings targets which were supposed to force difficult decisions, the 100th Congress decided to make the targets easier to meet with the amendments described above.[21] Under the Reaffirmation Act, Congress has two extra years in which to balance the budget.

The real effect of the Reaffirmation Act's changes will be to generally postpone the difficult spending and tax decisions until future years. For instance, under the Reaffirmation Act, the new deficit reduction target required only $23 billion in savings for fiscal 1988. This was even $15 billion less than the savings called for in the fiscal 1988 budget resolution which missed the original Gramm-Rudman-Hollings target for that year by $23 billion.

---

[18] Rauch, *supra* note 8, at 247.

[19] *Id.*

[20] The Omnibus Budget Reconciliation Act of 1986, Pub. L. No. 99-509, 100 Stat. 1874.

[21] *See supra* note 7.

In budgets since the passage of Gramm-Rudman-Hollings, Congress has fudged, ignored, and then eased the Act's deficit cutting requirements. One is therefore compelled to ask: why won't it happen again and again, each time it is cards-on-the-table time?

Even if real deficit reduction were to occur through Gramm-Rudman-Hollings, the economic assumptions underlying the Act might actually be assumptions undermining it. Gramm-Rudman-Hollings is a scientific approach to deficit reduction. However, at its core it is based on the imprecise art of economic forecasting. The annual savings targets set by the law which must be met by Congressional action or the automatic sequestration order, are based on estimates of federal receipts and outlays. These, in turn, depend on assumptions about general economic conditions. There are several variables which must be factored into these assumptions, including interest rates, the gross national product, the unemployment rate, the international trade deficit, and inflation.

A budget estimate is therefore bound to be extremely sensitive to changes in economic conditions. This seriously complicates budget planning because the inevitable errors in forecasting the economy's performance necessarily lead to errors in the budget forecast.

Obviously, errors do occur. For example, in January 1987, both the Congressional Budget Office and the Office of Management and Budget found the deficit to be running above $170 billion, nowhere near the $151 billion that had been forecast at the time the fiscal 1987 Budget Reconciliation Bill was enacted. However, it appears that we had it right in the first place, as the August 1987 sequester report produced by the same two budget agencies have since recalculated the fiscal 1987 deficit to be about $157 billion. Between 1980 and 1986, the Congressional Budget Office underestimated the deficit every year by an average of almost $50 billion.[22]

The consequences of forecasting errors can be enormous. A one percentage point reduction in economic growth will add $9 billion to the deficit, according to the Congressional Budgeting Office.[23] Thus, even despite the best efforts of its authors,

---

[22] Rauch, *CBO's Wishful Thinking*, NATIONAL JOURNAL, Mar. 7, 1987, at 550, 551.

[23] Wehr, *Gramm-Rudman 'Sequester': Nothing Doing*, CONG. Q., Aug. 22, 1987, at 1947.

the reliance on economic assumptions in Gramm-Rudman-Hollings might itself doom its efficacy. As a result of imprecise assumptions, any year's compliance with the set deficit target could easily result in a short-lived ephemeral victory overturned by the "real" numbers.

Another area in which Gramm-Rudman-Hollings may trip over its own feet is in producing a healthy economy, the overarching goal of deficit reduction. Gramm-Rudman-Hollings requires that the automatic targets be met only when the economy is growing at more than one percent per year.[24] Yet if the economy were to grow at only one and a half percent per year, for instance, unemployment would likely be on the rise as would other symptoms of national economic ill-health. At a time of such limited growth, a further contraction of the economy that would result from large budget cuts would only make things worse. Interest rates would rise, costing the Federal government billions; unemployment would rise, slowing the economy down even more as spending for unemployment increased and tax revenues declined. Yet under Gramm-Rudman-Hollings, the cuts would still come.

There is no denying that hard choices need to be made about the deficit. However, Gramm-Rudman-Hollings is not a real solution to the deficit. Clearly, if Congress wants to get around the law, it can. In fact, Congress has done just that and shows every sign of continuing to do so.

The forces to produce deficit reduction were set in place two hundred years ago by the Founders of this Nation. Article I of the Constitution squarely confers the power to tax and spend—the power to make the tough choices—upon the Congress.[25] And as the Supreme Court has reconfirmed, "[in] the framework of our Constitution, the President's power [is] to see that the laws are faithfully executed."[26] Thus, "[the] principle of separation of powers was not simply an abstract generalization in the minds of the framers."[27] The Congress and its authority to write laws is balanced against the President and the President's prerogative to veto legislation. The Founders intended these

---

[24] Gramm-Rudman-Hollings provides that during times of recession, defined as less than one percent national growth, the requirement to meet the deficit reduction targets is suspended. *See supra* note 2, § 254.

[25] U.S. CONST. art. I, § 8.

[26] Youngstown Sheet and Tube v. Sawyer, 343 U.S. 579, 587 (1952).

[27] Buckley v. Valeo, 424 U.S. 1, 124 (1975).

two institutions to work individually and then come together to make the "hard choices" that are the backbone of a representative government.

Since 1980 we have had a President who has been unwilling to participate with Congress in the legislative process as envisioned by the Founders. In 1981, President Reagan pushed through the largest tax cut in the history of this Nation[28] while also initiating a massive increase in defense spending. This is what created the deficit problem in the first place. In asking that the Federal deficit be eliminated, the President steadfastly focuses on spending cuts alone. He will not explore the truly difficult decision of raising revenue. Thus, Congress momentarily convinced itself that it too could ignore its constitutional mandate to make tough decisions by agreeing to the sequestration order in Gramm-Rudman-Hollings.

The hard decisions Gramm-Rudman-Hollings seeks to engender will only be made when the chemistry is right between the individuals who make up the Congress and the Executive. Gramm-Rudman-Hollings ignores this altogether. This indifference is the final reason why Gramm-Rudman-Hollings cannot work.

[28] Economic Recovery Tax Act of 1981, Pub. L. No. 97-34, 95 Stat. 172.

THE POLITICS OF THE ENACTMENT AND IMPLEMENTATION OF GRAMM-RUDMAN-HOLLINGS: WHY CONGRESS CANNOT ADDRESS THE DEFICIT DILEMMA

JOHN W. ELLWOOD*

In the fall of 1985, Congress passed and President Reagan signed the Gramm-Rudman-Hollings Balanced Budget and Emergency Deficit Control Act.[1] This law was passed after Congress and President Reagan failed for five years to reduce meaningfully the largest peacetime federal budget deficits since the Great Depression.[2] As originally enacted, Gramm-Rudman-Hollings encompassed two classes of changes in the budget process. The first set of changes created a "doomsday machine"

* Associate Professor of Public Policy and Management, The Amos Tuck School of Business Administration, Dartmouth College. A.B., 1964, Franklin & Marshall College; M.Poli. Sci., 1967, The Johns Hopkins University; Ph.D., 1972, The Johns Hopkins University.

A previous version of this Essay was presented at the 1987 annual meeting of The American Economics Association on December 28, 1987 in Chicago, Illinois. Research for this Essay was supported by a Tuck Associates Research Grant.

This Essay would not have been possible without the kind assistance of Professors Joseph White and Aaron Wildavsky. Much of the historical material has been drawn from their book-length manuscript, The Battles of the Budget: From the Last Year of Carter Through Gramm-Rudman-Hollings and Tax Reform. They also were kind enough to make available their interviews of many of the participants who affected the passage and implementation of the first version of Gramm-Rudman-Hollings. I re-interviewed many of the same individuals to gain their views on the passage and implementation of the second version of Gramm-Rudman-Hollings. Although one can not help but be affected by the White-Wildavsky analysis, the views expressed in this Essay are mine. You are urged to read the full White-Wildavsky budgetary history once it is published.

[1] The Balanced Budget and Emergency Deficit Control Act of 1986, Pub. L. No. 99-177, 99 Stat. 1037 [hereinafter Gramm-Rudman-Hollings].

[2] From fiscal year 1934 through fiscal year 1940 the federal budget deficit equalled 3.47% of the Gross National Product (GNP). During the 1950's the average deficit equalled .42% of GNP. This statistic rose to .8% of GNP during the 1960's and to 2.43% of GNP during the 1970's. From fiscal year 1980 through fiscal year 1986, the federal deficit equalled 4.5% of GNP. OFFICE OF MANAGEMENT AND BUDGET, BUDGET OF THE UNITED STATES: HISTORICAL TABLES: FISCAL YEAR 1988 table 1.2 (1987).

A similar pattern occurs for the standardized-employment deficit as a percentage of potential GNP. From fiscal year 1956 through fiscal year 1959 this ratio equalled 0.1%. During the 1960's it rose to an average of 1.0%. It further increased to an annual average of 1.7% during the 1970's. From fiscal year 1980 through fiscal year 1986 it rose to a yearly average of 2.8% of potential GNP. CONGRESSIONAL BUDGET OFFICE, THE ECONOMIC AND BUDGET OUTLOOK: AN UPDATE 102, table C1 (1987).

Although some economists have argued that these standard measures of the federal deficit overstate the magnitude of the deficits of recent years, their corrections have had little impact on the public and congressional perceptions of the size of the deficit problem. *See, e.g.*, R. EISNER, HOW REAL IS THE FEDERAL DEFICIT? (1986).

of "across-the-board" budget reductions that would take place unless the President and Congress agreed to their own deficit reduction package. The automatic cuts would occur in any year that the budget package agreed upon failed to reduce the federal deficit to a maximum amount set by the Act—$144 billion in fiscal year 1987, $108 billion in fiscal year 1988, $72 billion in fiscal year 1989, $32 billion in fiscal year 1990, and zero in fiscal year 1991.[3] The second set of modifications formally incorporated into law the procedural changes that had been adopted informally over the past decade and created a series of new rules and points of order to strengthen the power of the congressional budget committees.[4]

Two years later, faced once more with the need to raise the limit on the federal debt and confronted with the prospect of very large automatic budget cuts (which the Act refers to as sequestrations), congress enacted and President Reagan signed the Balanced Budget and Emergency Deficit Control Reaffirmation Act of 1987.[5] This Act raised the federal debt-limit to a high enough level to avoid another debt-limit increase during the Reagan presidency, and modified the Gramm-Rudman-Hollings deficit reduction targets and procedures to make them more politically feasible.

This Essay addresses the politics behind the creation and implementation of Gramm-Rudman-Hollings and the Reaffirmation Act. Public choice explanations of congressional behavior put great stress on congressional credit claiming in order to maximize chances of reelection, but this Essay points out that on some issues where credit claiming is impossible, congressmen seek to avoid blame. This Essay posits that congressmen

---

[3] For detailed descriptions of the procedures of Gramm-Rudman-Hollings, see Sequestration Report for Fiscal Year 1986—A Joint Report to the Comptroller General of the United States, 51 Fed. Reg. 1,922–41 (OMB and CBO 1986); Sequestration Report for Fiscal Year 1988—A Joint Report to the Temporary Joint Committee on Deficit Reduction, 52 Fed. Reg. 31,534–53 (OMB and CBO 1987); Sequestration Report for Fiscal Year 1988; Notice, 52 Fed. Reg. 38,683–93 (CBO 1987); Keith, Changes in the Congressional Budget Process Made by the Balanced Budget Act (P.L., 99-177) (Congressional Research Service, Report No. 86-713 GOV, May 13, 1986); Davis & Keith, Congressional Budget Process Reform (Congressional Research Service Issue Brief, Order Code IB87196, Updated Oct. 7, 1987).

[4] For the best detailed descriptions of the non-sequestration procedural changes enacted as part of Gramm-Rudman-Hollings, see Davis & Keith, Debt-Limit Increase and the 1985 Balanced Budget Act Reaffirmation: Summary of Public Law 100-119 (H.J. Res. 304) (Congressional Research Service, Report No. 87-865 GOV, Oct. 29, 1987).

[5] The Balanced Budget and Emergency Deficit Control Reaffirmation Act of 1987, Pub. L. No. 100-119, 101 Stat. 754 [hereinafter the Reaffirmation Act].

also seek to maximize the achievement of their ideological goals and their power within Congress. Only by taking these last two points into account can one understand the key roles played by Senator Phil Gramm (R-Tex.), Congressman Dan Rostenkowski (D-Ill.), and the frustrations caused by the budgetary deadlocks of the 1980's.

The first part of this Essay presents a brief fiscal and political analysis of the adoption and implementation of Gramm-Rudman-Hollings. The following part provides an overview of the law's major provisions and describes the events leading to the adoption of the Reaffirmation Act. The final part evaluates the automatic sequestration procedures of Gramm-Rudman-Hollings as a version of an expenditure-limitation or balanced budget amendment.

I. The Adoption and Content of Gramm-Rudman-Hollings and the Reaffirmation Act

Dramatic changes giving rise to political frustration have punctuated the budgetary history of the 1980's. Within this decade Congress enacted the largest peacetime defense buildup in American history. It also passed the largest tax reduction in American history (The Economic Recovery Tax Act of 1981, or ERTA)[6] immediately followed by the largest peacetime tax increase in United States history (The Tax Equity and Fiscal Responsibility Act of 1982, or TEFRA).[7] The size of the defense buildup and the magnitude of the ERTA tax cut, even after TEFRA's tax increase, combined with the effects of the 1982-1983 recession, created a series of budget deficits of unprecedented size for peacetime history. A series of congressional reconciliation and appropriations measures failed to reduce the deficit as promised, and the President and Congress still have not been able to reconcile their policy differences on ways to break the budgetary log-jam created by economic developments and legislative-executive conflict since 1981.

The key problem of federal budget-making since 1982 has been nominal budget deficits in excess of $200 billion and structural budget deficits in excess of $100 billion, which comprise three percent of potential GNP. Each year, after much debate

[6] Economic Recovery Tax Act of 1981, Pub. L. No. 97-34, 95 Stat. 172.

[7] Tax Equity and Fiscal Responsibility Act of 1982, Pub. L. No. 97-248, 96 Stat. 324.

and delay, Congress would enact a budget resolution containing a multi-year plan promising a balanced budget, or at least a trend of decreasing deficits. The battle of the budget resolution would be followed inevitably by individual battles over reconciliation bills to implement the entitlement and tax portions of the accepted deficit reduction plan, and appropriations bills incorporating budget cuts for those accounts funded on an annual basis. Over the course of this decade, the enactment of reconciliation and appropriations legislation took increasingly longer until in recent years reconciliation bills were either not passed or were enacted half way through their initial fiscal year, and appropriations bills were grouped into single omnibus bills enacted as year-long continuing acts. Although consideration of the budget has increasingly dominated the congressional calendar, recent budgetary action appears to have had a minimal effect on the deficit, which still exceeds $150 billion despite the political battles and service reductions.

II. WHY WASN'T THE DEFICIT REDUCED?

Two factors account for most of the failure to reduce the deficit significantly prior to the passage of Gramm-Rudman-Hollings: the performance of the economy, and the policy disagreements between President Reagan and Senate Republicans and House Democrats. The deep recession of 1981-1982 led to a slower average growth rate for the first half of the decade—an annual average of 2.5%—than during the previous ten years—2.8%.[8] Each year the President and Congress would assume faster economic growth than would eventually occur. Even the most positive economic event associated with the recession, the dramatic decline in the rate of inflation, hurt efforts to reduce the federal deficit beginning in calendar year 1984 because ERTA indexed the rates of the income tax.[9]

The policy deadlock between the President and Congress had an even greater effect on the failure to reduce the deficit. Pres-

[8] Compounded growth rates calculated from ECONOMIC REPORT OF THE PRESIDENT table 2-B, 250 (1988).

[9] Prior to the indexation of the individual income tax rate, inflation would cause individuals to fall into higher tax brackets. This would lead to additional revenues for the federal government. Thus, prior to indexation, inflation helped the federal government reduce the deficit. In addition, the decline in the rate of inflation occurred faster than budget planners anticipated.

ident Reagan came to office seeking to lower taxes (particularly marginal tax rates), increase spending for the military, and significantly reduce the size of the federal government through large domestic budget cuts. Because of the President's refusal to support a significant tax increase and the Democrats' refusal to reduce major middle-class and means-tested entitlements, the deficit problem had to be solved by reductions in budget accounts that contributed less than twenty percent of total federal spending. This policy deadlock worsened over time as net interest on the public debt increased in proportion to federal spending, and as both President Reagan and the Democrats learned that those who advocated tax increases (such as Walter Mondale in 1984) or favored cuts in middle-class entitlements such as Social Security (such as Republicans in 1982 and 1984) suffered at the polls.

The budgetary effects of this deadlock can be seen in data which show the sources of the differences between actual budget totals and those promised in the first budget resolutions that were enacted for fiscal years 1982 through 1986 (Gramm-Rudman-Hollings was enacted in 1985). From 1981 (fiscal year 1982) to the enactment of Gramm-Rudman-Hollings in calendar year 1985 (fiscal year 1986), the actual federal deficit exceeded the planned congressional deficit by an annual average of $47.7 billion. Of this amount, $32.5 billion (68%) was due to less vigorous economic performance than assumed by Congress when it enacted its budget plan, $12.9 billion (27%) was due to the failure to enact policy changes that were promised in the budget resolution, and $2.4 billion (5%) was due to estimating errors by the technical staffs of the budget committees and the Congressional Budget Office (CBO).

Policy disagreements contributed more to the failure to reduce the deficit by a significant amount as the years progressed. For fiscal years 1982 and 1983, an annual average of $5 billion, or only six percent of the difference between Congress' budget plan and actual spending, was due to failure to enact policy changes that were assumed in the budget resolution. For fiscal years 1984, 1985, and 1986, however, deficits arising from promised but unenacted policy changes rose to an annual average of $18 billion, or sixty percent of the difference between what was promised in the budget resolution and what actually occurred.

Congress and the President were firmly deadlocked, especially over defense expenditures, where the President's 1986

budget proposed an increase of six percent in budget authority above inflation. Democrats in the House continued to block attempts to reduce spending from non-means tested entitlements and the Republicans were particularly wary of being attacked as being in favor of cutting Social Security. Moreover, the President refused to go along with a tax increase and Democrats, painfully aware of the fate that befell Walter Mondale in the 1984 presidential election, refused to bring up the subject. The non-defense discretionary programs that remained had already been trimmed: in real terms their outlays were already reduced by over fifteen percent between fiscal years 1980 and 1985. Unless something gave way there just was not enough money left in the budget to achieve real deficit reduction.

The budgetary history of the 1980's is not the only key to understanding the passage of Gramm-Rudman-Hollings, however. It is necessary to examine as well the overall political atmosphere within which the Act was formulated and adopted. The same major factors that motivate elected members of Congress in other areas of activity affect their actions in the policy disagreements and the political compromises of the deficit reduction process. These factors are: (1) the desire to maximize prospects for reelection and (2) the need to avoid blame for unpopular legislation.

III. Maximizing Reelection through Credit-Claiming

Traditionally, legislators increase their chances of reelection by "credit claiming"—by taking credit for government measures that benefit the constituents they represent. The easiest way to credit-claim is to supply public goods and services directly to one's constituents. Such credit claiming, however, has become increasingly difficult in the 1980's. Although President Reagan has been unable to reduce the size of the government considerably, his refusal to raise taxes has limited the ability of legislators to create new programs or increase significantly the dollar levels of existing efforts. Moreover, accounts that are funded through annual appropriations bills have tended to suffer the largest budget reductions since 1981, and it is the programs funded by these accounts that are the most highly targeted to individual districts and states. Thus, the very programs that

involve the greatest amount of credit-claiming are the ones that have borne the biggest spending cuts.

Some observers have argued that the American public will support elected officials who endorse actions that will lead to smaller deficits. However, a consistent pattern of public opinion calls this view into question. In the best analysis of the public's inconsistent attitudes toward taxes and public spending, David O. Sears and Jack Citrin analyzed the attitudes of Californians during the Proposition 13 tax revolt of 1978 and 1979. They concluded:

> [S]ubstantial majorities of the California electorate wanted cutbacks in government spending and taxes, and expressed strong preferences for a smaller or less powerful government bureaucracy, while at the same time (and by equally strong majorities) requesting additional services in most areas of government responsibility. On the face of it, the public seemed to want something for nothing. This paradoxical mixture of attitudes prevailed throughout the period of the California tax revolt. And the same mentality is evident in the attitudes of Americans nationwide.[10]

Sears and Citrin explain this apparent inconsistency by pointing out that a majority of those sampled believed that their taxes could be lowered without a reduction in services through the elimination of government "waste." The authors point out that, "[t]he same mass belief in widespread waste in government could be seen when the public was asked how much the government could cut spending without reducing services. At all stages of the tax revolt, Californians overwhelmingly believed that it was possible to cut spending without reducing services. . . ."[11]

Finally, Sears and Citrin found that the services least favored by the electorate were those that were "means-tested," awarded only upon demonstration of need. On the other hand,

> programs whose benefits are available to everyone, at least in principle, such as police and fire and schools, are more widely favored than those with specialized clienteles, such as public housing and welfare. Universal access to such programs makes them expensive, of course, so we are left with the irony that the public supports expansion of the

---

[10] D. SEARS & J. CITRIN, TAX REVOLT: SOMETHING FOR NOTHING IN CALIFORNIA 44 (1982). For similar findings with national data, see Ladd, et al., *The Polls: Taxing and Spending*, 43 PUB. OPINION Q. 126 (1979).

[11] D. SEARS & J. CITRIN, *supra* note 10.

costliest governmental responsibilities, while simultaneously demanding reduced taxes.[12]

IV. Implications for Credit-Claiming

To the extent that Sears and Citrin are correct, congressmen face a no-win situation when it comes to deficit reduction. To economists and budget experts, the arithmetic of the federal budget mandates increased taxes, reductions in non-means-tested middle-class entitlements, or both, in order to achieve significant deficit reduction. But these actions are precisely those with the least public support. Even though the costs of limitations on the growth of non-means-tested entitlements would be relatively diffuse, the public would provide little if any support for the legislator who votes for the reduction in growth.

So far the public has not had to face the fiscal trade-offs that economists have highlighted. For the public, the perceived existence of massive amounts of "waste" eliminates the need for hard choices. While many policy makers have advocated an education effort to make sure that the public realizes that spending cannot be significantly reduced without reductions in real services, few elected officials are willing to confront the next election knowing that their efforts at real deficit reduction have drastically reduced their ability to engage in credit-claiming.

V. Blame Avoidance

When politicians cannot engage in credit claiming they turn to various strategies of avoiding blame.[13] One blame avoiding strategy is to redefine the issue. Thus, many Republicans (and

[12] *Id.* at 49.

[13] Kent Weaver suggests that politicians will engage in blame avoiding behavior with respect to a particular policy when their interests conflict with those of their constituents, when as many constituents are harmed as are benefited, when some constituents are especially adversely affected, or when embracing the policy achieves no benefit whatsoever. K. Weaver, Automatic Government: Indexation and the Politics of Blame (1988). He contends that politicians will engage in blame avoiding behavior when (1) "there is a zero-sum conflict among the policymakers' constituents"; (2) "when all possible alternatives have strong negative consequences for at least some of the policymakers' constituents"; (3) "when constituency opinion is [so] overwhelmingly on a single side of an issue . . . that there is little credit to be derived from agreement with it"; and (4) "when the personal or policy interests of the policymaker and clientele are opposed." *See generally id.* When faced with a non-credit-claiming situation policymakers can turn to several blame avoiding strategies.

some Democrats) saw Gramm-Rudman-Hollings as a vehicle to turn the issue of voting for a $2 trillion debt limit into one of doing something about the deficit. A second strategy is to continue to throw resources at the problem in order to secure the status quo. Thus, while economists cannot understand why legislators are creating a much worse problem for themselves in the future by refusing to deal with the deficit now, such action is sensible for the elected official when blame avoidance is a paramount concern.

A third strategy is to pass the blame to other individuals or institutions. Many legislators were attracted to Gramm-Rudman-Hollings because it offered the opportunity to shift the blame from Congress to the President or to technicians at CBO or OMB. Of course, the President could, at the same time, seek to use Gramm-Rudman-Hollings to shift the blame from the White House to Congress.

A fourth strategy calls for all policymakers to agree to simultaneous support of the unpopular policy. Under this strategy policymakers seek strength in numbers. The Senate Republicans appeared partially to adopt this strategy in the spring of 1985 when they sought to limit the COLAs of the major middle-class entitlements. Under a fifth strategy, when decisionmakers feel that their policy positions are at odds with those of a majority of their constituents they can seek process and policy solutions that limit their discretion. In the case of deficit reduction, for example, they might create a "doomsday machine" such as Gramm-Rudman-Hollings that will bring about changes which they support but their constituencies oppose. Finally, the classic blame avoidance strategy is to pass the buck—and what better way to receive this passed buck than an automatic process. Weaver points out: "Congress set in motion a process which months or years later was to cause cuts to be made automatically, distancing themselves from the blame. Even the officials who would be responsible for sequestering funds are simply following a mandated formula, so they cannot be blamed."[14]

VI. Blame Avoidance: Conclusions

Blame avoidance can explain the reluctant support for Gramm-Rudman-Hollings. Because they were unable to claim

[14] K. Weaver, *supra* note 13, at 25.

any credit on deficit reduction measures with their constituents, legislators eventually had to turn to a mechanism that minimized their blame. An automatic doomsday machine was a perfect candidate.

But blame avoidance cannot explain the behavior of those policy entrepreneurs who actively sought the spotlight for their efforts to enact Gramm-Rudman-Hollings or the Reaffirmation Act. It cannot explain Senator Gramm's (R-Tex.) behavior nor that of Senator Domenici (R-N.M.) and his colleagues who voted for the Senate budget resolution in 1985. Nor can it explain Chairman Rostenkowski's actions in insisting on a "hard trigger" in the rewrite of Gramm-Rudman-Hollings in 1987. To explain these actions we have to turn away from the maximization of constituency support and toward the other two motivations—maximization of one's policy beliefs and maximization of one's power and position.

VII. Events Leading to Gramm-Rudman-Hollings

In 1985 the Senate Republican leadership put forth a budget resolution that challenged spending for two sacred programs, defense and Social Security. The resolution, which passed by a single vote, would have held increases in budget authority for defense to three percent above inflation and canceled the cost of living allowance for all indexed programs, including Social Security, for one year.

The efforts of the Senate Republican leadership and the House's moderate Democrats to deal with the budget were undercut by the President and the Speaker of the House. Instead of supporting his Senate leadership, President Reagan opposed the defense aspect of the compromise. Speaker O'Neill (D-Mass.) and the House Democratic leadership then backed off from any agreement that would reduce Social Security payments. Massive amounts of time and political capital had been spent only to re-establish the budgetary deadlock. By August 1985 it would have been hard to find a committee chairman or a party leader in either chamber who approved of the budget process.

It was within this environment that Congress, and particularly the Senate, faced the need to raise the debt limit to $2 trillion. During the 1970's the periodic need to raise the debt limit had

caused crises on the House floor as members cast symbolic votes against raising the limit in the wake of the consequences of their prior support for large expenditures. In order to avoid this yearly (and occasionally multi-yearly) charade, the House amended its Rule XLIX so that the Clerk of the House would "(engross) and (send) to the Senate a joint resolution adjusting the public debt limit to the amount set forth in the budget resolution" after agreement was reached on the final budget resolution.[15]

Raising the debt limit above $2 trillion was potentially embarrassing to Senate Republicans. Senator Phil Gramm, however, saw the need to pass the new limit as an opportunity to force his reluctant colleagues to accept a legal version of a constitutional amendment to limit spending and prohibit federal budget deficits. During the 1970's and 1980's Congress had enacted a series of such laws. Most simply stated that in the absence of war (and possibly a recession), Congress could not enact appropriations after a certain date which would lead to a federal budget deficit. Senator Gramm himself had offered such proposals in the past: in 1979 he tried to add a balanced budget requirement to a House debt limit bill, and in 1980 he co-sponsored a bill with then Majority leader Jim Wright (D-Tex.) which would have required the President to sequester funds if the budget was not balanced.[16]

Senator Gramm rewrote the bill that he had proposed in 1980 to require a balanced budget over a five year period. He then contacted the White House and the Senate Budget Committee to tell them of his intention. Democratic Senators Warren Rudman (R-N.H.) and Ernest Hollings (D-S.C.) also joined as cosponsors. The Treasury and the Senate leadership did not take the effort seriously, but the Senate's rules required the Senate to vote on Senator Gramm's proposal. Moreover, the proposal had a great deal of symbolic appeal. By supporting it President Reagan could better withstand Democratic attacks on the $2 trillion federal debt level and show that he was serious about deficit reduction.[17]

[15] E. DAVIS & R. KEITH, DEBT-LIMIT INCREASE AND 1985 BALANCED BUDGET ACT REAFFIRMATION 3 (1987).

[16] J. White & A. Wildavsky, The Battles of the Budget: From the Last Year of Carter Through Gramm-Rudman-Hollings and Tax Reform 32 (unpublished manuscript).

[17] White and Wildavsky also report that prior to Senator Gramm's contacting the administration, the White House staff had been working on a multi-year deficit reduction package that would be proposed to offset the anticipated negative publicity that would result from the debt limit fight. *See id.* at 39.

Once introduced, the proposal quickly became impossible to stop, because Senator Gramm was amenable to changes in order to gain support. When Senator Gramm agreed to place Social Security beyond the limits of the automatic budget reduction process, the proposal's passage in the Senate was assured. Many Senators believed that the House Democrats would kill the proposal anyway, but the House leadership's initial vote count indicated anything but defeat. The House leaders, therefore, chose to go to the Senate-House Conference and push for the exemption of most entitlements and means-tested programs from sequestration and for the addition of language that would leave the sequestration process vulnerable to a court test of its constitutionality.

Gramm-Rudman-Hollings was virtually rewritten by the conferees. From a technical standpoint, Senator Gramm's original proposal was so vague that it was unworkable. In threatening sequestration of outlays in order to achieve necessary deficit reduction, it did not take into account program to program variations in timing of budget authority. These variations occur because Congress does not directly control outlays but rather grants authority to agencies to enter into obligations; only when those obligations come due does the Treasury make a disbursement. Each year, over fifteen percent of the Treasury's outlays result from budget authority and obligations granted in previous years.[18] This process makes across-the-board cuts very difficult to achieve.

Political considerations also required extensive rewriting of the bill. Many participants saw an opportunity to achieve some of their goals through the Gramm-Rudman-Hollings structure and process. Senator Gramm and other conservatives saw an opportunity to create a smaller domestic public sector. President Reagan viewed Gramm-Rudman-Hollings as good because it avoided tax increases while guaranteeing domestic cuts (instead of merely promising them).[19] On the other side, "liberals hoped that it would make deeper cuts in defense."[20] Upon first seeing the bill, Representative Richard Gephardt (D-Mo.) exclaimed, "Can this be?—it's the undoing of the Reagan revolution, en-

---

[18] OFFICE OF MANAGEMENT AND BUDGET, BUDGET OF THE UNITED STATES GOVERNMENT: HISTORICAL TABLES, FISCAL YEAR 1988 table 8.1, 8.1(1)-(2).

[19] *Id.* at 46.

[20] *Id.*

suring a defense cut and almost certainly a revenue increase."[21] Senate and House Republicans who had been undercut by the President's compromise with the Speaker saw Gramm-Rudman-Hollings as a mechanism to force the President and the Democratic leadership to begin making the hard choices necessary to achieve significant deficit reduction.

On December 12, 1985, after a series of massive conferences that lasted most of the fall, the Gramm-Rudman-Hollings Act became law. Some legislators, such as Senator Gramm, saw it optimistically as the way to a balanced budget. Others, such as Senator Rudman, saw it merely as "a bad idea whose time has come."[22] House Majority Leader Tom Foley (D-Wa.) was even less generous: he expressed the opinion that the bill was "about the kidnapping of the only child of the President's official family that he loves (the defense budget) and holding it in a dark basement and sending the President its ear."[23] One conservative Republican Congressman summed up the events leading to Gramm-Rudman-Hollings even more critically:

> And the process [of 1985] led to Gramm-Rudman-Hollings as well—one of the most idiotic developments in my years in the legislative business . . . the Senate did a good job [on the fiscal year 1986 budget resolution], the President rejected it, we [the House] kicked them [the Senate] in the teeth, and the Senate did Gramm-Rudman-Hollings in return. Then we screwed it up more over here by taking half the budget out.[24]

VIII. Major Provisions of Gramm-Rudman-Hollings

Gramm-Rudman-Hollings contained five major provisions. First, it extended the federal debt ceiling to over $2 trillion. Second, it created multi-year deficit targets that would lead to a balanced budget at the end of a five year period. Third, it created a procedure under which all non-exempt projects, programs and activities (PPAs) would be subject to automatic cuts (sequestration) if the President and Congress failed to agree upon a budget plan by a certain date. Fourth, it fashioned a new timetable aimed at speeding up budgetary action so that the President and Congress could avoid the sequestration process.

---

[21] *Id.* at 44.
[22] *Id.* at 46.
[23] *Id.* at 9.
[24] Interview with unnamed Member of Congress (Dec. 21, 1987).

Finally, it incorporated many procedural changes within the budgetary process. Arguably the most significant of the procedural changes was a series of requirements that would force authors of substitutes and amendments to reconciliation bills and appropriation acts to include budget cuts or tax increases to compensate for any deficit increase that their bills would produce.

A. *Multi-Year Deficit Targets*

Table I contains the multi-year deficit targets of Gramm-Rudman-Hollings and the Reaffirmation Act. The time patterns of the targets of Gramm-Rudman-Hollings and the Reaffirmation Act reflect election year politics. The acts were enacted at the end of the first session of the 99th and 101st Congresses, respectively. Thus, in each case the law would go into full effect during an election year. The conservative Republican authors of each law sought to shift the hardships of budget cuts beyond the election horizon, whereas the House Democrats sought to "front-load" the pain so that, should a sequestration occur, its effects could be blamed on the sitting Republican President and his congressional allies.[25]

B. *Sequestration Procedure*

The sequestration process remains the heart of Gramm-Rudman-Hollings. It is congress' attempt to create a doomsday machine that threatens to cause so much pain that the various parties to the budget debate will agree to a compromise rather than suffer the consequences of automatic implementation. Under Gramm-Rudman-Hollings, the CBO and the Office of Man-

[25] Robert Hartman has described the process that led to the time pattern of the Gramm-Rudman-Hollings targets:

> Last fall's [fall 1985] best deficit estimate for FY 1986 was $192 billion. Originally Gramm-Rudman-Hollings sequestrations were to be triggered if the deficit exceeded the target by 7 percent in the first year. Thus, if the deficit target was set at $180 billion, the expected deficit would just avoid the sequestration. That $180 billion number became the first-year target because the sponsors' intent was not to activate the procedures until 1987. The pre-ordained end point, five years later, was a target of zero. 180 divided by 5 is 36, so the deficit sequence was 144, 108, 72 etc.

R.W. Hartman, *Gramm-Rudman-Hollings After One Year* 11 (paper delivered at the 1986 Annual Meetings of the Association for Public Policy and Management, Austin, Tex., Nov. 1, 1986).

agement and Budget (OMB) would independently estimate the degree to which congressional action had failed to reach the year's budget target. If disagreement occurred, the Comptroller General (the head of the General Accounting Office) was empowered to decide in favor of either agency.

To calculate the sequestration amount (the percentage by which each non-exempt account would be automatically cut), the maximum deficit amount allowed for a given year under Gramm-Rudman-Hollings would be subtracted from the baseline deficit approved by the Comptroller General. If the baseline deficit exceeded the maximum deficit by $10 billion, the difference (excess deficit) was divided equally between defense and non-defense programs. In both the defense and non-defense areas, many budget accounts were exempt so that the automatic reductions only affected accounts funding just over a third of budget dollars. In the non-defense area the cuts of certain other programs, including guaranteed student loans, foster care and adoption assistance, and medicare and other health programs, were governed by a series of special rules that further limited the size of the automatic cuts for these accounts. Once these exceptions were considered, the size of the reductions within the defense and non-defense accounts was determined by cutting a uniform percentage from outlays for the PPAs in each category (defense and non-defense). Differences between CBO and OMB estimates of the required sequester were to be eliminated through averaging. Once a sequestration process was ordered, the President and Congress would still have until the beginning of the fiscal year (October 1) to enact legislation in order to avoid the sequestration.

The original sequestration process faced three problems: there were serious questions about its constitutionality, it threatened to violate the separation of powers by shifting the power to make decisions with significant budgetary consequences into the hands of unelected technicians, and it left many openings for evading its restrictions.

In the summer of 1986 the Supreme Court voided the automatic sequestration mechanism of Gramm-Rudman-Hollings by holding that it was a violation of the separation of powers for the Comptroller General to order the President to make budget reductions. In *Bowsher v. Synar*, the Court found that the Congress had assigned an executive function to an officer—the Comptroller General—who was subject to removal only by Con-

gress.[26] Congress was thus forced to rely on the back-up sequestration procedure that had been built into the law. For fiscal years 1986 and 1987, therefore, if sequestration was necessary, Congress would have to vote to bring it about on its own, without a third party scapegoat. Congress so voted in fiscal year 1986 when the size of the sequestration was capped at $11.7 billion. It avoided the difficulty of a vote the following year through the adoption of spending reductions and increased taxes along with a one-time benefit of $20 billion from the Tax Reform Act of 1986 and $15 billion in artificial adjustments such as asset sales, accelerated revenue collections, and shifting of outlays from one fiscal year to another. By calendar year 1987, however, these one-time savings had come back to haunt Congress, since the effect of their inclusion in the fiscal year 1987 budget was to increase the required deficit reduction for fiscal year 1988.[27] This was a major factor behind the re-design of Gramm-Rudman-Hollings in 1987.

The sequestration process also shifted a great deal of power to unelected technicians at CBO and OMB. Determination of the baseline from which the required sequestration would be calculated became a crucial policy function. Economic assumptions, spend-out rates from budget authority to outlays, determination of how to treat appropriated entitlements, and many other technical issues were left to technical agency staff. The potential budgetary impact of these decisions is great. In August 1987, for example, OMB and CBO differed by $33.1 billion in their baseline estimates. Even after attempts were made to add greater specificity to the baseline and to eliminate many discretionary factors in the Reaffirmation Act, CBO's Assistant Director for Budget Analysis, James Blum, had to inform the Reaffirmation Act conferees that reasonable technicians could still differ by $20 to $25 billion.[28] This was unwelcome news to politicians attempting to create a sequestration process that could bring about a $23 billion deficit reduction that very year.

---

[26] 478 U.S. 714 (1986). *See* CONGRESSIONAL RESEARCH SERVICE, THE LIBRARY OF CONGRESS, SUMMARY AND ANALYSIS OF THE RAMIFICATIONS OF *BOWSHER V. SYNAR*, THE GRAMM-RUDMAN-HOLLINGS DEFICIT REDUCTION ACT CASE (1986).

[27] The deficit reduction targets of Gramm-Rudman-Hollings were absolute levels. Thus, spending or tax increases that had only a one-year effect would not help Congress in the next budget: in the following year, Congress would have to make the additional deficit reductions required by Gramm-Rudman-Hollings plus make the reductions that it had been spared from making in the prior year.

[28] Interview with James Blum (Dec. 21, 1987).

This discretion also gave the executive branch and Congress the opportunity to create baselines and resulting sequestrations that favored their policy choices. Forty-one percent of the $33.1 billion difference between the 1987 OMB and CBO estimates was due to differing economic assumptions, with the remainder due to conceptual and technical differences. So pervasive was the congressional belief that OMB was manipulating its numbers that the initial Senate rewriting of Gramm-Rudman-Hollings in 1986 attempted to specify the values of individual economic variables, in effect tying the Administration's hands.

C. *Timetable*

Gramm-Rudman-Hollings modified the congressional budget process timetable. Under the new law, the President was required to submit his budget on the first Monday after January 3. The reports by the various committees of each chamber to their budget committee were required on February 15, a month earlier than before. The completion of the budget resolution process was also moved up by a month. If the budget process was not completed by May 15, appropriations bills could come to the floor of the House. Reconciliation action was to be completed by June 15, and House action on the thirteen annual appropriations bills was to be completed by June 30.

This process was put in place so that all spending and tax action would be included in the baseline of the initial CBO and OMB reports that were due on August 15. But just the opposite occurred. Since Gramm-Rudman-Hollings sequestrations would be made from the baseline of enacted legislation, it was in everyone's interest to enact spending and tax legislation after the baseline had been determined. As a result, since the passage of Gramm-Rudman-Hollings the number of enacted annual appropriations bills has dropped to zero. Although the Reaffirmation Act attempts to remove this unintended incentive by pro-rating sequestration and granting credit when the appropriations level is below the baseline level of the final sequestration order, many observers of the process feel that it is still in the interest of the appropriations committees to hold back their bills until after sequestration has taken effect.[29]

[29] Interview with unnamed source (Dec. 21, 1987).

IX. Events Leading to the Adoption of the Reaffirmation Act

The adoption of Gramm-Rudman-Hollings did not significantly change the politics of the budgetary process; the same forces that created the deadlock that led to Gramm-Rudman-Hollings were still present after its enactment. When the Supreme Court declared the automatic sequestration procedures unconstitutional, the need for voluntary political compromise remained and the same forces that prevented such a compromise prior to 1985 were still present. Thus, it is not surprising that in the two years of Gramm-Rudman-Hollings' implementation, Congress enacted few long-term deficit reductions and was even more delinquent in meeting its budgetary timetable.

Because of the $33 billion in one-time deficit reductions that eliminated the need for sequestration in calendar year 1986, Congress faced the need for even larger deficit reductions for the following budget cycle. With the 1988 election approaching, therefore, a way was needed—at a minimum—to rewrite the deficit limits of Gramm-Rudman-Hollings. The Senate passed such a rewrite in 1986, but the House refused to act. All participants were aware, however, that another debt limit extension would be needed by May 15, 1987. Thus, all parties expected Gramm-Rudman-Hollings to be on the legislative agenda in 1987.

Most of the same actors were again involved, although this time all the participants had more time to develop their proposals. Once again Senator Gramm and other conservative Republicans saw Gramm-Rudman-Hollings as a means to reduce federal domestic spending as well as to close the deficit. House conservative Democrats—led by Representative Buddy MacKay (D-Fla.)—continued to focus on the deficit and were willing to accept many policy compromises to achieve deficit reduction. The House leadership and liberal Democrats more than ever saw a rewrite of Gramm-Rudman-Hollings as a lever to force President Reagan to come to the bargaining table. The Senate Republican leaders continued to want to play the "responsible" (but politically dangerous) role of cutting popular programs, but they had become increasingly disenchanted with the ability of a Gramm-Rudman-Hollings process to bring about such budgetary responsibility; many of them supported the re-

write only as a vehicle for attaching further budget reform proposals.

To this mix was added one important additional actor—Chairman Dan Rostenkowski of the House Ways and Means Committee. Prior to the 1987 rewrite of Gramm-Rudman-Hollings, Congress enacted a fiscal year 1988 budget resolution that required a $36 billion reduction in the deficit. As part of this package, the tax committees were instructed to report out tax bills that would raise $19 billion in fiscal year 1988. This put Chairman Rostenkowski in a very difficult position, for he knew that raising $19 billion in revenue would require real tax increases (increases greater than those from the extension of taxes that were scheduled to expire and other small items), and that as long as President Reagan opposed any real tax increase, the Ways and Means Committee Chairman would face the probability of being badly beaten on the floor of the House. Chairman Rostenkowski needed a rewrite of Gramm-Rudman-Hollings to lower his tax target from $19 billion to a figure (under $10 billion) that could be raised without significant new or increased taxes.

Democratic Congressional leaders, represented by House Majority Leader Tom Foley, were better prepared this time around. Their goal was to create a rewrite that would put the ball—and hopefully the political cost—in the President's court.

Conservative Republicans were placed in a difficult position. On the one hand, they were in favor of any plan that would lead to reductions in domestic spending and the size of the deficit. But they also faced the fact that they were entering an election cycle with a politically weakened President, in part because of the Iran-Contra Affair. In private, therefore, they—and the White House—wanted a Gramm-Rudman-Hollings rewrite that would appear to lead to deficit reduction but in actuality would require minimum budget action in the two remaining budget cycles of the Reagan presidency. Senator Gramm and the White House, therefore, proposed going from a current law baseline (where only those programs where inflation adjustments are required by law are inflated) to a current policy baseline (where only those programs where inflation adjustments are prohibited by law are not inflated). The Reaffirmation Act appears to be a two-year solution to get the political process through the Reagan era.

This constellation of forces in Congress led to three major proposals. Congressman Foley's contained a "soft trigger." Under his plan, CBO would send its sequestration report to the Comptroller General who, in turn, would "advise" the President that a sequestration was needed. The President would then have the choice of either invoking the sequestration or explaining why he refused to do so.[30]

The Foley plan was clearly a political gambit. Its goal was to place the onus on the President. Even among House Democrats, it never gained much support, because many legal observers questioned whether it would be constitutional following the Court's ruling in *Bowsher v. Synar*,[31] and because it did not address Chairman Rostenkowski's need to lower the revenue reconciliation from $19 billion to a number that could be reached without a major tax increase. Although it was discussed and debated in the House Democratic caucus, it was never introduced as a formal bill.

Chairman Rostenkowski championed the "hard trigger" approach. The Rostenkowski plan involved: (1) a rewrite of the deficit limits so that the tax bill of 1987 would not exceed $10 billion, (2) a revision of the sequestration procedure to place it fully in the President's hands, (3) an attempt to close all known loopholes that could be used by the President and OMB to avoid across-the-board cuts, and (4) a reliance on disclosure and the resulting political fallout to keep the President and OMB in line.[32]

The Senate plan became a joint effort of Senators Gramm and Chiles (D-Fla.).[33] This plan also called for a "hard trigger," but initially it was a two-year rather than a five-year effort. It also contained a variety of budget process "reforms" championed by Senator Pete Domenici and other Senate Budget Committee members. Most of these "reforms" were stripped from the bill in conference.

---

[30] *See generally* M. Rosenberg, Fixing Gramm-Rudman-Hollings: The Foley Proposal (Congressional Research Service, American Law Division, July 13, 1987) (unpublished manuscript).

[31] Bowsher v. Synar, 478 U.S. 714 (1986).

[32] Interviews with members of the House Ways and Means Committee staff (Dec. 22, 1987).

[33] Chiles also proposed a separate plan that would have mandated specific values for a variety of variables that are used in deriving economic assumptions and baselines. Interview with Rick Brandon, Staff Director, Senate Budget Committee (Dec. 22, 1987).

The drafting of the Reaffirmation Act followed a pattern similar to that of Gramm-Rudman-Hollings. The real work, once again, occurred in conference. Chairman Rostenkowski, in an unusual move for a committee chairman, supported a successful effort offered by House Minority Leader Robert Michel (R-Ill.) on the House floor to instruct the House conferees to support a "hard trigger." There is general agreement that the final bill owes more to the desires of the House than to those of the Senate.

The conference agreement implied the possibilities of a large reduction in defense procurement, a significant tax increase, or both. A battle took place within the administration over whether the President should sign the measure. Faced with the need to show his support for deficit reduction in a time of uncertain credit markets and the fact that at a minimum the rewrite would eliminate the need for seeking another rise in the debt limit during the remainder of his term, President Reagan signed House Joint Resolution 324 (the Reaffirmation Act) into law on September 29, 1987.[34]

The view has been expressed that the Reaffirmation Act should be seen as a two-year law.[35] Everyone expects the next President to be a player at the budget bargaining table. The participants in the rewrite process continue to see the Reaffirmation Act as a mechanism to avoid short-term problems rather than as a fundamental procedural change needed to solve the nation's deficit dilemma.

After the Reaffirmation Act was enacted, the stock market crashed on October 19, 1987. This event, rather than the Gramm-Rudman-Hollings rewrite, finally brought the Administration to the budget bargaining table. But the resulting budget summit did not lead to actions that will solve the long-term deficit problem. In fact, when one-time savings are excluded, the size of the deficit reduction in the summit agreement is equal to the size of the sequestration that it avoided. If anything, the results of the summit agreement suggest that the size of the sequestration determines the upper limit of possible deficit reduction in a given year.

---

[34] The Balanced Budget and Emergency Deficit Control Reaffirmation Act of 1987, Pub. L. No. 100-119, 101 Stat. 754.

[35] *See supra* note 1.

X. The Difficulty of Constructing a Semi-Autonomous Doomsday Machine

The difficulty with expenditure limitations and balanced budget amendments is that they seek to cap the effects of politics rather than change the incentives that cause those effects. A major claim by the advocates of Gramm-Rudman-Hollings is that it does affect political incentives. According to its proponents, this occurs in three ways. First, because of the $10 billion leeway, smaller deficit reductions are required (at least in the short-run) when Congress enacts its own priorities than when sequestration occurs. Second, when Congress enacts its own package of deficit reduction measures, it can include tax increases, thus reducing the size of the needed spending reductions. Third, when enacting its own reduction package, Congress can impose its own priorities on the budget cuts.

Unfortunately, these incentives do not compensate for the perceived advantages of creative accounting and smoke and mirrors. So far, in Gramm-Rudman-Hollings' three years of implementation, Congress has continued to come up with creative new ways to avoid the full impact of deficit reductions. The response of the drafters of Gramm-Rudman-Hollings has been to write provisions into the Reaffirmation Act prohibiting as many end-runs as can be identified. But the one-time reductions in last fall's budget summit agreement suggest that Congress cannot make itself do something neither it nor its constituents want it to do.

The design of Gramm-Rudman-Hollings also illustrates the difficulty of designing a credible semi-autonomous doomsday machine. Once a doomsday machine is semi-autonomous, all participants in the game know that it will not be invoked if it is truly too terrible and threatening. This failure was present in Gramm-Rudman-Hollings. Under the original Act, after the first year, sequestration would occur equally in all defense PPAs. This provision would have led to massive troop-level reductions, since Gramm-Rudman-Hollings' sequestration levels would be measured in net outlays for each PPA and, because of termination costs, many additional troops would have to be let go in order to achieve the appropriate outlay savings. All participants had to know that any President would not fire a quarter of the Army, Navy, or any other military branch just to achieve an arbitrary deficit reduction target. As originally designed, there-

fore, Gramm-Rudman-Hollings was too threatening. In order to solve this problem the revision of Gramm-Rudman-Hollings has given the President leeway to shift cuts in military personnel accounts to other defense accounts.

At the other extreme, a semi-autonomous doomsday machine can be too weak. This was clearly the case in 1985 (fiscal year 1986) when, rather than create the incentives to bring all parties to the bargaining table, the initial round of sequestration was accepted by Congress as the lesser of two evils.

The Gramm-Rudman-Hollings experience so far would seem to indicate that the level of sequestration should be viewed as the maximum attainable deficit reduction. Even with the stock market crash of October 19, 1987 acting as a spur, the summit agreement appears to be about as large as the automatic sequestration would have been. What is clear so far is that no one has been able to devise a solution that would dictate the appropriate sequestration size that would bring all the reluctant parties to the table. The solution to the deficit dilemma, therefore, will arise from changes in the policy positions of the President and the working majorities of the House and the Senate rather than from changes in budget procedures.

TABLE I
DEFICIT TARGETS OF GRAMM-RUDMAN I AND II
(In Billions of Dollars)

	Gramm-Rudman I		*Gramm-Rudman II*	
Fiscal Year	*Original Target*	*Year-to-Year Decrease*	*Revised Target*	*Year-to-Year Decrease*
1986	171.9	—		
1987	144	27.9		
1988	108	36	144	—
1989	72	36	136	8
1990	36	36	100	36
1991	0	36	64	36
1992			28	36
1993			0	28

Source: Davis & Keith, Debt-Limit Increase and 1985 Balanced Budget Act Reaffirmation: Summary of Public Law 100-119 (H.J. Res 324) 15 (Congressional Research Service, Report No. 87-865 GOV, October 29, 1987).

GRAMM-RUDMAN-HOLLINGS: AN IMPERFECT LAW THAT WORKS

ROBERT W. KASTEN, JR.*

The Balanced Budget and Emergency Deficit Control Act of 1986, otherwise known as the Gramm-Rudman-Hollings Act,[1] has been roundly denounced as a mere fiscal fig leaf, a gimmick contrived in the vain hope that legislators can postpone indefinitely the tough fiscal policy choices required by the federal budget deficit. These criticisms come from all points on the ideological spectrum, from conservatives who fear rollbacks in the defense buildup to liberals claiming that the Gramm-Rudman-Hollings process will savage discretionary social spending. Others charge that Gramm-Rudman-Hollings has precipitated a breakdown in the legislative process, resulting in government by stop-gap spending measures and continuing resolutions.

Yet despite its weaknesses, the Gramm-Rudman-Hollings Act has worked. The Act has imposed a measure of fiscal discipline in the budget process that otherwise would not exist. Since the enactment of Gramm-Rudman-Hollings in 1986, the annual growth of government spending has declined from double-digit rates in the late 1970's and early 1980's to the lowest rate in two decades.[2] As a consequence, the federal budget deficit has declined significantly.

The continued success of the Gramm-Rudman-Hollings Act depends on the mix of economic policies used to reduce the deficit. In addition, the Congressional budget process must be further reformed to restore some measure of budgetary power to the Executive Branch. The erosion of the President's role in budgetary matters in recent years has contributed to the breakdown of the budgetary process and the subsequent rise in public spending.

---

* Member, United States Senate (R-Wisc.). B.A. University of Arizona, 1964; M.B.A. Columbia University, 1966. Senator Kasten is a member of the Senate Budget and Appropriations Committees.

[1] The Balanced Budget and Emergency Deficit Control Act of 1986, Pub. L. No. 99-177, 99 Stat. 1037 [hereinafter Gramm-Rudman-Hollings]. The Act was amended by the Balanced Budget and Emergency Deficit Control Reaffirmation Act of 1987, Pub. L. No. 100-119, 101 Stat. 754.

[2] *See* CONGRESSIONAL BUDGET OFFICE, REPORT TO THE SENATE AND HOUSE COMMITTEES ON THE BUDGET, THE ECONOMIC AND BUDGET OUTLOOK, FISCAL YEARS 1989–1993 36 (1988) [hereinafter 1989–1993 BUDGET OUTLOOK.]

HISTORICAL PERSPECTIVE ON THE BUDGET PROCESS

A brief history of the Congressional budget process will help illuminate the implications of the Gramm-Rudman-Hollings Act.

The Constitution of 1787 empowered the President to veto spending bills passed by Congress. However, the boundaries of presidential discretion over the expenditure of appropriated funds were not clearly defined by the Constitution's framers. From 1789 until the collapse of the Nixon Presidency, the President retained the power to "impound" (refuse to spend) money appropriated by Congress. Congress had no formal power to overturn this action. Presidents exercised this impoundment power (along with the veto) to check Congress' tendency to overspend. For example, Presidents Kennedy and Johnson used impoundments to reduce projected Congressional spending by 6% and 5.4% during their respective Administrations.[3] For the most part, Congress grudgingly assented to these impoundments.

President Nixon pushed impoundment powers to their outer limits by reducing and terminating federal programs against the explicit wishes of Congress. In response to Nixon's alleged abuse of his impoundment authority, Congress passed the Congressional Budget and Impoundment Control Act of 1974.[4] This Act eliminated the presidential prerogative of impoundment and replaced it with weaker substitutes. The 1974 Budget Act also overhauled the entire budget process to promote planning and coordination in fiscal policy by creating the Congressional Budget Office and the House and Senate Budget Committees, which were charged with developing an overall budget plan (called a "budget resolution") that would serve as a guide for the Appropriations Committees of each chamber.

The 1974 Budget Act has not worked well. Since 1974, federal spending has almost quadrupled, from $269 billion to over $1 trillion.[5] According to Dr. William Orzechowski, an economist with the United States Chamber of Commerce, Congressional spending has surpassed its own budget resolutions by an average of $25 billion per year. Congress has exceeded its deficit goals over the same period by an average of $48 billion.

---

[3] Unless otherwise indicated, all data for this Essay have been taken from sources on file with the Minority Staff of the Senate Budget Committee.

[4] Pub. L. No. 93-344, 88 Stat. 97 (1974).

[5] *See* 1989–1993 BUDGET OUTLOOK, *supra* note 2, at 136.

The 1974 Budget Act concentrated budget power in the Congress by severely limiting presidential impoundments. In effect, the 1974 Budget Act removed any effective external check on Congress' tendency to overspend. Congressional use of multi-billion dollar omnibus appropriations bills has further diluted the President's control over federal spending. The unseemly combination of omnibus appropriations and restricted presidential impoundments has been an underlying cause of the growth in federal spending and deficits.

Frustration with the budget process in addition to the specter of skyrocketing expenditures and rising deficits in the early 1980's led to the enactment of the Balanced Budget and Emergency Deficit Control Act of 1986. Gramm-Rudman-Hollings was approved with wide bi-partisan support in the Senate by a vote of sixty-three to thirty-six. This Act was an attempt to provide a procedural framework to eliminate deficits over five years by restoring greater impoundment authority to the President.

Gramm-Rudman-Hollings established fixed ceilings on the annual federal budget deficit and reformed several legislative procedures to enhance the likelihood that Congress would abide by the annual ceilings. Gramm-Rudman-Hollings created an "action-forcing" mechanism (the "sequester") which would impose upon the President a duty to make automatic across-the-board spending reductions if Congress failed to meet its deficit target. In the event of a sequester, federal spending programs would be reduced by a uniform percentage. Defense and certain domestic programs would be reduced equally to sustain these budget cuts.

Since 1986, the Gramm-Rudman-Hollings Act has been revised to ensure the constitutionality of the sequestering process, and the deficit targets have been moved forward one year. The ultimate goal of this Act is a balanced budget by fiscal year 1992.

THE ECONOMIC AND POLITICAL IMPACT OF GRAMM-RUDMAN-HOLLINGS

Thus far, Gramm-Rudman-Hollings has been a qualified success. It has significantly reduced the growth rate of federal spending. From 1980 to 1985, federal expenditures grew at an

annual rate of 10.5%.[6] Federal spending as a percentage of Gross National Product (GNP) reached an historic high of 24% in 1985.[7] Without Gramm-Rudman-Hollings, the prospect for meaningful deficit reduction was grim.

Two years of Gramm-Rudman-Hollings have reversed these alarming trends. In 1986, the sequestering mechanism cut $11.7 billion from the federal budget. Congress further facilitated the budget reduction by resurrecting the 1985 reconciliation bill which saved another $11.5 billion. As a result, Gramm-Rudman-Hollings helped reduce spending growth to 4.6% in fiscal 1986.[8] This favorable trend continued in fiscal 1987, as spending growth slowed to 1.2%—the smallest annual increase in over two decades.[9] Gramm-Rudman-Hollings helped achieve the largest single-year reduction in the federal budget deficit in history—from $221 billion in 1986 to $148 billion in 1987.[10]

The Act contains a "deficit-neutral" provision which requires offsetting spending reductions or tax increases to pay for new spending above the budget resolution's outlay levels. This provision strengthens the incentives for fiscal responsibility. Sixty votes are required in the Senate to waive this deficit-neutral provision. The stringent voting requirement has blocked the consideration of several budget-busting bills on the Senate floor.

Gramm-Rudman-Hollings' deficit-neutral provision has improved the political dynamics among members of Congress and special interest groups. Amendments to increase spending for specific programs are becoming more difficult to pass, as competing interest groups emerge to oppose the required offset to the amendment—the spending cut or tax increase. In the past, Congress simply passed on the higher costs by increasing the deficit, engendering little interest group opposition.

Indeed, the political magic of Gramm-Rudman-Hollings is that it forces interest groups to compete against one another rather than against the amorphous and poorly represented body of taxpayers. Gramm-Rudman-Hollings restores the taxpayer's power in the budget process.

---

[6] *See* 1989–1993 BUDGET OUTLOOK, *supra* note 2, at 138.
[7] *See id.* at 141.
[8] *See id.* at 136.
[9] *Id.*
[10] *See id.*

Nevertheless, the Act has had some undesirable effects. It has contributed to the increased use of budget gimmickry to achieve the required deficit savings. These gimmicks include shifting payments from the end of one year to the beginning of the next, underestimating program costs, and inflating projected budget baselines to soften the impact of a sequester. In addition, Congress has achieved immediate savings at the expense of increasing the deficit in the out-years, thereby exchanging short-term savings for long-term increases in the deficit.

Gramm-Rudman-Hollings creates an incentive to reduce the deficit, but it also creates a dangerous temptation to increase taxes rather than reduce spending. Taxes can be increased to help meet the annual deficit ceiling, but increased taxation may dampen economic activity and lead to collection of lower revenues. Therefore, a tax increase is not an effective solution to reducing the deficit.

It is vitally important for policymakers to recognize the relationship between fiscal policy and the productive economy when deciding how to meet the Gramm-Rudman-Hollings deficit reduction goals. Failure to adhere to these economic fundamentals can lead to policies that increase the budget deficit and reduce economic opportunity.

Perhaps the greatest drawback of Gramm-Rudman-Hollings is that it focuses on the size of the budget deficit and diverts attention from the true threat of excessive government spending. A growing body of evidence from around the world shows that rising levels of government spending inhibit economic growth. High levels of public spending and correspondingly high taxes absorb private resources and diminish productive economic activity.

Statistics from industrial countries over the past twenty years reflect the inverse relationship between government size and economic growth. The government share of Gross Domestic Product (GDP) in the seven major Western European countries rose from an average of 32.6% in the years 1965–1969 to 41.2% during the years 1980–1984. Not surprisingly, economic growth in these nations fell over the same period from an average of 5.0% to 1.8%. While the United States is fortunate to have a total public sector size of about 35% of GDP, we must reduce the public sector's claim on private resources even further in order to improve the prospects for long-term economic growth.

THE 1987 BUDGET COMPROMISE

The Gramm-Rudman-Hollings process had mixed results in 1987. The budget compromise of last fall proved that while the Act has been an effective tool for fiscal restraint, Congress retains the ultimate responsibility for making it work.

With the Democrats in control of the Senate in 1987, many conservatives feared that Gramm-Rudman-Hollings was going to be used primarily as a vehicle to raise taxes, contrary to the intent of one of its original sponsors, Senator Phil Gramm (R-Tex.). These fears were realized when the Senate Budget Committee approved a budget resolution which called for the largest single-year tax increase in history, $18.5 billion in fiscal 1988 and a whopping $98 billion over the next four fiscal years. Under this budget plan, total federal spending would grow 4.2% in fiscal 1988, more than twice the spending growth in fiscal 1987.[11]

Faithful to his 1984 campaign pledge, President Reagan vowed to veto any tax hike. He knew that he could count on receiving the thirty-four Senate votes necessary to sustain a veto. The President indicated that he would accept a Gramm-Rudman-Hollings sequester before he would sign a tax increase.

Gramm-Rudman-Hollings was working. Congressional tax hikers were on the defensive. It seemed that Congress was about to achieve deficit reduction the right way, by reducing expenditures. Then came "Black Monday." On October 19, 1987, the Dow Jones Industrial Average plunged 508 points. Conventional wisdom put the blame for the market panic on Congressional inaction on the federal budget deficit, and this breathed new life into the "deficit reduction at any cost" orthodoxy in Congress. Contrary to conventional wisdom, the deficit was not the primary cause of the stock market crash.

The budget deficit had declined over 30% over the year preceding Black Monday, from $221 billion in fiscal 1986 to $150 billion in fiscal 1987.[12] When state and local government finances are included, the general government deficit as a percentage of GNP dropped to a 6-year low in 1987—3.4% of GNP.[13] Thus, the deficit was shrinking in the years leading to the market crash.

---

[11] *See id.* at 67.
[12] *See id.* at 136.
[13] *See id.* at 137.

Other macroeconomic factors influenced the October market plunge. These factors include sluggish economic growth in Western Europe, lack of coordination on international exchange rates, and fear of tax increases and protectionist legislation. Another important cause of the crash was the Federal Reserve's restrictive monetary policy throughout 1987. Nonetheless, hysteria over Black Monday led to an agreement between the President and Congressional leadership on a budget plan that proposed to reduce the budget deficit by $76 billion over two years.

The savings were derived primarily from defense cuts and tax increases. There were virtually no substantive reductions in non-defense spending. Most of the non-defense savings consisted of accounting gimmicks, one-time asset sales, or savings that had already been achieved by Congress. Federal spending in fiscal 1988 would increase by $50 billion over fiscal 1987.

Many policymakers stigmatized the sequestering process as the least desirable cure for the ailing market because, they claimed, it represented a "failure" in political leadership. This argument was used to force the President to accept a tax increase and avoid any real progress on reducing the growth of government spending. Their contention is simply untrue. A sequester is better for the economy than a tax increase, and the stock market would have responded positively to a sequester, interpreting it as a positive indicator.

First, a sequester is a permanent reduction in the spending base. Unlike the tax increases in the budget compromise, a sequester would release $23 billion in economic resources to the private sector. American businesses would have $23 billion more with which to finance domestic investment and spur economic growth. It is highly unlikely that the budget compromise will yield even $10 billion in reduced demands on the private sector.

Second, the markets did not react against the $11.7 billion sequester that occurred in fiscal 1987. It is hard to imagine how a sequester of $23 billion would have caused a stock market tailspin. The markets are looking for real results—not agreement for agreement's sake. They are looking for reduced pressure in the capital markets, and a sequester would have done the job.

Because the budget compromise failed to restrain spending and raised taxes, the fiscal 1988 budget deficit is likely to be much higher than the $150 billion dictated by the budget agreement. As an unfortunate consequence, the growth rate of spend-

ing will rise in 1988 after three consecutive fiscal years of decline.

THE FUTURE OF GRAMM-RUDMAN-HOLLINGS AND DEFICIT REDUCTION

Far from being an act of legislative futility, Gramm-Rudman-Hollings is an imperfect law that has worked. The events of the past year, however, show that Gramm-Rudman-Hollings is just a process, a set of rules and guidelines which depend on Congressional intelligence and good faith for their successful implementation. The legislative chaos which occurred last year was a result of Congress' inability or unwillingness to adhere to rules, requirements, and deadlines of the Act. The fact that the deficit will be higher in 1988 is not a sign of the failure of Gramm-Rudman-Hollings itself, but rather of the failure of Congressional will to reduce the deficit.

To be sure, Gramm-Rudman-Hollings and the budget process need to be strengthened and improved. We need further budget reforms to correct the institutional imbalance between the executive and the legislative branches in money matters, and to enhance the political incentives for fiscal responsibility. These reforms include allowing the line-item veto, enhancing rescission powers, and enacting a host of other procedural changes to plug the loopholes in the budget process.

Congress can achieve a balanced budget by 1992 if it aims for price stability and economic growth. The budget deficit cannot possibly be reduced without continued economic growth, lower interest rates, and job creation. Congress should also consider freezing federal spending for one year. By holding spending in check, the natural growth in tax revenues resulting from economic expansion would dramatically reduce the deficit. Freezing domestic spending alone would save $20 billion in one year.[14] Congress should look to innovative, growth-oriented approaches to deficit reduction, such as a cut in the capital gains tax rate to 15%, as opposed to being taxed as ordinary income,

[14] *See* Testimony of Prof. Lawrence Lindsay of Harvard University before the Heritage Foundation's Ad-Hoc Committee on Taxation of Capital Gains (Feb. 2, 1988).

which would raise between $4 and $8 billion in federal tax revenues.[15]

Gramm-Rudman-Hollings, combined with an employment-oriented national economic policy, can substantially reduce the deficit. We should give the process a chance, and not get sidetracked into illusory and counterproductive quick fixes—like tax increases.

[15] *Id.*

IS GRAMM-RUDMAN-HOLLINGS AN EXERCISE IN LEGISLATIVE FUTILITY?

DAN QUAYLE*

The Congressional budget and appropriations process has for years been perilously close to collapse. The procedure that Congress employs to raise and spend money is inefficient and time consuming. The budgetary process has produced fiscal gridlock, and it has left little time for other legislative activity. Responding to these problems, Congress passed the Gramm-Rudman-Hollings Balanced Budget and Emergency Deficit Control Act,[1] which instituted a set of procedural changes that catalyzed significant change. The Gramm-Rudman-Hollings procedural reforms have helped to reduce the political pressures that often lead to fiscal irresponsibility and breakdown of political consensus. These and other proposed procedural reforms promise to remove many obstacles that have prevented us from prudently allocating federal resources. The evidence to date indicates that Gramm-Rudman-Hollings has succeeded. Since its enactment, the growth in nominal federal spending, an estmated 2.1 percent for fiscal year 1987, has reached its lowest level since 1965.[2] The Act's strictures have forced a reluctant Congress to make the hard political choices necessary to achieve meaningful deficit reduction.

A HISTORY OF BUDGET-PROCESS REFORM

Spending process reform has traditionally resulted from reactions to periods of fiscal crisis. Before the Civil War, the taxwriting committees—the Ways and Means Committee in the House, and the Finance Committee in the Senate—exercised jurisdiction over the federal budget. During the Civil War, the

* Member, United States Senate (R-Ind.). B.A. DePauw University, 1969; J.D. Indiana University School of Law, 1974. Senator Quayle is a member of the Budget Committee, Armed Services Committee, and Labor and Human Resources Committee.

[1] The Balanced Budget and Emergency Deficit Control Act of 1986, Pub. L. No. 99-177, 99 Stat. 1037 [hereinafter Gramm-Rudman-Hollings or the Act]. The Act was amended by the Balanced Budget and Emergency Deficit Control Reaffirmation Act of 1987, Pub. L. No. 100-119, 101 Stat. 754.

[2] Unless otherwise indicated, all data used in the Essay are drawn from sources on file with the Senate Budget Committee.

public debt grew from $90 million to $2.7 billion. Even though this debt was halved by the end of the century, the rapid growth of the national debt prompted Congress to establish two separate appropriations committees—the House Committee in 1865 and the Senate Committee in 1867—to control the federal purse strings.

Public sentiment for a national budget system swelled after World War I. During that war federal outlays skyrocketed from $700 million to $18 billion. The exhorbitant cost of financing World War I led to the passage of the Budget and Accounting Act of 1921,[3] which required the President to submit budgetary recommendations to Congress through a newly established Bureau of the Budget, now the Office of Management and Budget (OMB). The 1921 Budget Act also established the General Accounting Office (GAO), an independent, nonpolitical agency to serve Congress in investigations, audits, legal accounting, claims settlements, and to make recommendations for more effective fiscal management. Furthermore, it clarified and formally outlined the jurisdiction of the appropriations committees. The 1921 Budget Act's delegation to the President of a larger role in the budget process led to debate between those who believed that the President had been granted too much authority in budgeting and those who believed that Congress could better balance revenues and outlays if the executive branch provided recommendations and a detailed accounting of all programs.

After World War II, during an era of Congressional reform, an ill-fated Joint Budget Committee was established under the terms of the Legislative Reorganization Act of 1946.[4] This panel was comprised of members of the Senate Finance Committee, the House Ways and Means Committee, and both appropriations committees. It was charged with reviewing the President's budget and recommending its own ceilings for expenditures and appropriations. This committee reported its first budget in February of 1947 but disbanded two years later.

During the Nixon Administration, Congress faced a budget crisis that threatened to strip it of its control over government expenditures: the President refused to spend money which Congress had appropriated. While Presidential impoundment had been exercised regularly since Thomas Jefferson's Administra-

---

[3] 42 Stat. 20 (1921).
[4] 60 Stat. 812 (1946).

tion,[5] President Nixon's refusal to spend an unprecedented $18 billion (or about 7.3% of total spending) was viewed by Congress as overstepping the bounds of executive authority. In addition, Congress had become skeptical of the accuracy of the economic and budget projections made by the OMB and relied upon by the President in formulating budget proposals. It desired a more systematic process under which it could consider and control budget policy. In response to these concerns, the budget process was overhauled by the Congressional Budget and Impoundment Control Act of 1974.[6]

The 1974 Budget Act addressed the budget process concerns by the following actions: providing a procedure for rescissions and deferrals; establishing the Congressional Budget Office (CBO) to provide economic forecasts and estimates independent of the executive branch; and creating House and Senate Budget Committees to set overall revenue and spending levels for each major function of the federal government. Congress could act affirmatively rather than respond to each of the President's spending recommendations which allowed for little account to be taken of each measure's individual impact.

Under the new budget process, broad spending and revenue levels would be set for eighteen different categories so that Congress would have a fiscal benchmark from which to make policy decisions. The budgets developed by Congress under the 1974 Budget Act set recommended levels of spending, called functions, in each of these categories. However, the functional aggregate figures did not bind Congress or its committees to any policy changes. The authorization committees still exercised jurisdiction over developing programs to meet national needs, while the appropriations committee of each chamber determined the funding for those programs. The expectation that the 1974 Budget Act would allow Congress to balance the budget proved illusory. During the first ten years of the Act, the deficit increased over four-fold.

---

[5] Thomas Jefferson first withheld $50,000 intended for maintaining Navy gunboats on the Mississippi River. Presidential impoundments continued through the era of Franklin D. Roosevelt, who was the first chief executive to impound congressionally appropriated funds to accomplish public policy objectives rather than fiscal management. Roosevelt impounded $500 million that Congress had earmarked for public works projects. Presidents Kennedy and Johnson reduced federal spending through impoundments by 6% and 5.4%, respectively.

[6] Pub. L. No. 93-344, 88 Stat. 297 (1974).

GRAMM-RUDMAN-HOLLINGS

In 1985, faced with yet another fiscal crisis in the form of a deficit in excess of $200 billion and a national debt nearing $2 trillion, Congress adopted the Gramm-Rudman-Hollings Balanced Budget and Emergency Deficit Control Act.[7] A consensus was reached, although the final version of this legislation was significantly different from what the authors and original supporters had envisioned. Despite strenuous outcry from opponents, this drastic change in the way that Congress budgets and spends its money was adopted in relatively short order.

Regardless of the charges of its critics, Gramm-Rudman-Hollings is not a budgetary meat-axe. It is simply a change in the budget process. Under Gramm-Rudman-Hollings, Congress for the first time is required to look for revenues to offset additional spending items. In addition, Gramm-Rudman-Hollings has set limits above which the deficit may not rise for each of five successive years. If Congress fails to meet the deficit reductions contemplated by the targets, an automatic across-the-board reduction, known as a sequester, is triggered. Congress then has one month to enact changes sufficient to avoid a sequester, or the across-the-board cuts will go into effect. These procedural reforms force Congress to focus on reducing the size of the deficit when formulating its budgetary priorities.

The deficit has been reduced by more than $60 billion from its high point at $220 billion in fiscal year 1986. Furthermore, the discipline of Gramm-Rudman-Hollings forced the final agreement on the latest two-year $76.1 billion deficit reduction package passed in the waning hours of the first session of the 100th Congress.[8] Although it was Wall Street's "Black Monday"—when the Dow Jones Industrial average fell 508 points on October 19, 1987—that brought the leadership of both houses of Congress to the bargaining table with the Reagan Administration, it was Gramm-Rudman-Hollings that compelled negotiators to complete the task. Negotiations collapsed more than once, but the threat of a sequester's across-the-board cuts brought the negotiators back together until a final compromise was crafted.

---

[7] *See supra* note 1.
[8] Pub. L. No. 100-203, 101 Stat. 1330 (1987).

The political logic of Gramm-Rudman-Hollings is actually quite simple. It automatically provides a scenario worse than that of most proposed budget cuts. Without the threat of a sequester, it is unlikely that the spending reduction would have come to fruition. While virtually everyone will agree that the aggregate federal budget must be trimmed, few are willing to sacrifice programs that directly benefit themselves. The public's voting and interest group lobbying efforts provide a centrifugal counterweight to Congressional consensus in crafting a significant deficit reduction package. Yet Gramm-Rudman-Hollings induces both consensus and deficit reduction.

Notwithstanding increases in federal spending, the economy has seen a trend towards deregulation during the past eight years, which in turn has spurred prosperity. Paradoxically, concurrent with record budget deficits, our economy has enjoyed record growth. Employment is at an all-time high; the rate of inflation has been cut from more than 12% at the beginning of the decade to about 4% in 1987;[9] median family income is on an upward trend—reversing the decline of the early eighties; and we have achieved sixty-three months of continuous economic expansion. Such an expansion has never before been achieved during peacetime. Those benefitting from economic expansion are reluctant to suffer cuts in programs affecting them. To them, the consequences of continued budget deficits seem nebulous and distant, and the need for measures like Gramm-Rudman-Hollings seems remote.

WHAT THE FUTURE HOLDS

Gramm-Rudman-Hollings has forced Congress to take measures to alleviate the budget crisis. In order to comply with the second year provisions of the amended version of the Act, Congress faces the task of paring $36 billion from the approximately $157 billion deficit as estimated by the Congressional Budget Office.[10] In order to meet this and the remaining four years of Gramm-Rudman-Hollings targets, difficult choices must be made.

---

[9] CONGRESSIONAL BUDGET OFFICE, REPORT TO THE SENATE AND HOUSE COMMITTEES ON THE BUDGET, THE ECONOMIC AND BUDGET OUTLOOK, FISCAL YEARS 1989–1993 36 (1988).

[10] *See id.*

Budget and spending process reform may make some of the crucial choices politically easier for legislators. As Benjamin Franklin reportedly said, "Once the people find they can vote themselves money, that will herald the end of the Republic." In fact, legislators proposing elimination of programs have faced sharp campaign attacks by special interests and coalitions. In this way, Congress, and by extension the Federal Treasury, has been held political hostage. However, the threat of the Gramm-Rudman-Hollings' automatic sequester gives Congress the incentive to make difficult political choices and confront their constituents with the prospect of reduced federal subsidies.

The original Gramm-Rudman-Hollings budget reduction targets clearly have not been met. Rather than face the required cuts that were circumvented in previous years through financial "smoke and mirrors" and accounting gimmickery, Congress has resorted to revising the Gramm-Rudman-Hollings timetable and deficit targets. Still, the framework of Gramm-Rudman-Hollings remains, and the threat of a sequester continues to impose discipline on the budget process. Gramm-Rudman-Hollings insulates legislators politically by distancing them from direct criticism over budget cuts. Such insulation is a necessary ingredient to successful deficit reduction, as long as it is accomplished through our complex democratic political system.

The success of Gramm-Rudman-Hollings in insulating and disciplining Congress argues for further procedural reforms. Several years ago, I first proposed the idea of a two-year budget cycle to ease some of the time pressures that have hampered the budgeting process. I have also strongly advocated a Constitutional amendment to balance the budget and a proposal to give the President line item-veto authority. More recently, I have offered a spending reform measure that would give the President enhanced rescission authority and force Congress to scrutinize more closely many of the projects that it routinely lumps into foot-thick, multi-billion dollar omnibus spending measures known as "continuing resolutions." All of these, while no sure cure for deficit reduction, make difficult but necessary choices somewhat easier.

GOVERNING IN THE POST-LIBERAL ERA: GRAMM-RUDMAN-HOLLINGS AND THE POLITICS OF THE FEDERAL DEFICIT

RANDALL STRAHAN*

During the 1980's, "Gramm-Rudman" has become one of the most important new terms in the American political lexicon. Since its enactment in 1985, the budgeting law to which this shorthand term refers has engendered a great deal of controversy, and at times has been an important factor in national policymaking by creating the prospect of automatic spending cuts if Congress fails to meet legally-mandated deficit levels. Others in this Symposium will no doubt address the constitutional issues raised by the GrammRudman-Hollings balanced budget law and the important effects it has had on budgeting routines. The primary purposes of this essay, however, are to examine the underlying political conditions that explain the emergence of this new approach to budgeting and to assess its continuing influence in national policymaking.

The thesis of this essay is that a mechanistic budget-cutting law has assumed a prominent role in national policy debates primarily because of the inability of the nation's political leadership to define a new public philosophy for American society in the 1980's. A public philosophy entails a widely-accepted view of the role of government in society, which is grounded in an interpretation of the basic goals embedded in American political culture (primarily liberty, equality and progress). As political scientist Samuel Beer has pointed out, a public philosophy in a democratic political system "serves to give definition to problems and direction to government policies for dealing with them."[1] By the late 1970's, a liberal public philosophy that viewed the federal government as an expansive force for re-

* Assistant Professor of Political Science, Emory University. Ph.D, University of Virginia, 1986. The Author would like to thank Professor Richard F. Doner for his comments, and acknowledge research support by the Lynde and Harry Bradley Foundation.

[1] Beer, *In Search of a New Public Philosophy*, in THE NEW AMERICAN POLITICAL SYSTEM 5 n.1 (A. King ed. 1978). *See also* Ceaser, *The Theory of Governance of the Reagan Administration*, in THE REAGAN PRESIDENCY AND THE GOVERNING OF AMERICA 57–87 (1984); Heclo, *Reaganism and the Search for a Public Philosophy*, in PERSPECTIVES ON THE REAGAN YEARS 31–63 (J. Palmer ed. 1986).

dressing inequality and other ills in American society had ceased to provide the basic orienting principles for the national agenda. Yet, despite dramatic policy changes engineered by the Reagan Administration in 1981, the conservative public philosophy offered by Ronald Reagan and his supporters as an alternative to liberalism has proven only partly successful in structuring the national policy debate in subsequent years.

The years following the Reagan Administration's 1981 tax and budget cuts may be characterized (after the fashion of the times) as the *post-liberal* era in American politics. The defining characteristics of this period include the existence of unprecedented peacetime budget deficits together with the absence of a dominant public philosophy. Extremely large budget deficits have made clear definition of essential domestic and international responsibilities the principle task of governing at a time when the nation's political leadership is sharply divided over the role of government in society and America's role in the world. The difficulty of governing under these conditions has engendered a whole series of new organizational and procedural forms (including informal "gangs" and summits and increased use of bipartisan commissions), of which Gramm-Rudman-Hollings has been one of the most visible.

If, as this essay will argue, the Gramm-Rudman-Hollings Act (or Gramm-Rudman-Hollings) is primarily a symptom of an underlying disarray among Americans and their political leaders over the proper extent of governmental involvement in society and in the international sphere, it is important in a symposium of this type to consider not only the constitutional and procedural aspects of this controversial budgeting statute, but also how the political impasse of the post-liberal era developed and what the prospects are for its resolution. With the goal of placing Gramm-Rudman-Hollings within this broader political context, this essay looks briefly at the sequence of events that ushered in the post-liberal era, examines why a mechanistic budget-cutting procedure won the support of Ronald Reagan and overwhelming majorities in Congress in the mid-1980's, and considers the prospects for the continued prominence of Gramm-Rudman-Hollings or something like it in the future.

I. The Decline of Liberalism

From the 1930's until the 1970's, a liberal public philosophy dominated the national political agenda in the United States. In

its initial formulation as the rationale for Franklin Roosevelt's New Deal of the 1930's, liberalism stressed the need for the active use of the power of national government to manage the economy and to offset some of the societal inequalities created by concentrations of wealth and power. During the 1960's, however, Presidents John F. Kennedy and Lyndon B. Johnson redefined the prevailing liberal public philosophy, by expanding dramatically the range of problems considered proper objects of federal intervention. Domestic programs that came into being during the Great Society rested on the view that the federal government should not only manage the economy and provide security against inequalities in economic power, but also be held responsible for the very presence of inequality and poverty, and a whole host of other problems in society.[2]

The ambitious liberalism of the Kennedy-Johnson years began to founder in the 1970's. In part, liberalism fell prey to internal contradictions that developed as its agenda matured. Beer, for example, notes that the expansive redefinition of liberalism during the 1960's was guided in part by technocratic social science and a new participatory egalitarianism. As Beer demonstrates, these were views of governing that proved incompatible with one another and ultimately highly critical of many Great Society programs.[3] In a similar vein, James W. Ceaser has argued that the embracing of an antihierarchical or reformist view of governmental authority by many liberals in the late 1960's and early 1970's had the effect of weakening the national governmental institutions that were responsible for managing an increasingly congested liberal agenda.[4]

Others emphasized the fiscal constraints and deteriorating economic conditions in the 1970's in explaining the declining influence of liberalism.[5] Indeed, as early as 1967 the escalating costs of the Vietnam War and domestic commitments began to create sharp conflicts both within Congress, and between Congress and the President. Conflicts over budgetary control continued to escalate during the Nixon years, leading to the enactment of the Congressional Budget and Impoundment Control Act of 1974.[6] Even with the creation of the new budget process,

---

[2] *See* Beer, *supra* note 1, at 5–32.
[3] *Id.* at 15–37.
[4] *See* Ceaser, *supra* note 1, at 68–78.
[5] *See, e.g.*, T. B. EDSALL, THE NEW POLITICS OF INEQUALITY (1984).
[6] For a detailed account of the "budget wars" that led up to the 1974 budget reforms, see A. SCHICK, CONGRESS AND MONEY: BUDGETING, SPENDING, AND TAXING (1980).

and the onset of an extended period of sluggish economic growth, the entitlement programs that were the main fiscal legacy of Great Society liberalism continued to expand.

By the end of the 1970's, it was clear that liberalism no longer set the terms of the debate over federal domestic policies. Discussions of new federal programs to ameliorate inequality or other problems were all but eclipsed by debates over how to deal with persistent budget deficits, a stagnant economy, growing frustration with federal regulation and bureaucracy, and middle class "tax revolts." A shift toward a more centralized budget process in Congress in 1980 also indicated a heightened sense of the need to exercise greater control over spending generated by existing federal programs.[7] As the decade ended, liberal orthodoxy still held many adherents—both inside and outside government—but they found themselves in the unfamiliar position of debating policy proposals such as "supply-side" tax cuts that rejected many of the premises that had defined the national agenda over the previous two decades.

II. Ronald Reagan and the Conservative Alternative

Ronald Reagan campaigned in 1980 on a platform designed to offer a clear alternative to the liberalism of the 1960's and 1970's. In place of the idea of an interventionist national government as the engine of greater equality and social progress, Reagan emphasized individual liberty, private initiative, unfettered markets, and familial and local bonds as the primary means of bettering society. According to this view, the role of the federal government is limited primarily to managing a "safety net" of social programs and maintaining the ability to project American power into the world to contain the influence of the Soviet Union and support individual liberty. To realize the goals of this conservative public philosophy, Reagan called for a major buildup in defense spending and a new economic program to reduce inflation, stimulate growth, and balance the federal budget by 1984.

---

[7] On the shift to a more centralized budget process in 1980 and subsequent years see A. Schick, Reconciliation and the Congressional Budget Process (1981); *see also* Ellwood, *The Great Exception: The Congressional Budget Process in an Age of Decentralization*, in Congress Reconsidered 315–42 (1985).

The centerpiece of Reagan's economic program was a tax cut consisting of a thirty percent reduction in marginal rates on personal income over three years and reductions in business taxes through an accelerated depreciation plan. Also included were major cuts in domestic spending, reductions in government regulation of the economy, and calls for monetary restraint. In justifying a tax cut of this magnitude while simultaneously calling for balanced budgets, Reagan relied on the "supply-side" doctrine that substantial economic growth could be fostered by removing economic disincentives created by high marginal tax rates and non-productive government expenditures.[8]

As a number of observers have noted, the enthusiastic embrace of supply-side economics by candidate Reagan occurred in part because this economic doctrine linked the conservative political goal of reducing government intervention in society to large tax cuts and the promise of sustained economic growth and balanced budgets.[9] In addition to providing a politically attractive rationale for an assault on big government, however, the embrace of supply-side economics also ensured that the future standing of the Reagan Administration—and of the conservative public philosophy it championed—would depend to a considerable extent on the ability to deliver on the promise of economic prosperity. As Hugh Heclo and Rudolph G. Penner have written: "It seems fair to say that no incoming administration had ever before staked so much on a specific, comprehensive economic program."[10]

Together with the general consensus that liberal economic policies had failed in the 1970's, the outcome of the 1980 election—a decisive Reagan victory over incumbent Jimmy Carter, a Republican majority in the Senate for the first time since the mid-1950's, and a substantial Republican gain in the House of Representatives—created highly favorable conditions for enactment of the new economic program. In 1981, the Reagan Administration skillfully used the congressional budget process to achieve major changes in spending priorities. Use of the reconciliation procedure, having been strengthened by Congress

---

[8] For a more detailed examination of the various components of the Reagan economic program, see Heclo & Penner, *Fiscal and Political Strategy in the Reagan Administration*, in THE REAGAN PRESIDENCY 21–31 (F. Greenstein ed. 1983).

[9] *See id.* at 25–28; *see also* Ceaser, *supra* note 1, at 84–86; H. STEIN, PRESIDENTIAL ECONOMICS 235–62 (1985); Quirk, *The Economy: Economists, Electoral Politics, and Reagan Economics*, in THE ELECTIONS OF 1984 162–65 (M. Nelson ed. 1985).

[10] *See* Heclo & Penner, *supra* note 8, at 22.

the year before, was especially important. Through this procedure, most of the domestic spending cuts could be presented in a single package. The result was, as Allen Schick has pointed out: "Members had to vote for or against the President's program, not for or against particular cutbacks."[11] Overall, in 1981 Congress reduced nondefense spending around $45 billion for fiscal year 1982, while increasing the authority for new defense expenditures by a real rate of approximately ten percent.[12]

A modified version of the President's tax proposal was also enacted in August 1981. It was estimated that by 1984, when all of its rate reductions were implemented, the Economic Recovery Tax Act of 1981 would reduce annual federal revenues by approximately $160 billion. By indexing tax rate brackets for inflation, the act promised to reduce future revenues even further. Without indexing, federal revenues had increased from year to year not only from a larger GNP due to economic growth, but also from "bracket creep," where inflation increases taxpayers' income, pushing them into higher and higher tax brackets without a real increase in income. With indexing, the tax base would no longer yield these revenue increases from inflation as it had in the past.

The one area in which spending changes proposed by the Administration were notably unsuccessful was Social Security. Despite the program's impending financial crisis, a proposal to reduce Social Security spending by $82 billion over five years met almost universal condemnation from congressional Democrats and found little support among Republicans. As a result, the Reagan Administration withdrew its proposal in September 1981 calling instead for the creation of a bipartisan commission to develop an alternative plan. A Social Security package including benefit reductions and revenue increases was finally enacted in 1983, but not before congressional Democrats enjoyed considerable success in portraying Republicans as insensitive to the needs of politically-powerful elderly voters who were the program's chief recipients.[13]

---

[11] Schick, *How the Budget Was Won and Lost*, in PRESIDENT AND CONGRESS: ASSESSING REAGAN'S FIRST YEAR 26 (N. Ornstein ed. 1982).

[12] Calculations of the effects of budget changes in 1981 and later years depend heavily on the assumptions that are used. *See* Ellwood, *The Size and Distribution of Budget Reductions*, in REDUCTIONS IN U.S. DOMESTIC SPENDING 33–70 (1982).

[13] For a detailed account of the politics of Social Security reform, see LIGHT, ARTFUL WORK (1985).

By the end of 1981 the basic elements comprising the Reagan economic plan were in place. But contrary to the promises of the election campaign, the immediate aftermath was a major economic recession. With the onset of the recession in late 1981 came spiralling deficit projections. After this brief period in which the new conservative ideas about governing appeared to dominate national policymaking, massive budget deficits signalled the onset of a new pattern in American politics.

III. DEFICIT POLITICS, 1982–1984

In February 1982, the Congressional Budget Office projected that budget deficits under existing policies would exceed $100 billion in fiscl year 1982 and $150 billion in fiscal year 1983. Even with major policy adjustments, during these years the actual deficits ran even higher: $127 billion in fiscal year 1982; $208.9 billion in fiscal year 1983; and $185.3 billion in fiscal year 1984. Despite these fiscal pressures, the President's response in 1982 and for the remainder of his first term was to reassert the basic tenets of a conservative public philosophy and its associated budgetary priorities. The President stated in his February 1982 budget message:

> . . . [O]ur budget deficits will be large because of the current recession, and because it is impossible in a short period of time to correct the mistakes of decades. But our incentive-minded tax policy and our security-based defense programs are right and necessary for long run peace and prosperity, and must not be tampered with in a vain attempt to cure deficits in the short run.[14]

Consistent with this position, budgets submitted by the Administration between 1982 and 1984 requested real increases in defense spending of over ten percent and proposed no basic changes in tax policy or major revenue increases. Domestic programs bore the brunt of any deficit reduction efforts that were endorsed by the White House as each budget cycle began. After 1981, it became clear that the coalition which allowed the President to dominate the budget process during his first year could not be reassembled. The 1982 Congressional elections strengthened the hand of the House Democratic leadership, one

[14] Budget Message of the President, WEEKLY COMP. OF PRES. DOC. 134 (Feb. 8, 1982).

of the primary remaining outposts of unreconstructed liberalism, and not even the Republican leadership in the Senate was willing to embrace the President's budget priorities.

Although deficit politics worked to undermine the coalition in Congress that had supported Ronald Reagan's conservative agenda in 1981, divided partisan control of the two chambers and divisions within the two parties over how to deal with deficits made it extremely difficult for Congress to formulate coherent alternatives independent of the White House. As Charles O. Jones observed, the consequences for the remainder of Reagan's first term were ". . . the gradual disintegration of the budgetary process and the emergence of scores of members in business for themselves."[15] The reconciliation procedure in the budget process did provide the framework for a package of tax increases and spending reductions engineered by Republican Senators Pete Domenici (R-N.M.) and Robert Dole (R-Kan.) in 1982, but major revenue increases mandated by the reconciliation procedure in 1983 failed to be enacted. In 1984 the budget process was immobilized for months due to inter-chamber differences over defense spending, although Dole and House Ways and Means Committee Chairman Dan Rostenkowski (D-Ill.) took the lead in formulating a new round of tax increases and domestic spending cuts. As the 1984 elections approached, conflict over spending measures included in a massive continuing resolution delayed adjournment and produced a brief shutdown of executive agencies before agreement between the House, the Senate, and the White House could be reached.

After 1981, the three major institutions involved in federal budgeting—the White House, the House of Representatives, and the Senate—fell into a recurring pattern of budgetary politics. The Reagan White House refused to initiate any basic changes in policy direction in response to growing deficits, House Democratic leaders sought to protect domestic programs by proposing increased revenues and a slowdown in the pace of defense spending, and Senate Republicans—although generally sympathetic to the President's goals of strengthening defense, restraining domestic spending, and keeping taxes low—were often cast in the difficult role of seeking out grounds for compromises needed to head off even higher deficits. The en-

---

[15] *See* Jones, *A New President, A Different Congress, A Maturing Agenda*, in THE REAGAN PRESIDENCY AND THE GOVERNING OF AMERICA 271 (1984).

actment of substantial deficit reduction packages in the election years of 1982 and 1984 reflected a general consensus that large deficits were both economically and politically undesirable, but the divergent views of the role of the federal government held by the President on the one hand and congressional liberals on the other, made agreement on a more comprehensive program of deficit reduction impossible to achieve.

As the 1984 election approached, Congressional Budget Office forecast deficits in the $200 billion range for years to come. Any hopes that the election might produce a breakthrough in the established pattern of deficit politics proved unfounded, as the outcome reinforced rather than resolved the political stalemate of the post-liberal era. This situation encouraged a new group of political entrepreneurs in the Senate to develop a strategy to force compromise between the parties in this unresolved debate over the nation's public philosophy.

IV. THE 1984 ELECTION

During the 1984 Presidential election, Democratic nominee Walter Mondale attempted to focus the campaign debate on a specific program of deficit reduction. In his acceptance speech at the July Democratic Convention, Mondale stated:

> If this Administration has a plan for a better future, they're keeping it a secret. . . . Whoever is inaugurated in January, the American people will have to pay Mr. Reagan's bills. The budget will be squeezed. Taxes will go up. . . . Let's tell the truth. It must be done. Mr. Reagan will raise taxes and so will I. He won't tell you, I just did.[16]

On September 10, 1984 Mondale outlined a detailed deficit reduction plan which included slowing growth in defense spending to three to four percent after inflation, modest cuts in some domestic programs, and tax increases on corporations and middle and upper income households that altogether would yield around $85 billion in new revenues by fiscal year 1989.[17]

Ronald Reagan, on the other hand, avoided discussion of a specific deficit reduction program during the campaign. In re-

---

[16] *Mondale Accepts Presidential Nomination*, CONG. Q. WEEKLY REP., July 21, 1984, at 1793.

[17] For the details of Mondale's deficit reduction proposals, see Fessler & Tate, *Economics: Spotlight on the Budget Deficits*, CONG. Q. WEEKLY REP., Sept. 15, 1984, at 2251–59.

sponse to the Mondale plan, the President reiterated his proposals for constitutional amendments to require balanced budgets and create a line item veto, and continued to emphasize economic growth and restraint in domestic spending as the primary means of reducing deficits. In September, Reagan indicated that his economic program in the second term would be a "continuation of what we've been on."[18]

In response to Mondale's challenges, Reagan made two specific pledges during the campaign. First, he stated that tax increases of any sort would be acceptable only as a "last resort" and that any attempt by Congress to raise income tax rates on individuals would be vetoed.[19] Second, in response to Mondale's charge in the October 7, 1984 debate that the Administration planned additional cuts in Social Security, Reagan promised never to cut the benefits of the program's recipients. In this way, the President took a clear stance against two approaches that were likely to be important for any long-term deficit reduction plan.

The outcome of the 1984 election was an impressive landslide victory for Ronald Reagan but resulted in only a modest gain of fifteen seats for Republicans in the House and a net loss of two seats for the President's party in the Senate. Not surprisingly for a period when the nation was at peace and the economy had rebounded strongly from the 1981–1982 recession, the voters in 1984 chose continuity over change at the national level. From the perspective of deficit politics, however, the election reaffirmed a configuration of power in Washington where the budget priorities implied by Ronald Reagan's conservative public philosophy were unlikely to win majority support in Congress.

Most analysts of American public opinion also concluded that the landslide re-election of President Reagan did not mean that a conservative public philosophy had displaced the idea of a broad federal role in society in the minds of Americans. By 1980 many Americans had lost confidence in traditional liberalism and the ability of the Democratic party to manage the economy, but most still viewed the domestic responsibilities of the federal government in broader terms than those advocated by Reagan.[20]

---

[18] *Id.* at 2257.

[19] *See Budget Deficit and Taxation: Radio Address to the Nation*, WEEKLY COMP. OF PRES. DOC. 1113–14 (Aug. 13, 1984).

[20] *See* Ladd, *The Reagan Phenomenon and Public Attitudes Toward Government*, in THE REAGAN PRESIDENCY AND THE GOVERNING OF AMERICA, *supra* note 15, at 221–49.

Opinion polls taken during Reagan's first term found increasing support for more federal spending on environmental, health, education, welfare, and urban programs, while support for increased defense spending steadily declined. Thus, as William Schneider noted in 1985, "Trend data on policy preferences indicate that public opinion since 1980 has moved away from Reagan's positions."[21] The one area where the conservative public philosophy did appear to strike a responsive chord was tax policy. Both before and after the 1984 election, opinion polls have generally shown large majorities opposed to raising taxes to reduce deficits.[22]

Therefore, as Ronald Reagan began his second term, divisions within Congress, and between Congress and the White House over budget priorities, reflected underlying ambiguities in the views of the American public. While the ambitious liberalism traditionally associated with the Democratic party had been discredited and little support was to be found for major tax increases, the sharply limited view of federal domestic responsibilities advocated by Reaganite conservatives had not changed the expectations that the federal government would remain engaged in a broad range of domestic activities. These conditions, together with the continued pressure of $200 billion deficits, made it inevitable that 1985 would be another difficult year for budget decisionmaking in Congress.

V. THE STRANGE CAREER OF GRAMM-RUDMAN-HOLLINGS

In keeping with statements made during the campaign, the budget submitted to Congress by the Reagan Administration in February 1985 fit the same basic pattern as those of earlier years. Defense spending increases of about six percent after inflation were requested, along with major cuts in domestic spending, including elimination of the Small Business Administration and subsidies for Amtrak. Other than a request for $3.6 billion in user fees, no new taxes were included in the President's budget for fiscal year 1986, nor were any changes in Social Security proposed.

---

[21] *See* Schneider, *The November 6 Vote for President: What Did It Mean?*, in THE AMERICAN ELECTIONS OF 1984 228 (A. Ranney ed. 1985); *see also Opinion Roundup*, 10 PUBLIC OPINION 30–33 (1987).

[22] *See Tax Americana*, 8 PUBLIC OPINION 19–29 (1985); *see also Opinion Roundup*, *supra* note 21, at 37.

The congressional response to the President's fiscal year 1986 budget also assumed a familiar pattern, with members of both parties rejecting the plan as unrealistic, and Senate Republican leaders taking the initiative in formulating an alternative. With the White House intervening at key points to head off Senate proposals to eliminate Social Security cost of living adjustments (COLAS) for a year and to increase revenues, the two chambers finally reached an agreement in August on a budget resolution that provided no increase in defense spending beyond inflation and required smaller domestic cuts than the President had requested. The congressional budget plan (some parts of which were not enacted until the following year) projected a budget deficit of $171.9 billion for fiscal year 1986.

What was new in deficit politics in 1985 was a proposal by Senators Phil Gramm (R-Tex.), Warren B. Rudman (R-N.H.), and Ernest F. Hollings (D-S.C.) to legislate prescribed annual deficit reduction levels, backed up by automatic spending reductions if deficits exceeded specified levels. When the Senate took up legislation in September of that year to approve additional federal borrowing authority needed to keep the government solvent, the Gramm-Rudman-Hollings proposal was initially introduced as an amendment. Despite widespread concern over the constitutionality and wisdom of the measure—even Senator Rudman described it as "a bad idea whose time has come"—frustration in the Senate over the impasse in budget politics resulted in passage of the amendment by a 75-24 margin on October 10, 1985. Pressures to support the balanced-budget legislation proved irresistable in the House as well, and by December 11, 1985 a revised version of the Gramm-Rudman-Hollings procedure had won approval by large margins in both chambers. The final measure, which revised existing budget procedures and required deficits to be reduced in annual increments to reach zero by FY 1991, was signed into law on December 12, 1985.

The enactment of Gramm-Rudman-Hollings did *not* represent the emergence of a new consensus on national budget priorities. Some in Congress thought the President would acquiesce to tax increases rather than risk defense cuts that would occur if deficit targets were not met; others believed public resistance to new taxes would force Congress to accept deeper cuts in domestic spending. Even the authors of the procedure disagreed over the effects it would have on deficit politics. When Hollings was

asked in a colloquy with Senator Bill Bradley (D-N.J.) if the threat of automatic spending cuts would result in a tax increase, he replied: "Right." "And it does so in the income tax system," Bradley continued. "You are right again, Senator," Hollings responded.[23] Senator Gramm, on the other hand, advocated adoption of the procedure as a means of securing further spending reductions. As Gramm explained during the floor debate:

> This bill does not dictate decisions, but it does force decisions. . . . I have a vision of America's future that is a vision of America growing where more and more people can provide things for themselves. And I believe that vision can be realized by controlling spending.[24]

Senate Budget Committee Chairman Pete Domenici's comments during floor debates on the amendment perhaps best reflected the combination of frustration and resignation that helped propel Gramm-Rudman-Hollings through Congress. "It is not the greatest way to manage a magnificent country," Domenici stated in arguing for the bill's passage. "But we have political gridlock at this point, and the existing processes of our government, executive and legislative, invite the continuation of gridlock."[25]

In signing the bill, President Reagan made it clear that his support for Gramm-Rudman-Hollings did not signal any change in the administration's priorities. "I want you all to know," he stated in a speech on December 18 1985, "that when I sat at my desk in the Oval Office and signed Gramm-Rudman-Hollings, I kept my veto pen ready in the top drawer. It's sitting there right now waiting for any tax increase that might come my way." He also emphasized the importance of maintaining growth in defense expenditures. As before, the focus remained on cuts in domestic spending as the solution to deficits. "We will meet the Gramm-Rudman-Hollings targets in the budgets that we submit to Congress, and we'll do it the right way—by cutting or eliminating wasteful and unnecessary programs."[26]

Contrary to the hopes of many of its original supporters, Gramm-Rudman-Hollings has not produced any grand compromise among the proponents of competing views of national

---

[23] 131 CONG. REC. S13051 (daily ed. Oct. 10, 1985).
[24] *Id.* at S17389–90.
[25] *Id.* at S17386.
[26] *Balanced Budget Legislation: Remarks to Congressional Supporters of the Legislation, December 18, 1985*, WEEKLY COMP. OF PRES. DOC. 1513 (Dec. 23, 1985).

priorities. A modest spending cut ($11.7 billion) required under the law for fiscal year 1986 was implemented, but the difficult test of meeting the $144 billion deficit target required for fiscal year 1987 was avoided when the Supreme Court on July 7, 1986 found the Gramm-Rudman-Hollings procedures for implementing automatic cuts unconstitutional. Prior to the Court's decision, Congress had rejected a Reagan Administration budget calling for substantial defense increases (eight percent after inflation), deep cuts and/or terminations in domestic programs, and sales of federal assets. Aside from some user fees proposed by the Administration, no tax increases were adopted by Congress in late 1986, and few of the domestic spending cuts proposed by the Administration were enacted. This left action by Congress to hold defense spending increases below the rate of inflation as the major deficit reduction effort for fiscal year 1987.

In 1987 the President and Congress did agree to a deficit reduction package including new revenues, but it took a revived Gramm-Rudman-Hollings procedure and a dramatic economic event, the "Black Monday" stock market crash of October 19, 1987 for this to occur. Congressional Democrats, in control of both chambers after the 1986 election, had initially enacted a fiscal year 1988 budget resolution calling for $19.3 billion in new taxes. A second budget measure enacted in September 1987 restructured the Gramm-Rudman-Hollings law to remedy the constitutional flaws identified by the Supreme Court, and also revised the congressional budget resolution for fiscal year 1988. In keeping with a new, less stringent set of deficit targets established by the Gramm-Rudman-Hollings "fix," the revised congressional budget called for $23 billion in deficit reduction for fiscal year 1988, including $12 billion in new revenues. These figures were superceded in turn by a two year plan developed through the informal "budget summit" convened by House and Senate leaders and the White House to calm financial markets following the October stock market dive.

In a press conference on October 22, 1987, the President announced his agreement to participate in bipartisan talks, explaining: "I'm putting everything on the table with the exception of Social Security."[27] As the negotiations proceeded, the primary goal became one of meeting or exceeding the Gramm-Rudman-

[27] *President Faced Questions on Budget, Persian Gulf Policies*, CONG. Q. WEEKLY REP., Oct. 24, 1987, at 2626.

Hollings targets for fiscal year 1988-1989, in order to reassure domestic and foreign investors that the national governing process was capable of making significant headway on chronic budget deficits. Even with the pressures generated by "Black Monday," and the presence of action-forcing deadlines created by the pending implementation of automatic cuts under Gramm-Rudman-Hollings, a major deficit reduction package proved difficult to craft. Republican participants in the summit sought to reach the targets through larger cuts in domestic discretionary and entitlement programs and smaller tax increases; Democrats tried to win agreement on a larger tax increase in order to shield domestic programs as much as possible. In the end, a $76 billion two-year package including $23 billion in new taxes was agreed to in November, 1987, but the package consisted of only incremental changes in tax and spending policies.

After three years under Gramm-Rudman-Hollings, no "grand compromise" has been reached on deficit reduction, and the debate among the nation's leaders over national priorities remains unresolved. What has been achieved since the law has been in place is a series of policy adjustments that should place deficits on a downward path—barring any unforeseen economic shocks. Given the difficulties of governing without a dominant public philosophy and with the configuration of power that has been present in Washington during these years, this is no small achievement. It is impossible to predict whether Congress will meet the more difficult Gramm-Rudman-Hollings deficit reduction targets in future years, or if not, when it will abandon the current timetable. Nonetheless, some more general observations about the future course of budget politics are offered by way of conclusion.

VI. Gramm-Rudman-Hollings and Governing in the Post-Liberal Era

The future role of Gramm-Rudman-Hollings in national policymaking will depend primarily on economic conditions that will determine the size of future budget deficits, on the partisan and ideological dynamics created by electoral outcomes in 1988, and on the progress of the continuing debate over the nation's public philosophy. Calls to resolve the deficit impasse through new reforms in the congressional budget process, or constitu-

tional amendments to require balanced budgets or a line item veto, assume that the problems of governing during the 1980's reflect procedural or organizational flaws in national institutions. To be sure, institutional arrangements are important in structuring budget politics. Decentralizing reforms in Congress during the 1970's left the institution in some respects ill-equipped to manage the priority-setting demanded by the massive deficits of the 1980's. The modest successes of Gramm-Rudman-Hollings also show that procedural arrangements may be important in encouraging action on deficit reduction when other conditions are favorable.

In organizational terms, there has already been significant movement toward greater centralization in congressional budgeting over the past decade, and the existing budget process with its reconcilitation mechanism has worked somewhat effectively to implement policy changes when majority coalitions in the House and Senate have reached agreement with the White House on a course of action. The primary obstacle to deficit reduction has been disagreement between the President and those who hold positions of power in Congress over the proper role of national government in society and in the world. Although further procedural reforms would undoubtedly have an impact on budget politics—including effects unanticipated by proponents—the current impasse in deficit politics is primarily a product of the breakdown of consensus on a public philosophy that began to occur in the 1970's. Resolution of the deficit impasse is at bottom a problem of political and intellectual leadership rather than one soluable through institutional reform of the type that has been proposed thus far.

The policy changes enacted early in the Reagan Administration transformed the national agenda in a direction that made revival of the ambitious liberalism of the 1960's and 1970's all but impossible. As Hugh Heclo argues:

> Reaganism . . . enjoys an advantage akin to political jujitsu, leveraging the weight of the governmental establishment against itself. The focus on revenue restraint (protection of previous tax cuts, indexation of tax rates) and the enduring deficit problem amount to a kind of institutionalized "no" in the political system.[28]

---

[28] *See* Heclo, *supra* note 1, at 49.

However, the same deficits have undermined support among the nation's political and intellectual leadership for the economic ideas on which Ronald Reagan sought to base a new conservative public philosophy, and public opinion on the role of government remains at some distance from the scaled-down ideal implied by unleavened Reaganism.

Seven years of deficit politics have demonstrated that neither the old liberalism nor a minimal government conservatism of the the Reagan variety can provide the framework for a governing coalition in American society today. Both the future role of Gramm-Rudman-Hollings, and the prospects for effective governing in the post-liberal era will depend to a considerable extent on the development of a public philosophy that better reconciles the hard realities of the present with the historic aspirations for liberty, equality and progress that animate American political life. As in the past, a new public philosophy must emerge over time through the attempts by political leaders to develop politically coherent and intellectually defensible responses to new circumstances. Gramm-Rudman-Hollings may contribute to this process by focusing attention on the fiscal situation, but continued progress in reducing budget deficits will be difficult without broader agreement on a new public philosophy.

GRAMM-RUDMAN-HOLLINGS AND THE BALANCED BUDGET ADMENDMENT: A PAGE OF HISTORY

DICK THORNBURGH*

One hundred and fifty years before today's growing and economy-threatening federal budget deficit became Congress's most pressing concern, budget and tax issues also dominated the public agenda. The dispute in the 1830's, however, was far different. The concern then was a huge budget surplus.[1]

While the problems are poles apart, the Congressional solutions have been similar. In the 1830's, Congress voted to impose automatic revenue reductions over a nine-year period. The law was successful and, as tariffs gradually declined, the budget was brought into balance.

The 1980's solution, The Balanced Budget and Emergency Deficit Control Act of 1985 ("Gramm-Rudman-Hollings"), calls for automatic elimination of the deficit over a six-year period.[2] Although only time will tell, I am not optimistic that Gramm-Rudman-Hollings' legislative solution to the budget deficit will be as successful as the 1830's legislation was in getting rid of the budget surplus. As we all know, cutting taxes and increasing spending are much easier than cutting spending and increasing revenues.

Reviewing federal government budget actions from the 1830's to the 1980's for this article proved once again that saying attributed to Yogi Berra—"It's deja vu all over again." Today's budget deficit affects the federal government's role by reducing resources so that state and local governments must resume some activities which had been taken over by the federal government. The surpluses of the 1830's also affected state-federal relations by their potential to strengthen the influence of the central gov-

---

* Director, Institute of Politics, Harvard's John F. Kennedy School of Government; Partner, Kirkpatrick & Lockhart, Pittsburgh, Pennsylvania; Co-Chairman, Citizens for a Balanced Budget Amendment; Governor of Pennsylvania, 1979–1987. B.Eng., Yale University, 1954; LL.B., University of Pittsburgh Law School, 1957.

[1] The surplus was caused by higher-than-expected revenues from the protective tariff initiated in 1815, which was the principal source of federal revenue in the 1830's.

[2] The Balanced Budget and Emergency Deficit Control Act of 1985, Pub. L. No. 99-177, 99 Stat. 1037, *amended by* The Balanced Budget and Emergency Deficit Control Reaffirmation Act of 1987, Pub. L. No. 100-119, 101 Stat. 754.

ernment.[3] The solution to the 1830's budget surpluses was Clay's Compromise Bill which

> would continue the tariff of 1832 with some major modifications. From January 1, 1834 . . . all duties over 20 percent would be reduced in biennial installments of one-tenth, with one half of the residue—six tenths of the whole—taken off January 1, 1842, the other half on July 1 of the same year. After that duties would be laid for the purpose of raising such revenues as may be necessary to an economic administration of the government.[4]

In other words, federal revenues were to be automatically reduced, on a phased-in basis over a nine-year period, to liquidate the surplus and fund only ongoing governmental operations. Does this sound vaguely familiar?

The Clay Compromise Bill was passed by both the House and the Senate and was signed into law by President Jackson. In the ensuing prosperity of the 1830's, the loss of tariff protection was hardly noticed and eventually the phased-in reduction of the surplus became a mere footnote to history, Clay having noted, "Now give us time; cease all fluctuations and agitations, for nine years to come, we can safely leave to posterity to provide for the rest."[5]

A century and a half later, rising deficits instead of surpluses were plaguing that posterity. But attempts to reduce the deficits and to balance the budget turned once again to the phased-in approach in 1985 with the Gramm-Rudman-Hollings Act. The Act went into effect for the 1986 fiscal year and outlined a five

---

[3] The swollen tariff collections and the budget surplus were prime sources of potential funding for Senator Henry Clay's "American system" of centrally planned and supported "internal improvements"—a massive infrastructure development program. See M. PETERSON, THE GREAT TRIUMVIRATE (1987) for a discussion of the contributions made to this debate by Senators Daniel Webster, Henry Clay, and John C. Calhoun, particularly Clay's proposals for solving the surplus problem.

As pointed out by Professor John Kenneth Galbraith, "The [embarrassing] surplus from the tariffs was deemed an urgent problem, the critical questions being whether to return funds to the states or to spend them, many thought unwisely or unconstitutionally, on internal improvements." J. GALBRAITH, ECONOMICS IN PERSPECTIVE 158 (1987).

According to Professor Peterson, President Andrew Jackson rejected both those alternatives and set out to reduce revenues to cover only the costs of operating government. Moreover, in his third annual message, he proposed measures to denationalize and simplify the role of the federal government.

To help achieve these goals, the Chairman of the House Ways and Means Committee, Guilian C. Verplank of New York, introduced a bill to withdraw the protective tariff over two years and eliminate the surplus precipitously. This proposal aroused a storm of sectional (north versus south) and sectoral (manufacturers versus planters) strife.

[4] M. PETERSON, *supra* note 3, at 226.

[5] *Id.* at 227.

year downward glide path which was amended to six years in The Balanced Budget and Emergency Deficit Control Reaffirmation Act of 1987 ("Reaffirmation Act"). The goal of these two Acts was a zero deficit by 1993.

Today's question is not whether the objectives of Gramm-Rudman-Hollings are good for the country. Rather, the key question today is whether Congress can meet even the amended Gramm-Rudman-Hollings' deferred targets and timetables, or whether budget waivers and a subsequently undiminished budget deficit is to become the order of the next five years—and thereafter.

It is my view that medicine stronger than Gramm-Rudman-Hollings will be needed this time, and that the prescription must be a Balanced Budget Amendment to the United States Constitution.

Despite the best efforts of its authors to address the very serious problem of the federal deficit, Gramm-Rudman-Hollings probably cannot accomplish its balanced budget goal. The problem is that Gramm-Rudman-Hollings cannot truly operate as planned—to force Congress to make the hard decisions necessary to reduce the deficit. Gramm-Rudman-Hollings cannot force Congress because it is a law and, like any other law, can be amended or waived by Congress.

For instance, current law requires Congressional passage of thirteen separate appropriations bills each year by October 1, each covering different aspects of federal spending. Since the beginning of this decade, Congress only has met the October 1 deadline it set for itself on ten of the ninety-one required bills.[6]

During the past two years, the new budget deadlines and deficit reduction targets of Gramm-Rudman-Hollings have been in effect. However, during the same two years, not one of those thirteen appropriations bills has been approved, and federal appropriations have been amassed into a single overall spending resolution.[7] Last fall Congress had to throw together four temporary emergency continuing resolutions just to keep the government from running out of money and shutting down.[8] Shortly before Christmas and three months late, Congress finally approved a 1,057 page budget continuing resolution for the re-

[6] State of the Union Address by President Ronald Reagan 3 (Jan. 25, 1988) [hereinafter Union Address] (on file at the Harvard Journal on Legislation).

[7] *Id.*

[8] *Id.*

mainder of the year.[9] Clearly, Gramm-Rudman-Hollings has not changed the Congressional pattern of missed budget deadlines.

Moreover, Congress has also gotten around Gramm-Rudman-Hollings' budget cutting requirements. The House and Senate regularly waive provisions of the 1974 Budget Act and the Gramm-Rudman-Hollings Act in approving new spending bills for which there are no new revenues.[10] The result of these waivers is the perpetuation of a massive budget deficit. According to a study by the House Operations Committee for Rep. Robert Walker (R-Pa.), the 99th Congress approved 106 budget waivers.[11] Walker concluded, "a clear record [exists] of Congress refusing to discipline itself through the Budget Act. Those who control the Congress are obviously the source of the problem."[12]

Because Congress can amend or waive any law, to balance the federal budget we absolutely need an amendment to our federal Constitution. It seems that a balanced budget amendment is the only effective way to impose long-overdue fiscal discipline on Washington's "credit card" mentality.

The time has come to provide Congress and the President with the same structural tools and constraints that states have used effectively to balance their budgets, year-in and year-out. During the 1982–1983 recession, for example, forty-three states cut expenditures and forty-four raised taxes in order to keep budgets in balance.[13] It seems doubtful that these actions would have occurred without state constitutional requirements mandating balanced budgets, and without the executive and legislative discipline those provisions impose.

In fact, it has been documented. A 1987 study by the Advisory Committee on Intergovernmental Relations of state fiscal discipline mechanisms concluded that "more stringent balanced budget requirements are significantly associated with a reduc-

---

[9] *Id.*

[10] The Senate Republican Policy Committee found that the 1974 Budget Act or the Gramm-Rudman-Hollings Act were waived at least nine times last year alone to increase federal spending. SENATE REPUBLICAN POLICY COMMITTEE, SENATE RECORD VOTE ANALYSIS, 100th Cong., 1st Sess. S-1374, S-1451, S-1671, S-7186, S-7349, S-7409, S-8043, S-11026 (1987) (on file at the Harvard Journal on Legislation).

[11] May 26, 1987 News Release of Congressman Bob Walker, *Study Shows Budget Committee Makes Budget Rules Meaningless* 2 (on file at the Harvard Journal on Legislation).

[12] *Id.* at 2.

[13] NATIONAL GOVERNOR'S ASSOCIATION, POLICY POSITION REPORT 1986–1987 13 (on file at the Harvard Journal on Legislation).

tion in budget deficits and with an increase in the size of the surpluses."[14]

Meanwhile, the federal budget process, lacking any such discipline, has been out of balance in twenty-seven of the last twenty-eight years, and the total national debt has more than doubled in the 1980's alone.[15] As President Reagan stated in renewing his call for a balanced budget amendment in this year's State of the Union address, "the budget process has broken down, it needs a drastic overhaul."[16]

In this he is not alone. The National Governors Association, on a bi-partisan basis, has expressed support for the balanced budget amendment,[17] the presidential line-item veto,[18] and a separate capital budget[19] (differentiating investments from current outlays). All of these budget balancing tools are already available to most state governors and legislatures. Moreover, recent polls indicate that seventy-five percent of the American public favors a balanced budget amendment,[20] while the legislatures of thirty-two states have already called for a federal Consititutional convention to consider a balanced budget amendment.[21]

The balanced budget amendment has solid support, but doubters remain. To their most-frequently voiced objections I offer the following responses.

First, it is argued that the amendment would "clutter up" our basic document in a way contrary to the intentions of the Founding Fathers. This argument is clearly wrong. The framers of the Constitution contemplated that amendments would be necessary to keep it abreast of the times. The document has already been amended on twenty-six occasions. Moreover, one certainly can speculate that the notion of a federal government consistently spending more than it took in was so alien to the thinking of 1787 that a balanced-budget provision might well have been deemed superfluous. Indeed, one of the major preoccupations

[14] Advisory Commission of Intergovernmental Relations, Report: Fiscal Discipline in the Federal System—National Reform and the Experience of the States, Report #A-107 39 (July 1987) (on file at the Harvard Journal on Legislation).

[15] NATIONAL TAXPAYER'S UNION, REPORT: NOV. 1987 table: "Federal Budget Surpluses vs. Deficits" (research by Sid Tavler) (on file at the Harvard Journal on Legislation).

[16] State of the Union Address, *supra* note 6, at 4.

[17] NATIONAL GOVERNOR'S ASSOCIATION, *supra* note 13, at 14.

[18] *Id.* at 15.

[19] *Id.* at 33.

[20] New York Times, Dec. 1, 1987, at 24 (Poll Table).

[21] State of the Union Address, *supra* note 6, at 4.

at the time of the Constitutional Convention was dealing with the problem of how to liquidate the post-Revolutionary War debts of the states in an expeditious manner. The Treasury did not begin systematically to incur annual deficits until the mid-1930's, nearly a century and a half after the adoption of the Constitution.

Second, critics argue that the adoption of the amendment would not solve the deficit problem overnight. Serious proponents of the amendment have never claimed that it would. Obviously a period, such as the five years suggested by the nation's governors and orginally embodied in the Gramm-Rudman-Hollings Act, would be required for the full phase-in of a reduction to zero deficit. During this interim period, however, budget-makers would be constitutionally disciplined to meet declining deficit targets in order to reach a final balanced budget by the established date.

Third, it is argued that such an amendment would require vast cuts in social services, or the military, or other categories of expenditure. Not necessarily. These programs would have to be *paid for* on a current basis. Certainly, difficult choices would have to be made about priorities and levels of program funding in all areas, but the very purpose of the amendment is to discipline the executive and legislative branches actually to make these choices and not to propose or perpetuate vast spending programs without providing revenues to fund them. The amendment would, in effect, make Congress and the President more accountable for the spending decisions that they make while in office.

Fourth, critics say that a balanced budget requirement would prevent or hinder our capacity to respond to national defense or economic emergencies. This concern is easy to counter. Of course, any sensible requirement would feature a "safety valve" to exempt deficits incurred in responding to such emergencies—perhaps the requirement of a two-thirds or three-fifths "super majority" in both houses of Congress. Congress could also institute "rainy-day" funds to set aside current revenues during good times to be used for counter-cyclical purposes during economic retrenchment, as is currently done in over thirty states, or perhaps even for debt reduction as time went on.

Fifth, it is said that a balanced budget amendment would be "more loophole than law" and easily could be circumvented. The experience of the nation's governors suggests that this ar-

gument is also wrong. Balanced budget requirements are now in effect in all but one of the fifty states and have served them well.[22] Moreover, giving the President the line-item veto, which is available to forty-three governors, would assure that any specific congressional overruns (or loophole endruns) could be dealt with by the President. The public's outcry, the elective process, and the courts would also provide backup restraint on any tendency to ignore a constitutional directive.

In the final analysis, most of the excuses raised for not enacting a Constitutional mandate to balance the budget seem to rest on a stated or implied preference for solving our deficit dilemma through the "political process"—through responsible action by the President and Congress.

As noted above, this has been tried and found wanting, again and again.

I believe this country is ready for a simple, clear, and supreme directive that its elected officials fulfill their fiscal responsibilities, and I believe that a Constitutional amendment is the only instrument that will fulfill this need effectively. Years of experience at the state level argue persuasively in favor of such a step, and years of debate have produced no persuasive arguments against it.

A distinguished American and advisor to eight Presidents, Milton S. Eisenhower, in one of his last observations prior to his death said: "Without an amendment requiring the President and Congress to maintain a balanced budget, with exceptions to meet certain emergencies, our nation faces disaster."[23]

We can yet avoid such a disaster if all interested parties move quickly to do so. That's why former Colorado Governor Dick Lamm, a Democrat, and I have agreed to head up a bi-partisan grass roots effort, Citizens for a Balanced Budget Amendment, to work for the enactment of a balanced budget amendment.

Perhaps Thomas Jefferson said it best: "to preserve our independence, we must not let our rulers load us down with perpetual debt."[24] That is the aim of the Balanced Budget Amendment.

---

[22] Vermont is the exception.

[23] Wall St. J., Jan. 17, 1985, at 26, col. 3.

[24] Brief for The National Taxpayer's Union at 136 of affidavit 2, attachment 1 (citing Letter from Thomas Jefferson to Samuel Kerchival of 1816), Bowsher v. Synar, 106 S. Ct. 3181 (1986).

RECENT PUBLICATIONS

POLITICAL ETHICS AND PUBLIC OFFICE. By *Dennis F. Thompson.* Cambridge, Mass.: Harvard University Press, 1987. Pp. viii, 263, notes, credits, index, n.p.

Should Oliver North be criminally liable for the diversion of funds to the contras? How responsible is a president for the actions of subordinates? Is the private life of Gary Hart or the academic record of Joe Biden legitimately a matter of public concern? Did Robert Bork act unethically when he dismissed Archibald Cox as Special Prosecutor of the Watergate investigation? Many of the important political issues of the past year raise questions of the moral conduct of public officials. Dennis F. Thompson, a professor at the John F. Kennedy School of Government of Harvard University has written a new book, *Political Ethics and Public Office*, which attempts to develop a conception of political ethics which addresses the contemporary ethical dilemmas of public officials.

Thompson begins by disputing the belief that ethical considerations should not play a role in the policy decisions of public officials. He sees a mutual dependency between politics and ethics in a democracy. In his view, political ethics provide support for democratic politics by supplying criteria by which public officials can be judged and be held accountable (p. 3). Democratic politics supports political ethics, since collective deliberation is an important way in which our society resolves disputes concerning fundamental values (p. 3).

Thompson develops a conception of political ethics and applies it to ethical conflicts that democratic societies face. For example, when can paternalistic interventions regulating personal safety be justifiably made by public officials? He cautions that his approach does not yield a handbook for public officials, but instead a mode of analysis based on factors "that citizens as well as officials consider as they deliberate about decisions and policies" (p. 7). The unifying themes in his analysis are that individual responsibility must be maintained even within a complex bureaucratic organization and that public participation in the decision-making process is important. The latter theme is addressed by Thompson in his discussion of the first ethical

dilemma in *Political Ethics and Public Office*: the problem of democratic dirty hands (ch. 1).

Thompson's term, "democratic dirty hands," describes the ethical conflict that occurs when a public official must violate a shared moral principle in order to effectuate some public purpose (p. 11). Thompson feels that the key to dealing with the dirty hands paradox is to focus on accountability for such decisions. This can be done by retrospective review of past dirty-handed decisions, generalized discussion of future policies which require dirty-handed decisions, and oversight of such decisions. Congressional intelligence committees are an example of the oversight role, while the recent Iran-Contra hearings are a good example of the retrospective review approach. This latter approach, Thompson points out, is limited since the damage has already been done (p. 24).

Thompson attempts to reconcile the practical requirements for secrecy in government with his insistence that dirty-handed decisions pose a special ethical dilemma to democracies by undermining their ability to deliberate collectively about key issues. He uses nuclear deterrence as an example of this conflict between the need to have public participation on an important issue and the need to have both secrecy and uncertainty in order for it to be effective (pp. 33–38).

While Thompson is explicit in identifying the moral costs of deterrence to a democratic society, his attempt to show the other side of the balance is less than convincing. He suggests that the dirty-handed undemocratic nature of deterrence may "shift the balance of risks in ways that could encourage a democracy to take greater chances in negotiations on arms control and perhaps even to take some unilateral steps in the reduction of arms" (p. 38). This argument has a mechanical, almost syllogistic structure, and is totally devoid of the strategic consequences of abandoning deterrence and the moral costs of those consequences. These moral costs may not be as direct as the costs which Thompson outlines, but can be seen, for example, in Greece's ability to elect a government democratically after the United States deterred the Soviet Union from making it part of the Eastern Bloc. Also, assuming a massive reduction in nuclear weapons, the United States would presumably shift its reliance from nuclear to conventional weapons. Yet, in order to avoid war, the United States would still rely on deterrence, and

Thompson's criticisms of deterrence would seem to apply whether the deterrence is based on nuclear warheads or tanks.

Political ethics, according to Thompson, requires that we "stop the making of decisions that cannot be justified in public," and strengthen the imperfect but useful checks which provide for accountability (p. 39). He feels that, although we cannot eliminate the problems of dirty-handed decision making, we can alleviate them somewhat (p. 39).

While the chapter on dirty hands allows Thompson to focus on his theme of public accountability in making decisions, Thompson switches his focus to individual responsibility when dealing with the dilemma of "many hands" (ch. 2). "Many hands" refers to the difficulty faced in locating responsible parties in large complex organizations. Such an effort is considered important since citizens need to identify responsible persons in order to effectuate government accountability (p. 40). Thompson illustrates his discussion of accountability by describing the familiar ritual of the president, or other top official, who takes responsibility for a failed policy. For example, President Kennedy publicly accepted full responsiblity for the Bay of Pigs failures, while more recently, President Reagan took full responsibility for the failure to protect the United States Marines in Lebanon. Instead of generating a sincere effort to understand what caused the failure, the presidential ploy is little more than a ritualistic incantation calculated to cut short any public inquiry (p. 44).

Thompson's approach seeks to tie moral responsibility to an official if the elements of causation and volition are present. He deals with many of the issues which naturally arise when an individual official's conduct is being scrutinized. A sample of such issues includes: an advisor's claim that she lacked the authority to affect an outcome, the unintentional consequences of a well-intentioned policy, an official's complaint that his predecessor's policies created the problem, and the simple pleas of ignorance by an official. When addressing all of these issues, Thompson rearticulates his theme of personal responsibility.

Thompson uses a varied array of examples to illustrate these issues. For example, he notes how those advisors of President Johnson who were critical of his Vietnam policy were viewed as devil's advocates who reassured the president that his policies were receiving thorough consideration (p. 59). Thompson

finds fault with such officials who persisted in offering the same kind of advice, knowing that it was not being used as intended; however, he would excuse those officials who were unaware of the actual results or could do nothing to prevent such results.

Thompson's message has a reassuring aspect to it in that it tells one to reject feelings of futility or the "you can't fight city hall mentality." While his conception of political ethics has an optimistic tone, it also has a disturbing aspect to it. This can be seen in his discussion of the former New York City Mayor, Abe Beame (pp. 61–65). The Beame example is used to illustrate how focusing on individual responsibility can transcend offices and bureaucratic structures. Mayor Beame blamed the policies of his predecessors for New York's fiscal crisis in 1975. As Thompson points out, Beame in fact had major roles in previous administrations, and therefore such excuses rang hollow. The Beame example, however, illustrates Thompson's point just a little too well. By placing so much emphasis on individual responsibility, Thompson leads us down the same road that he himself criticizes in the presidential, "I take full responsibility" ritual; that is, by having someone to focus the blame on, one eliminates the difficult task of finding out what went wrong.

Additionally, Thompson claims that his conception of political ethics is meant to apply equally well to the marginal choices that public officials face in the day-to-day performance of their jobs (p. 7). However, lower-level officials who make these marginal choices are often so wary of being scapegoats, that bureaucracies, which are afflicted with this cover-myself-first mentality, are unable to get anything done in crises.

Thompson's focus on the individual is intelligent and useful, since it allows us to scrutinize the types of excuses made by officials to determine how valid they really are. But in the less pristine environment of contemporary politics, it can also lead to an increase in fingerpointing and scapegoating, which politicians and bureaucrats already resort to all too quickly.

Thompson further bolsters his ethical prescriptions by advocating the use of criminal laws to punish public officials who violate ethical duties. He promotes the Lockean idea that officials have a fiduciary responsibility to those they serve (p. 83). While few would argue that there is no place within the criminal law for public officials, Thompson goes one step further and argues that public officials should be subject to criminal punishment on the basis of negligent behavior. Once again such an

approach will invite scapegoating. Thompson reasons that while tough penalties for misconduct may discourage some people from accepting public office, it may also encourage more worthy people by making public service a more honorable calling (p. 83). Such reasoning is strikingly disingenuous. Consider the irony of a political system in which public officials subject to criminal liability for negligent behavior are being scrutinized by a press who are legally protected from liability in the absence of "reckless disregard for the truth." It is difficult not to see that such harsh scrutiny will result in a "chilling effect" on individuals seeking public office.

Political Ethics and Public Office is an ambitious book with laudable goals. Thompson's ethical discussions operate from "middle level principles," so while he occasionally gets abstruse, we are spared long philosophical digressions. Thompson's skillful use of examples underscores his desire to show that his conception of political ethics is meant to be a workable one—one which functions in non-ideal circumstances. Thompson's strong argument for individual responsibility can help clarify the ethical obligations of public officials and help citizens ask intelligent and probing questions in order to ensure accountability. Focusing on the individual official, however, may lead to an increase in partisanship, personal attacks, and superficial inquiries which focus only on whom to blame, and not on finding out what was the cause of the mistake in policy.

—*Stephen Bier*

MERGERS, SELL-OFFS, AND ECONOMIC EFFICIENCY. By *David J. Ravenscraft and F.M. Scherer.* Washington, D.C.: The Brookings Institution, 1987. Pp. xiii, 283, index, n.p.

Corporate mergers, acquisitions, and divestitures have become so common in recent years that the 1980's may be remembered in twentieth century business history as the "decade of restructuring."[1] In 1987 alone, the Wall Street Journal reports, the "top fifteen" advisers for mergers and acquisitions completed a total of 1,248 deals worth $342 billion.[2] Moreover, 1988

[1] Kiechel, *Corporate Strategy for the 1990's*, FORTUNE, Feb. 29, 1988, at 34.
[2] Wall St. J., Feb. 3, 1988, at 1, col. 6.

promises to be little different; merger and acquisition specialists see little or no indication of takeover activity waning in the near future.[3] What does all this activity mean for the businesses involved and the American economy as a whole?

The standard textbook answer is that mergers and acquisitions promote efficiency in some circumstances and impair it in others.[4] Mergers are generally characterized as "horizontal," "vertical," or "conglomerate."[5] According to economists, horizontal mergers can enhance efficiency by creating greater economies of scale. Similarly, vertical mergers can lead to greater efficiency by promoting vertical integration, giving a corporation as much control as possible over the entire production process.[6] A conglomerate merger, however, is unlikely to increase efficiency, unless it creates some economy of scale or provides complementary resources to a parent.[7]

In *Mergers, Sell-Offs, and Economic Efficiency*, David J. Ravenscraft and F.M. Scherer join a number of other analysts who take issue with this conventional wisdom.[8] The book describes in great detail the two economists' economic study of corporate takeovers and the spin-offs which often follow. Using accounting data of corporations which went through restructuring between 1950 and 1981, the authors analyze the wave of mergers which took place during the late 1960's and early 1970's. From this study, they argue that mergers and acquisitions of all types generally create economic inefficiency (pp. 211–12).

Scherer and Ravenscraft's study of mergers and acquisitions differs from many past analyses in its methodology. Prior studies of the economic effects of mergers have often relied on the

---

[3] Wall St. J., Mar. 3, 1988, at 3, col. 1.

[4] R. BREALEY & S. MYERS, PRINCIPLES OF CORPORATE FINANCE 703 (1984). As a recent government report indicates, there remain significant disagreements among economists and government officials about the effects of mergers and acquisitions. *See Corporate Takeovers: Public Policy Implications for the Economy and Corporate Governance, Report from the Chairman of the Subcomm. on Telecommunications, Consumer Protection, and Finance of the House Comm. on Energy and Commerce*, 99th Cong., 2d Sess. 8 (1986) [hereinafter *Public Policy*].

[5] In a horizontal merger, the parent company acquires a firm in the same line of business; in a vertical merger, the parent buys out a firm with which the parent could have had a buyer-seller relationship for raw materials or finished goods; and in a conglomerate merger, the parent takes control of a company which operates in an unrelated line of business. R. BREALY & S. MYERS, *supra* note 4, at 703–04.

[6] *Id.* at 704.

[7] *Id.* at 706–07.

[8] These analysts include law professors as well as economists. *See Impact of Corporate Takeovers: Hearings Before the Subcomm. on Securities of the Senate Comm. on Banking, Housing, and Urban Affairs*, 99th Cong., 1st Sess. 108 (1985).

prices of stock, the capital asset pricing model, and the efficient market theory of corporate finance (p. 5). Scherer and Ravenscraft, however, question the validity of the efficient market theory; they argue that stock prices may not perfectly reflect company asset values (p. 8). Scherer and Ravenscraft instead use the accounting figures of American manufacturing companies, most of which were gathered from a special Federal Trade Commission (FTC) data base. This data base includes many smaller companies and a broader range of firms than have been examined in other studies. The authors therefore attribute some of the differences between their results and those of others to this better sample (p. 66). Scherer and Ravenscraft also draw upon many individual case studies to add detail to their analysis and to test their hypotheses empirically. They admit that their use of company financial statements in this type of study is problematic, but argue that this information is much richer than stock price data, and it can therefore better illuminate the economic relationships studied (p. 12).

Scherer and Ravenscraft were well aware of possible problems with their accounting data and took many steps to ensure the validity of their analysis. Because the financial information of acquired manufacturing companies is not kept distinct from that of their parents, the two economists had to examine company performance at the level of each corporate operating unit, or "line of business," in the FTC data base (p. 13). To ensure the uniformity of the data, different methods of accounting for depreciation and inventory, allocation of fixed costs, and macroeconomic trends also had to be statistically controlled in their tests (pp. 16, 105). Finally, Scherer and Ravenscraft also had to cope with two different methods of accounting for acquired assets: the pooling-of-interest and purchase methods (pp. 13–14). Given these complicating factors, the authors do not claim that their data is free of error or distortions (p. 15); however, since they were aware of the problems inherent in their study, they probably reduced or minimized some biases.

In their second chapter, Scherer and Ravenscraft give the reader an overview of mergers and acquisitions. During the twentieth century, America has experienced four major merger "waves": one at the turn of the century, a second in the late 1920's, a third with its peak in 1968, and the present one, which is by far the largest (p. 21). Scherer and Ravenscraft next differentiate horizontal, vertical, and conglomerate mergers. In so

doing, they classified as "horizontal" mergers many of the corporate combinations which the FTC had previously labeled "conglomerate" mergers (pp. 23–24). Although this categorization is different from the norm, the two analysts believe it is more accurate (p. 24). Despite their analysis of all three types of corporate combinations, the authors focus on conglomerate mergers and the trend toward diversification in the 1960's and 1970's (pp. 212–13). Because they could only collect enough data for analysis from the period 1950–1977, however, Scherer and Ravenscraft generalize from their studies of the merger wave of the late 1960's to today's fourth wave (pp. 18, 219).

Scherer and Ravenscraft demonstrate that the firms acquired during the 1960's and 1970's were not only profitable, but also were substantially more profitable than the average firm in their industry (p. 66). Smaller businesses, which were picked for purchase, were especially profitable (p. 63), and companies with high rates of growth were also particularly favored (p. 53). Finally, the authors discovered that although those firms which were the target of tender offers were generally underperformers in their industry, they were nevertheless on average more profitable than the universe of all manufacturing industries (p. 70).

Most of these acquired companies, however, performed much worse after purchase. Average growth by acquisition was less profitable than growth by internal development (p. 95). Acquired firms' operating incomes and profitability fell significantly and cumulatively in the years following combination (pp. 93, 111, 168). Moreover, firms which were added to conglomerate corporations became the least profitable of all acquired companies (p. 99). Firms whose stock had been purchased through a tender offer also did poorly following their mergers; their profitability and market share generally stayed below those of other companies in their industry over the long run (pp. 102–03). Overall, average growth by acquisition was less profitable than growth by internal development (p. 95). There were, of course, some cases of acquired firms that performed well after merger and others in which the parent company was able to turn around a poor performer (p. 141). Nevertheless, most acquiring corporations seemed to have had serious indigestion after swallowing their prey (pp. 192–93).

Scherer and Ravenscraft provide several explanations why acquired companies tended to perform worse after merger. Some of the acquired firms had problems which the parent had

not been aware of, such as an inexperienced work force, a lack of technological expertise, or a group of entrenched managers who resisted change (p. 134). Many of these firms also experienced leadership and organizational problems (p. 135). In these latter cases, the parent corporation retained the upper management of the acquired company and merely installed a new chief executive officer (CEO). This new CEO, however, was usually an outsider who had great difficulty taking charge of the business and mastering the details of its operation. These new CEOs were often unable to maintain morale, since old managers tended to see new chief executives as authoritarian and unsympathetic (p. 136). Moreover, the managers of some acquired firms were put under pressure to achieve high short run profits and/or cash flow. This emphasis led to some foolish management decisions and aggravated other difficulties (pp. 139–40).

As a result of falling profitability and unfulfilled expectations, corporate divestures in the late 1960's and early 1970's skyrocketed (p. 161). Most of these sell-offs were the result of earlier combinations (p. 160). In fact, Scherer and Ravenscraft estimate that thirty-three percent of all acquisitions made in the merger wave of the late 1960's were eventually divested (p. 166).[9] In selecting the lines of business to cast off, parent companies were more likely to hold on to lines which had absorbed large amounts of research and development (p. 179) and firms which were added through vertical acquisitions (p. 183). The units which were sold, however, tended to be no less profitable before their acquisition than units which were retained (p. 65). Moreover, spin-off companies were often able to improve their performance once freed from conglomerate control (p. 157).

On the macroeconomic level, Scherer and Ravenscraft believe that the last merger wave seriously hurt the American economy. They argue that mergers put factors of production to inefficient use, creating estimated annual efficiency losses of $1.78 billion during the period of 1975–1977 (p. 202). In addition, Scherer and Ravenscraft contend that the effects of mergers created

---

[9] For an interesting comparison and deeper explanation, see Porter, *From Competitive Advantage to Corporate Strategy*, HARVARD BUS. REV. 43–59 (May–June 1987). Porter examined thirty-three large, well-known American companies over the period 1950–1986 and determined that most of these corporations divested many more acquisitions than they retained.

about a tenth of the decline in productivity suffered by the American manufacturing sector between 1968 and 1976 (p. 203).

In their summation, Scherer and Ravenscraft emphasize their findings from the last merger wave and warn that present acquisitions are likely to promote similar inefficiency (pp. 218–19). They admit that there have been some successful mergers in the past (p. 194),[10] but they firmly believe that these are isolated exceptions to the rule (pp. 211–12). They do, however, see some encouraging changes in current corporate behavior. Many of today's managers, they believe, are aware of the problems and risks of conglomerate acquisition, and, as a result, conglomerate merger activity (as opposed to horizontal and vertical merger activity) has declined (pp. 217–18).

Scherer and Ravenscraft nevertheless argue that still more has to be done. They call for better education of corporate management[11] and for several legislative changes (pp. 217, 221). In particular, they argue for a change in antitrust policy which would establish a tough presumption against large horizontal mergers; the parties would have to demonstrate that a combination would create "substantial efficiencies" which would be otherwise unattainable. Moreover, they would modify the Williams Act of 1968 and Securities and Exchange Commission (SEC) rules regulating tender offers to eliminate biases favoring takeovers. Specifically, the two analysts believe that current law compels shareholders to tender their stock for fear of being "frozen out," and should be altered to protect investors' interests while allowing them to indicate their real preferences for or against proposed mergers (p. 227). Finally, Scherer and Ravenscraft would require publicly traded companies to submit data annually to the SEC so that its economists can better evaluate the effects of mergers and acquisitions (pp. 227–28). These changes, the two authors suggest, would discourage mergers and acquisitions generally and force managers to evaluate more thoroughly their decisions to make corporate acquisitions (pp. 225–26).

Mergers, Sell-Offs, and Economic Efficiency is not likely to get much attention in legislative circles, because it is difficult to read. Unfortunately, the authors wrote their book using lan-

[10] In three of their case studies, Scherer and Ravenscraft found that horizontal mergers did produce greater economic efficiency (p. 150).

[11] They see managers graduating from business schools with "naive views of merger-making as a quick, easy road to wealth creation" (p. 217).

guage and terms best understood by fellow economists and business school professors. The chapters also go into great detail about the different statistical tests performed by Scherer and Ravenscraft, and while an economist interested in replicating their studies may want to be taken step-by-step through their analysis, the authors' conclusions are often buried in their discussion of the statistics.

Futhermore, Scherer's and Ravenscraft's claim that today's mergers and acquisitions are generally inefficient is not convincing for several reasons. First, the two economists' findings that mergers of all types are inefficient may be at least partly the result of their categorization of corporate combinations. Most analysts would agree that many of the conglomerate mergers of the 1960's were inefficient;[12] since the authors' system of classification has the effect of shifting what were considered conglomerate mergers into the horizontal merger category (pp. 23–24), the authors' statistical evidence that horizontal mergers are usually inefficient may be distorted by this change. Thus, if Scherer and Ravenscraft had not classified their data as they did, their studies might only prove what we already know: that most conglomerate mergers are inefficient.[13]

Second, even if the authors' data are correct, the factors which they cite as causing inefficiency in the mergers and acquisitions of the late 1960's have been reduced today. Scherer and Ravenscraft point to managerial problems and tight financial control as being responsible for the inefficiencies which followed the mergers of the 1960's. The authors, however, admit that managerial problems in acquired companies will occur less often with today's horizontal mergers because most newly appointed executives will be familiar with the acquired firm's line of business (p. 219).[14] Similarly, corporate executives in the 1980's are much more likely to give operating managers greater independence and not keep them under tight financial control.[15] Since

---

[12] *See, e.g., Public Policy, supra* note 4, at 41; R. BREALY & S. MYERS, *supra* note 4, at 706–09; Kietchel, *supra* note 1, at 34.

[13] *See*, R. BREALEY & S. MYERS, *supra* note 4, at 703, 706–09.

[14] For further evidence that mergers in the 1980's are generally horizontal, and that these acquisitions are part of rational corporate strategies, see N.Y. Times, Mar. 19, 1988, at 31, col 3.

[15] Kietchel, *supra* note 1, at 37. This may not be true, however, for those firms which are acquired in transactions that encumber them with high levels of debt; such debt puts enormous pressure on managers to cut costs and maintain tight financial control. *See, e.g.*, Pare, *The New J.P. Morgans*, FORTUNE, Feb. 29, 1988, at 50; Wall St. J., Mar. 15, 1988, at 29, col 2.

the command and control problems of the last merger wave occur less often in the 1980's, Scherer and Ravenscraft do not provide a compelling theory to explain why horizontal mergers today are not efficient.

Finally, on a more general level, the results of Scherer and Ravenscraft's study of the mergers of the 1960's and 1970's may not be applicable to the 1980's because takeover practices today are considerably different. A recent government report on corporate takeovers states:

> The difficulties in settling the debate [over the benefits and harms of mergers] are compounded by the fact that takeover transactions today are much different from the ones on which many existing economic studies are based. We have moved from the conglomerate merger wave of more than a decade ago, to a wave of divestitures, spinoffs, financial restructurings, and leveraged buyouts.[16]

The two analysts do not make any real attempt to grapple with issues such as junk bonds, golden parachutes, major asset sales, leveraged buyouts, and other aspects of corporate combination in the 1980's which were not present in the 1960's. Without a discussion of these issues, this study of corporate takeovers loses its applicability to the mergers and acquisitions of today.

While *Mergers, Sell-Offs, and Economic Efficiency* gives us a thorough view of the last merger wave, its conclusions about the mergers of the 1980's lack certainty and credibility. Because of their possible miscategorization of merger types and the changes between past and present merger waves, the reader doubts the applicability of their discussion of the conglomerate mergers of the 1960's to current mergers. Thus, while their study reinforces the lessons of the past, it does not provide convincing evidence on the merits of corporate restructuring today.

—*Robert S. De Leon*

---

[16] *Public Policy*, *supra* note 4, at 8.

THE PRESIDENTIAL CAMPAIGN, THIRD EDITION. By *Stephen Hess*. Washington, D.C.: The Brookings Institution, 1988. Pp. viii, 118, bibliography, index. $22.95 cloth, $8.95 paper.

Once again we are in the midst of an election year, and once again politicians, journalists, political scientists, and many voters are bemoaning the way we pick our presidents. Pundits from all walks of life complain that our current system of choosing our leader discourages the truly capable from running and leaves us with a slate of telegenic, overly-ambitious political hacks. Our choice, these observers claim, is driven by the ability of any given candidate to manipulate television successfully and to pander sufficiently to special interests on both the right and the left to raise funds to buy the necessary television time. This vision of our electoral politics presents a fatalistic view of our hope for a strong future leadership that will be able to address the problems facing this country.

Stephen Hess is one of the most experienced observers of our presidential selection process, and he rejects the conventional thesis that television and money have permanently corrupted our electoral process. Hess, who has served four presidents and has been a student of presidential politics for more than thirty years, calls himself an "optimistic fatalist" (p. 2). While he does not profess to be pleased with the way in which we select our presidents, he is unconvinced that any changes in the method of selection will have any real impact on the results. This is a new position for him, for in the first edition of this essay, published in 1974 in the wake of Watergate and post-Vietnam political distrust, he suggested massive electoral reform, some of which was codified in the election reform acts passed before the 1976 election. Looking back on those reforms, Hess now argues that procedural changes will not improve the results (p. 118) and have not improved the public perception of the procedure. "I am skeptical of quick fixes," he writes, "of proposals that put a patch on party rules or juggle delegate selection timetables or outlaw negative ads" (p. 2).

Pulling examples from all of American history, but particularly from the period since World War II, Hess demonstrates that technological, financial, and moral changes that have had surface impacts on presidential selection appear to have had little or no effect on the results. Hess reaches this conclusion by

examining, in succession, those who run for president, how they came to run for president, the process of selecting nominees, and finally, the relationship between candidates and the media and candidates and the political parties.

After this review, Hess concludes that our current system of choosing the president, while not perfect, and certainly not pretty, works as well as any system we have ever used or as well as any we could imagine. His one regret is the decline in the strength of the parties which, he believes, could work to motivate voters and provide information to both candidates and voters in order to reduce the impersonal nature of presidential campaigns.

While Hess's conclusions are enlightening, his real contribution in this essay is to point out some truths and debunk some myths about our political elections. The first, and perhaps most telling of all the truths, is that almost all presidential nominees in the twentieth century have taken approximately the same route to that position. According to Hess, all but one, Wendell Wilkie, could be called professional politicians (p. 38). Even Dwight Eisenhower, often cited as the principal example of the non-political president, was in in Hess's view the most political of generals (p. 43). Most nominees have started in politics early and worked their way through the system, not as legislative leaders, but as strong team players (p. 32). The Senate, the Vice-Presidency, and the State House have been the best stepping stones to the nomination (p. 104). That presidential aspirants and candidates are professional politicians does not pose a problem to Hess. In fact, people who choose politics as a profession are often best suited to be president. They tend to be gregarious, risk-takers, consensus builders, and open to the public spotlight that is so glaringly focused upon our presidents. Even among politicians, those who choose to run for president are an elite and self-selected group. These few must be willing to take enormous risks and put themselves through extreme physical and emotional hardship. In essence, they must really want the position. Those who do not have such a strong desire, do not run.

In this country we have tended to undervalue our politicians (pp. 5–9), viewing them as beneath most professionals and as inherently open to our suspicion. This is both good and bad, Hess argues. We should not hold politics out as the ultimate in professions, for to do so might drain the best and brightest away from more productive pursuits. However, to devalue it too much

will result in an inability to attract capable people into the profession from which our leaders inevitably will emerge. As Hess explains, "[I]f the menu of [presidential] choices does not include the best and the brightest, the failure lies in how society attracts its citizens into the political profession" (p. 40).

The biggest myth that Hess manages to debunk is that television has radically altered the election process. We have come to blame the media in general and television in particular for most of the ills in our electoral system, yet, as Hess spends a chapter pointing out, the media, particularly the use of television advertising, has not had a noticeable impact on results. "Advertising," Hess points out, "could not save six Republican senators in 1986 who outspent their opponents by an average of 75 percent and still lost their seats" (pp. 69–70). There is no doubt that television plays a role in the political process, and there is no doubt that TV has hurt some candidates, most notably Richard Nixon in 1960. But while Hess recognizes that one of the most important skills for a president to have is the ability to communicate his message to the people (p. 23), he does not believe that television has the power to select our candidates or our presidents. Hess notes, "[T]here is no evidence that free media—presenting news and other public service programs on TV—have caused the nomination or election of a single candidate" (p. 76). The failure of Bruce Babbitt's campaign this winter, one which attracted substantial positive coverage from all the networks and major newspapers, should reinforce this conclusion.

Furthermore, simply looking good on television is not enough to help a candidate. The most telegenic of recent presidential candidates, John Lindsay, who ran in 1972, was a complete bomb with the voters (p. 68). With the exception of Ronald Reagan and John Kennedy, those candidtates we have selected since television has become a crucial element in campaigns—Eisenhower, Stevenson, Kennedy, Nixon (twice), Goldwater, Johnson, Humphrey, McGovern, Ford, Carter, Reagan, and Mondale—have not been, as a group, any more or less attractive than any random group of white males in America (p. 103).

Television has had effects, certainly. Television time requires money, which has increased the need for fund-raising. Television allows candidates who manage to do well in early primaries to get a great deal of free publicity in larger states (for example, Gary Hart in 1984 or Jimmy Carter in 1976). Moreover, televi-

sion tends to trivialize the message of many candidates by forcing them to condense their message to the 100 seconds of the average television news story (p. 77). At the same time, however, television has drastically increased voters' opportunities to see the candidates. Prior to television, the vast majority of voters never saw or, prior to radio, heard the person for whom they were voting. They voted because of party affiliation or, less frequently, based on newspaper stories that were often as brief as television tidbits. Presently, the exposure is vast, particularly this year with the extraordinary number of televised debates that have been aired. While debates may not be the best test of political or presidential skills, Hess says, debates "give more Americans than ever before a greater opportunity to learn about those who would be president. This is probably the most positive thing to come out of the age of television" (p. 77).

Hess also debunks another favorite political myth: that presidential elections have become excessively long. There is no doubt that the electoral process now seems endless. Nonetheless, elections have always begun years before the actual vote. Andrew Jackson began to run three years prior to the 1828 election (p. 47), Franklin Roosevelt two years before the 1932 election, and John Kennedy began his presidential bid, for all intents and purposes, in 1956 after his senate victory (p. 48). It is just that television has made the campaigns seem longer due to the intense coverage so early in the process (pp. 48–49).

Having noted that television has not and probably will not dramatically change the outcomes of our elections, Hess addresses what is, for most of us, the key question that comes to mind when we look at our electoral process: "Why Great Men are Not Chosen Presidents" (p. 96). Hess notes that only in times of crisis have we elected presidents who have been considered truly great leaders. The examples always given are George Washington, Abraham Lincoln, and Franklin Roosevelt. Hess does not answer the question directly, but instead points to the extreme difficulty in obtaining the job and the difficulty our governmental structure creates for someone who really wants to rule from the White House (p. 9). Hess's reaction is not to bemoan this result, but to focus instead on how to increase the quality, and potentially the quantity of the pool of candidates. He bemoans the decline of the role of political parties in presidential selection because he believes they serve the functions of policing candidate ethics and assuring that party

affiliation means more than just a label. "In short . . . parties have an obligation to protect their brand names, and their consumers" (p. 95).

In the end, even this is mere tinkering.

> A change in process may have some effect on which contender wins a specific nomination, and some presidential attributes are tested by the process. But regrettably for voters, journalists, social scientists, and students, the process will neither predict nor determine the chances of the winners' turning out to be great presidents (p. 118).

Hess, then, has provided a good primer on the realities of presidential selection. Particularly in this unusual year when neither party's nomination will be held by an incumbent president, Hess's essay provides a useful reminder of the historical and practical realities of the process. This can inform our criticism, and, if we wish, our praise of the way we choose our presidents.

—*James W. Lowe*

THE PROSECUTORS: INSIDE THE OFFICES OF THE GOVERNMENT'S MOST POWERFUL LAWYERS. By *James B. Stewart*. New York: Simon & Schuster, 1987. Pp. 355, index. $19.95, cloth.

James B. Stewart has solved a mystery: Why do lawyers leave cushy jobs in law firms, take a fifty-percent pay cut, and go to work in government offices where they may have to type their own briefs? The answer, as Stewart shows in *The Prosecutors*, is that putting away criminals is good, clean fun.

Using the format that made his book *The Partners*[1] a bestseller, Stewart conveys a sense of the profession by devoting a chapter to each of six investigations. In both form and substance, he gives the prosecutorial response to Alan Dershowitz's book, *The Best Defense*.[2] Stewart purports to provide a "reasonably representative sample," though he skews his case studies to show white shoe criminals trying to outsmart celebrity

[1] J. STEWART, THE PARTNERS: INSIDE AMERICA'S MOST POWERFUL LAWFIRMS (1983).

[2] A. DERSHOWITZ, THE BEST DEFENSE (1982).

prosecutors.[3] Consequently, Stewart's table of contents reveals a caseload that would make any prosecutor salivate: Bribery at McDonnell Douglas; The Hitachi Sting; Insider Trading At Morgan Stanley; The CBS Murders; The Bank of Nova Scotia Tax Shelter Scheme; and The Investigation of Edwin Meese.

Stewart's made-for-television selection of cases is forgivable, however, since it makes for fascinating reading, and he links the seemingly disparate cases with several themes. He reveals the political aspects of the job, manifested internally as infighting, and externally by translating convictions into future votes. He also dissects the symbiotic relationship between the prosecution and the media. Finally, Stewart hints at what drives prosecutors. Unlike their television counterparts, Stewart's prosecutors spend only a fraction of their time in trial. Thus, the road from bail to jail detours through criminal discovery, grand jury hearings, and plea negotiations, among other less glamorous tasks. Stewart defines enough of the rudiments to assist the layman without boring the lawyer. He also explains tactical as well as legal consequences of the decisions a prosecutor confronts.

I. BRIBERY AT MCDONNELL DOUGLAS[4]

This chapter concerns the first major criminal prosecution of illegal foreign payments. The case's novelty magnified the workload, but the Assistant United States Attorneys (A.U.S.A.) who were assigned the case thought "we'll probably never have another case like this in our lives" (p. 34). McDonnell Douglas used bribes to obtain bids from the Pakistani state airline by tacking a $500,000 bribe per plane to each plane's price tag. The Foreign Corrupt Practices Act of 1977 did not cover the timespan of the bribes, so the prosecutors charged McDonnell Douglas with defrauding the Pakistani Government (p. 50) and its investors (p. 22). The company responded that bribery was an accepted business practice and a precondition to competing in the Middle East (p. 19).

The prosecutors traced the authorization for the bribes up the corporate ranks. Stewart unmasks the power prosecutors wield

[3] For example, much of the book discusses the United States Attorney for the Southern District of New York, Rudolph Giuliani.

[4] *See* United States v. McDonnell Douglas, No. 79-516 (D.D.C. Feb. 9, 1981).

in deciding if, when, and whom to indict (p. 46). The decision intertwines legal, tactical, and ethical dilemmas that point in different directions. For example, in considering whether to indict the eighty year-old founder and chief executive officer of McDonnell Douglas, an A.U.S.A. recalls: "You can't just ask yourself, 'Do I have the facts to indict this man?' You have to ask,'Should I, given his age and health?'" (pp. 42–43).

The McDonnell Douglas investigation also shows that the defense may not be the prosecutor's only opponent. The propriety of indicting individual executives in addition to the corporation generated debate within the Department of Justice (p. 37). Likewise, the A.U.S.A.s feared that their case was being "fixed" when they learned that Senator John Danforth[5] (D-Mo.) had called their boss, Rudolph Guiliani, who was at the time Associate Attorney General. Guiliani then met with defense counsel before notifying the A.U.S.A. With typical understatement, Stewart writes: "To the prosecutors, it looked like an ominously deliberate and secretive effort to get rid of the case" (p. 69). The A.U.S.A.s wrote a memo of protest to Guiliani and sent copies to the entire fraud section of the Department of Justice (p. 71). A series of leaks and counter leaks to the press escalated the tensions between Guiliani and the A.U.S.A.s (p. 74).[6]

The McDonnell Douglas investigation ended in 1981 with a plea bargain; the corporation pled guilty to fraud and false statements and paid a $55,000 fine and $1.2 million civil penalty. The indictments against the corporate executives were dropped (p. 83).

II. The Hitachi Sting[7]

In this chapter Stewart examines how the government tried to purge Silicon Valley of industrial espionage. Lack of hard evidence prompted the government to try infiltration. The result

[5] McDonnell Douglas employs many of Senator Danforth's constituents.

[6] Guiliani wrote a memo, which became public, to rebut the A.U.S.A.'s charges of impropriety. He contended that the assistants had "displayed a disrespect for the facts and an immature petulance that gives me pause as to the judgments they may have made. . . ." Furthermore, he noted that "no statute, regulation, policy or cannon of ethics prevent[s] an attorney from presenting his views to government officials at any level on matters of mutual concern" (p. 77).

[7] *See, e.g.*, United States v. Cadet, 727 F.2d 1453 (9th Cir. 1984).

was "Operation Pengem"[8] (p. 94). The Federal Bureau of Investigation (FBI) and the United States Attorney's Office for San Francisco set up a bogus computer brokerage firm which they called "Glenmar Associates." They learned that Hitachi had obtained top secret information belonging to IBM, one of their competitors. The prosecutors were "brimming with enthusiasm over what seemed to be Hitachi's positive eagerness to commit a crime." This provided the evidence of predisposition necessary to rebut an entrapment defense (p. 101).

The sting consisted of a secretly videotaped exchange of Hitachi executives paying $500,000 to Glenmar executives for sensitive IBM information (pp. 87–88). The government indicted Hitachi for transporting stolen property across state lines. As in the McDonnell Douglas case, the simplicity of the Hitachi indictment contrasts with the covert and international complexities of the investigation (p. 112).

One strange wrinkle of the case was that the A.U.S.A. assigned to the case refused to comply with a defense discovery request, even after a judge so ordered (p. 120). His act jeopardized the case and humiliated the San Francisco United States Attorney's Office. The incident may have been linked to IBM's assistance with the sting. The government may have refused to comply with the discovery request to avoid disrupting ongoing investigations or to hide a rumored connection between IBM's aid to Operation Pengem and the government's decision to drop the IBM antitrust suit (p. 122).[9] The defense exploited the rumors by arguing that IBM "had improperly infiltrated the executive branch of the United States government" (p. 121).

The Hitachi Sting ended in a plea bargain: The individual executives pleaded guilty and were promised that the government would not seek prison terms. The Hitachi corporation was fined $10,000—"a pittance" (p. 125). Stewart reports speculation of a secret settlement between Hitachi and competitor IBM, whereby Hitachi may have paid IBM $300 million. If true, Stewart suggests that the plea really was a bargain. Hitachi prevented a public trial that would have aired forty-eight hours of videotapes, showing its executives haggling over how much they would pay for stolen trade information (p. 126).

---

[8] The acronym stands for "Penetrate Gray Electronics Markets" (p. 94).

[9] Stewart finds the rumored quid pro quo unlikely (p. 121).

III. INSIDER TRADING AT MORGAN STANLEY[10]

Stewart opens this chapter with two good-looking investment bankers plotting "the [almost] perfect crime" while playing chess at the Harvard Club of New York. He examines how the United States Attorney's Office for the Southern District of New York investigated and prosecuted a case that became precedent for the onslaught of insider trading cases of the 1980's.[11]

The New York Stock Exchange (NYSE) stock watch had detected a pattern of stock trading suggesting that "a large number of people were receiving advance information about numerous targets of takeover attempts" (p. 137). It looked like a pyramid scheme. The United States Attorney's office took over the investigation because, unlike the Securities and Exchange Commission (SEC), it could immunize witnesses and thereby eliminate their need—or ability—to invoke their Fifth Amendment right against self-incrimination. The prosecutors spent seven years filling in the names at the top of the pyramid, gambling that the witnesses they immunized were not the scheme's ringleaders (p. 139). Stewart shows the intensive detective work that a prosecution entails. The A.U.S.A.s searched for evidence in high school applications, telephone and bank records (p. 141), and the *Harvard Business School Yearbook* (p. 164). The plot finally unraveled when the three original insider-traders broke their promise to share the information only with each other—their circle expanded to 75 (p. 165).

Questioning a target's employer, however, posed a dilemma for the prosecutor: "I knew I was about to take a major step, and I might ruin his life. It's not easy. I'm human, and I felt bad" (p. 144). Shortly thereafter, this future defendant was fired, his wife divorced him on his honeymoon and became a witness against him (p. 146).

The prosecutors structured their indictment to fit the hole that the Supreme Court left open in its *Chiarella* opinion[12] (p. 153): namely, the misappropriation theory. The defense argued that the charge's novelty showed that the prosecution had over-

---

[10] *See* United States v. Newman, 664 F.2d 12 (2d Cir. 1981), *cert. denied*, 464 U.S. 863 (1983).

[11] For an appraisal of the case's significance, see R. CLARK, CORPORATE LAW 353–54 (1986).

[12] The indictment was carefully phrased, because it was the first charge where the securites fraud victim was not a buyer or a seller, but rather an employer (p. 166). *Cf.* Chiarella v. United States, 445 U.S. 222 (1980).

reached, depriving the defendants of notice of the crime. This resulted in dismissal of the indictment and a series of pre-trial appeals. Meanwhile, the prime target absconded to Europe.

IV. THE CBS MURDERS[13]

This is the only chapter in which Stewart discusses assistant district attorneys (A.D.A.s) prosecuting state crimes. Three CBS employees were murdered in 1982 where they parked their cars on Pier 92 in New York City. Another CBS employee witnessed a man dragging an unknown woman's body in front of a van, and then saw his CBS colleagues, who went to investigate, shot in the head (p. 183). It is the story of a sleuth dodging the exclusionary rule. "The entire case probably turned on the outcome of the suppression hearing" (p. 202). Stewart focuses on how the A.D.A. obtained evidence needed to catch and convict the murderer, and then link the murders to a diamond fraud scheme which the victims had witnessed.

Little of the prosecutor's time was spent persuading juries. "It was the painstaking and sometimes mundane accumulation of the evidence—the poring over of phone records, the sifting through parking tickets, the detailed laboratory tests—that now made the conviction seem within reach" (p. 202). The A.D.A.'s persistence paid off when he saw the murderer sentenced to 100 years in prison, one of the longest sentences ever given in New York City. The ultimate reward, however, was the chance to do it again. The A.D.A., Gregory Waples, was assigned to prosecute Bernhard Goetz (p. 225).

Stewart suggests how prosecutors pay for the diligence required to win a conviction. The A.D.A. "began to spend his free evenings and weekends" compiling evidence (p. 201). Like his counterparts in New York's private law firms, the A.D.A. worked six or seven eighteen hour days a week (p. 185), yet has "none of the material amenities associated with the yuppies of his generation . . . and spends what little leisure time he has with a few friends and his cat" (p. 226). He was divorced during the case.

---

[13] *See, e.g.*, People v. Margolies, 480 N.Y.S.2d 842 (Sup. Ct. N.Y. County 1984).

V. THE BANK OF NOVA SCOTIA TAX SHELTER SCHEME[14]

On the chance that the subject of tax may make the readers eyes glaze over, Stewart begins this chapter in the Cayman Islands. Two businessmen concocted a "money circle" whereby they borrowed money, ostensibly to invest, and received a four-dollar deduction for every dollar invested. Their timing coincided with the Department of Justice Tax Division's effort to showcase a "major tax shelter crackdown" (p. 238). Stewart shows the fine line between creative prosecutorial theories and overreaching. It was not clear that the scheme was illegal, but "the whole thing smelled bad" (p. 242). One of the businessmen was Denver entrepreneur William Kilpatrick; thus much of the investigation was centered in Denver. This created a turf battle between the Department of Justice tax lawyers from Washington and the local United States Attorney's Office (p. 247).

This chapter focuses not on the crime or the investigation, but on how a case can explode in a prosecutor's face. It is almost too painful to read. In particular, Stewart discusses the relationship between the prosecution and District Court Judge Fred Winner, to whom the case was assigned, and who *The American Lawyer* rated the "worst federal district court judge in the Tenth Circuit" (pp. 246–47).

The defense attorneys deflected attention from their client by alleging prosecutorial misconduct, thus putting the government on trial. When twenty-six of the twenty-seven counts of the indictment were dismissed on substantive grounds, the prosecutors "took the macho approach" and tried the single remaining count, knowing it would enrage the judge. During the trial, "relations between the prosecutors and the judge deteriorated steadily . . . the judge also evidenced a growing interest in some of the defense's allegations of prosecutorial misconduct" (p. 268). The tension errupted into a courtroom shouting match, with Judge Winner castigating "the lawyers from the banks of the Potomac" (p. 270). When the jury convicted Kilpatrick on the remaining count, Judge Winner invited motions on a new trial and granted an evidentiary hearing on prosecutorial misconduct. The Department of Justice refused the prosecutors' request for permission to move to disqualify Judge Winner. The

[14] *See, e.g.*, United States v. Kilpatrick, 821 F.2d 1456 (10th Cir. 1987); Blondin v. Winner, 822 F.2d 969 (10th Cir. 1987).

Denver United States Attorney's Office, meanwhile, stopped trying "to wrest control of the case from Washington" (p. 272). Even before the government presented its witnesses, Judge Winner announced that he had heard enough. He threw out the jury's guilty verdict, granted a new trial, and remanded the case to the lower court. Moreover, his opinion "lambasted" the prosecutors for eighteen pages.

The Department of Justice filed a motion asking Judge Winner to reconsider or delete the portion of his opinion attacking the prosecutors, given that the allegations were still unproven. He refused. The Solicitor General approved, recommending appeal of the judge's refusal to withdraw his opinion. "But the Justice Department didn't stop with its motion. It went ahead and shot itself in the foot" (p. 278). It asked West Publishing to omit the advance sheets containing the opinion from West's bound volume. West declined. The Department of Justice then requested the Court of Appeals for the Tenth Circuit to "halt publication" of the opinion. Shortly thereafter, *The New York Times* accused the Department of Justice of seeking a prior restraint. The Tenth Circuit reversed itself, the opinion was published, and Congress launched an inquiry. Following a full evidentiary hearing on remand before a new judge, the final tax evasion count was dismissed and the A.U.S.As were found guilty of prosecutorial misconduct and of interviewing witnesses outside the presence of their attorney (p. 282).

Stewart demonstrates that most prosecutors have enough ego and ambition to enjoy and even seek media exposure. In this chapter, however, the prosecutors sought publicity to save rather than advance their careers. When the Department of Justice forbade the A.U.S.A.s from appearing on the television show "60 Minutes," one prosecutor quit so that he could publicly defend himself. However, a fallout with the CBS producer prevented the interview, and the story aired without mentioning that the Department of Justice Office of Professional Responsibility had largely vindicated the charges of prosecutorial misconduct (p. 286).

VI. The Investigation of Edwin Meese

Stewart's tendancy to gossip, barely contained throughout the book, runs rampant in this chapter. We learn, for example, that

for their honeymoon, Mr. and Mrs. Edwin Meese took a tour of California state prisons (p. 289).

The chapter lacks focus and is premature. Stewart never clarifies whether he is evaluating Attorney General Meese as the nation's top prosecutor, whether he is examining the function of a Special Prosecutor, or whether he is exploring how prosecutors prosecute fellow prosecutors. He fails on all three fronts. Furthermore, this chapter remains literally unfinished. As of this writing, Attorney General Meese's legal travails continue,[15] and the constitutionality of the Special Prosecutor's Office remains unclear.

Stewart reiterates his theme about the political nature of prosecution. He notes that Senator Joseph R. Biden, Jr. (D-Del.), Chairman of the Senate Judiciary Committee, wrote Special Prosecutor Jacob Stein "asking explictly that he render an opinion concerning the 'propriety or ethics' of Meese's conduct, and not limit the allegations to allegations of criminal conduct" (p. 304). Stewart makes a chilling observation that several prosecutors strayed from their mandate to investigate whether Mr. Meese had broken any federal laws and instead dwelled on his qualifications to serve as Attorney General: "[T]hey thought it would be a public service to bring Meese down, even if he were eventually acquitted" (p. 341). Otherwise, this chapter drones on about cuff links and loans without adding any insight to the nature of prosecution.

Perhaps the investigation of Mr. Meese is an appropriate place to end *The Prosecutors*, because this chapter illustrates that a lawyer who decides to become a prosecutor swaps money for power. Stewart describes one prosecutor's self-perception as "a modern cowboy, doing today's equivalent of frontier justice" (p. 354). The job is inherently romantic. Who else would visit prisons on their honeymoon?

—*Bridget K. Maloney*

---

[15] *See, e.g.*, New York Times, Mar. 11, 1988, at A19, col. 3.

ERRATA

The editors would like to correct the following editorial errors occurring in Blum, *New Role for the Treasury: Charging Interest on Tax Deferral Loans*, Vol. 25, Number 1 (Winter 1988).

(1) The Author's Note printed at pages 2–3 is an earlier version of the following Author's Note, which should have appeared:

Author's Note

As this Article went to press, Congress enacted legislation that extends provisions for charging interest on tax deferral loans, as the Article recommends. In the Omnibus Budget Reconciliation Act of 1987, signed by the President on December 22, 1987, Congress extended the requirement that long-term contracts be reported under the percentage of completion method, and that errors be corrected with a lookback interest charge or credit, to 70% of the items under the contract (as compared with 40% under prior law). In addition, Congress imposed an interest charge to compensate for tax deferral enjoyed by use of the installment method with respect to a disposition of real property used in the taxpayer's trade or business or held for the production of rental income if the sales price is in excess of $150,000. The interest charge, however, applies only to the extent that the aggregate face amount of all such installment obligations that arise during a taxable year and are outstanding as of the close of that year exceeds $5 million. At the same time, Congress repealed the proportionate disallowance rule, that had limited the application of the installment method in some cases, but Congress provided special treatment of pledges of installment obligations.

Congress repealed the use of the installment method with respect to dispositions by dealers in personal or real property, thereby eliminating the need for an interest charge compensating for deferral of tax by such dealers. The legislation, however, retains the option under prior law for a dealer to use the installment method with respect to certain sales of time shares and residential lots if he pays an interest charge compensating for tax deferral.

The expanded use of interest provisions approved by Congress suggests the viability of this approach in other contexts as

well, such as professional fees and gain or loss from marketable securities, as discussed in the Article. Moreover, there is a need to focus on issues relating to the proper design of an interest charge under existing and possible new provisions, as is done in the Article.

Because the 1987 legislation was enacted as this Article went to press, the 1987 provisions are discussed in an Addendum to the Article, and the main body of the Article does not reflect these provisions.

(2) In footnote 51, on page 18, the following should appear after the second sentence:

But cf. BLUEBOOK, *supra* note 4, at 266 (personal interest "includes interest on underpayments of individual Federal . . . income taxes notwithstanding that all or a portion of the income may have arisen in a trade or business").

(3) In footnote 59, on page 21, in the second sentence, after the first comma, the following should appear:

in the case of a noncorporate taxpayer,

(4) In footnote 59, on page 22, the following should appear after the carryover sentence:

See Temp. Treas. Reg. § 1.163-8T(c)(1), 52 Fed. Reg. 24996, 25000 (July 2, 1987) (debt allocated to expenditures in accordance with use of debt proceeds).

(5) In footnote 255, on page 78, the beginning portion of the fifth line should read:

between taxpayers' standards of living).

(6) In footnote 298, on page 91, the citation in the second to the last line of the footnote should read:

See also Note, *Tax Reform Act*, *supra* note 2, at 794 & n.105

(7) On page 105, in the first paragraph of text, the reference to section 453(e)(4) should be to section 453C(e)(4).

(8) In the Addendum, at footnote 365, on page 108, the portion of the footnote coming after the sentence reading "*See supra* notes 51 & 59." should be replaced with the following:

See also Temp. Treas. Reg. § 1.163-9T(b)(2), 52 Fed. Reg. 48407, 48409 (Dec. 22, 1987) (personal interest includes interest paid on underpayments of individual Federal income taxes). The Temporary Regulations, issued as this Article went to press, provide further that personal interest includes (1) interest paid on indebtedness used to pay individual Federal income taxes, and (2) any interest charge paid under section 453C(e)(4) or section 1291(c) of the Code. *See id.* For an argument that the interest charge on a tax deferral loan should either be deductible or be determined at an after-tax interest rate, see *supra* notes 50–51 & 59 and accompanying text.