HARVARD JOURNAL
on
LEGISLATION

VOLUME 20, NUMBER 2 SUMMER 1983

PREFACE

STATUTES

ARTICLES

COMMENTS

BOOK REVIEWS

PREFACE

THE EXPLOSIVE GROWTH OF LAW THROUGH LEGISLATION AND THE NEED FOR LEGISLATIVE SCHOLARSHIP

ERWIN N. GRISWOLD*

From the modern point of view, the development of law in the nineteenth century was relatively simple. The approach was largely in analytical and historical terms. Students absorbed the works of Professors Austin and Maine, and practitioners gave little thought to anything but the cases. Dean Langdell treated cases as the basis for the scientific study of the law, while Dean Ames, as Dean Pound has pointed out, allowed "[i]n all his books references to legislation regularly [to] take the form of references to cases."[1] Justice Holmes did put emphasis chiefly on "experience,"[2] but this was primarily the experience of mankind, as reflected in the courts, and he did not contemplate much development of the law through legislation.

With the twentieth century came a greater emphasis on a "sociological" approach, a form of analysis championed by Dean Pound.[3] Pound recognized Holmes as "at least the forerunner of sociological jurisprudence in the United States,"[4] and further noted that "[i]n American sociological jurisprudence the outstanding work is that of Mr. Justice Cardozo."[5] In due course, sociological jurisprudence developed in a somewhat disordered way into "realism"—which may be the dominant approach today. But legal realism is hardly a system or a coherent approach to law. It is rather a device to encourage the judge to do what he wants or thinks wise, and continuity and predictability have largely disappeared from the law as determined and administered by the courts.

---

*Mr. Griswold was the Dean of the Harvard Law School from 1946 to 1967. He strongly supported the founding of the *Harvard Journal on Legislation* in 1963, and wrote a preface for its first issue. He served as Solicitor General of the United States from 1967 to 1973, and currently is a partner with Jones, Day, Reavis & Pogue in Washington, D.C.

[1] 3 R. POUND, JURISPRUDENCE 655 n.2 (1959).

[2] O.W. HOLMES, THE COMMON LAW 1 (1881).

[3] *See* Pound, *The Scope and Purpose of Sociological Jurisprudence* (pts. 1–3), 24 HARV. L. REV. 591 (1911), 25 HARV. L. REV. 140, 489 (1911–1912).

[4] 1 R. POUND, *supra* note 1, at 337.

[5] *Id.* at 338.

A greater change, however, that has become dominant in the past twenty years is the extraordinary increase in the development of law through legislation. Of course, there always have been statutes, and occasionally fundamental statutes like the Statute of Wills or the Statute of Frauds. About a century ago we began to enact statutes that developed the machinery of government, such as the Interstate Commerce Act of 1887 and the Sherman Act of 1890. Both of these statutes, in accordance with Holmes' dictum, arose out of "experience," but it was thought that the courts could not deal with this experience without legislative help. And in the early twentieth century we had the Packers and Stockyards Act, the Federal Trade Commission Act, the Clayton Act, and the Transportation Act, 1920. This legislative development of the law grew in the 1930's with the passage of the National Labor Relations Act, the Securities Acts, the Federal Communications Act, and the Internal Revenue Code of 1939, which for the first time sought to stabilize our tax law.

Development of the law through statutes continued during the post-war period with the Federal Rules of Civil Procedure, the Federal Rules of Criminal Procedure, the Uniform Commercial Code, the proliferation of regulations of administrative agencies (which are treated in many ways like statutes), the spread of zoning ordinances, and a host of other legislative developments. Much as those trained on the case method were prone to resist, our law was, before our eyes, becoming more and more stated in authoritative texts. Litigation, in appellate courts especially, increasingly became a matter of the interpretation of statutory provisions. Distilling meaning out of normative legislative or administrative commands—through the use of legislative history and other techniques—had little relation to the "unfolding" of the common law, which occurred through the development and the extension of principles.

Many fields of the law, of course, still follow a common-law approach to a considerable extent. One of these fields is negligence, although that area has been extensively preempted by workers' compensation statutes—and there is now a proposal for a federal statute on products liability. Slowly but surely, much of our law either has been taken over, or has been deeply invaded, by statutes. We all know that fact, yet we are perhaps not fully aware of it in our habitual processes of legal thinking.

In the past twenty years the development of the law through statutes has burgeoned, particularly in the ever-expanding federal field. The list of statutes is great and constantly growing. The promulgation of the following statutes and executive orders has made major changes in the law, not only in litigation before the courts, but also in the place where most law is first administered—in the law offices.

Civil Rights Act of 1964, Pub. L. No. 88-352, 78 Stat. 241.

Urban Mass Transportation Act of 1964, Pub. L. No. 88-365, 78 Stat. 302.

Exec. Order No. 11,246 (Federal Contract Compliance Program), 41 C.F.R. § 60 (1965).

Freedom of Information Act, Pub. L. No. 89-487, 80 Stat. 250 (1966).

Fair Housing Act of 1968, Pub. L. No. 90-284, 82 Stat. 73, 83.

Williams Act (relating to corporate takeovers), Pub. L. No. 90-439, 82 Stat. 454 (1968).

National Environmental Policy Act of 1969, Pub. L. No. 91-190, 83 Stat. 852 (1970).

Economic Stabilization Act of 1970, Pub. L. No. 91-379, tit. 2, 84 Stat. 796, 799–800.

Occupational Safety and Health Act of 1970, Pub. L. No. 91-596, 84 Stat. 1590.

Clean Air Act Amendments of 1970, Pub. L. No. 91-604, 84 Stat. 1676.

Presidential Election Campaign Fund Act, Pub. L. No. 92-178, § 801, 85 Stat. 497, 562 (1971).

Equal Employment Opportunity Act of 1972, Pub. L. No. 92-261, 86 Stat. 103.

Educational Amendments of 1972, Pub. L. No. 92-318, § 901, 86 Stat. 235, 373–74.

Federal Water Pollution Control Act, Pub. L. No. 92-500, 86 Stat. 816 (1972).

Noise Control Act of 1972, Pub. L. No. 92-574, 86 Stat. 1234.

Emergency Petroleum Allocation Act of 1973, Pub. L. No. 93-159, 87 Stat. 627.

Employee Retirement Income Security Act of 1974 (ERISA), Pub. L. No. 93-406, 88 Stat. 829.

Federal Election Campaign Act Amendment of 1974, Pub. L. No. 93-443, 88 Stat. 1263.

Commodity Futures Trading Commission Act of 1974, Pub. L. No. 93-463, 88 Stat. 1389.

National Health Planning and Resources Development Act of 1974, Pub. L. No. 93-641, 88 Stat. 2225 (1975).

Federal Insecticide, Fungicide and Rodenticide Act of 1975 (FIFRA), Pub. L. No. 94-140, 89 Stat. 751.

Education for All Handicapped Children Act of 1975, Pub. L. No. 94-142, 89 Stat. 773.

Hart-Scott-Rodino Antitrust Improvements Act of 1976, Pub. L. No. 94-435, 90 Stat. 1383.

Surface Mining Control and Reclamation Act of 1977, Pub. L. No. 95-87, 91 Stat. 445.

International Emergency Economic Powers Act, Pub. L. No. 95-223, tit. 2, 91 Stat. 1625, 1626–29 (1977) and Exec. Orders No. 12,276–12,285, 46 Fed. Reg. 7913–32 (1981); Exec. Order No. 12,294, 46 Fed. Reg. 14,111 (1981).

Futures Trading Act of 1978, Pub. L. No. 95-405, 92 Stat. 865.

Bankruptcy Reform Act of 1978, Pub. L. No. 95-598, 92 Stat. 2549.

Public Utility Regulatory Policies Act of 1978 (PURPA), Pub. L. No. 95-617, 92 Stat. 3117.

Economic Recovery Tax Act of 1981, Pub. L. No. 97-34, 95 Stat. 172.

Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA), Pub. L. No. 97-248, 96 Stat. 324.

This list is by no means complete or comprehensive. It has been compiled in large part simply by going through the opinions in cases decided by the Supreme Court in October Term, 1981 (that is, between October 1981 and July 1982). This list is striking enough, but it does not begin to disclose the extraordinary complexity of some of the statutory provisions, of which the National Health Care Planning and Resources Development Act, FIFRA, ERISA, and TEFRA may be cited as monumental examples, not to mention the present state of the Internal Revenue Code. Nor does this list include many amendments to these legislative and executive actions.

The causes of this growth in the development of law through statutes are not wholly clear. The ever-increasing size and bureaucratic tendencies of congressional staffs may play an important role, or they may be a result. Increasing demands from the public—or from special interest groups—for the government to do something—or to give something—have undoubtedly had their impact, but of course these demands are simply a part of the political process in a free and increasingly active society.

The statutes listed above occupy more than 1500 pages in the *Statutes at Large* and the *Code of Federal Regulations*. That volume is not much compared with the case law, which includes

the nearly 1000 pages each week to which the *Federal Reporter, Second Series* has now grown—not to mention all the other federal reporters, the reporters of all fifty states, and the reports of many state and federal administrative agencies. But the statutory provisions are far more concentrated. Statutory provisions, whatever they mean, do not just talk about the law; they *are* the law. And some of them are of incredible and, it would seem, inexcusable intricacy. Anyone who doubts this should take a look at section 2001 of ERISA, section 240 of TEFRA, or any of a good many provisions of the Internal Revenue Code.

Statutes are not the whole law, of course, but they have been a principal cutting edge of the law for much of the past twenty years. Problems arising under these statutes and orders, together with constitutional questions, have occupied much of the time and the attention of the Supreme Court. It should not be overlooked that constitutional questions also involve construction of a written text, though in the light of the fact that it is a "constitution" that is being interpreted. The process of interpreting statutes is not particularly stylized, and is subject to many of the political and social pressures that attended the "unfolding" of the common law. But it is a different process, with different premises; a system of law that is heavily based on statutes is different from that of the common law to which, over the years, much of legal education, and legal thinking, have been devoted.

For a long time, these developments had little impact at the Harvard Law School, and on legal education generally. Dean Pound taught a course on legislation in the 1920's, and published an outline of the course.[6] James M. Landis, after two years as assistant professor on the faculty, was appointed Professor of Legislation in 1928. He served in that capacity until his departure for Washington in 1933. Probably the greatest fruit of Landis' focus on legislation was his remarkable essay on *Statutes and the Sources of Law*.[7] The *Harvard Law Review* from time to time had *Notes on Current Legislation*,[8] and it even had a

[6]He had written earlier in the field. *See* Pound, *Common Law and Legislation*, 21 HARV. L. REV. 383 (1908); *see also* R. POUND, THE FORMATIVE ERA OF AMERICAN LAW 52–60, 93–97 (1938).

[7]Landis, *Statutes and the Sources of Law*, in HARVARD LEGAL ESSAYS 213 (1934), *reprinted in* 2 HARV. J. ON LEGIS. 7 (1965).

[8]*See, e.g.*, Note, *The Federal Wages and Hours Act*, 52 HARV. L. REV. 646 (1939); Note, *Revision of the Copyright Law*, 51 HARV. L. REV. 906 (1938).

significant *Developments* section on an important instance of innovative legislation.[9]

Legislation, however, remained the outsider, the disrupter of the more placid waters of the common law. I noticed this in my teaching, when I essayed to teach Federal Taxation. When questions were put to the students, almost all of them immediately began to think "great thoughts," and to raise questions of principle or analogy as they would do in a class on torts or contracts. I worked hard to get them to "look at the statute." I asked, "What does the statute say?"—and if there was nothing relevant in the statute, then the question was, "What does the regulation say?" No one had ever taught me about regulations when I was in law school, and I made considerable effort to help students to become familiar with them, to accept them, and to use them. A good many students have told me that they found that part of their legal education to be useful and of continuing value.

It was in this rather quiet situation that the *Harvard Journal on Legislation* made its appearance twenty years ago. The moving force for the idea came from the newly established Harvard Student Legislative Research Bureau, which was organized and operated by students to assist state legislators, city councilmen, charitable organizations, and others by preparing carefully considered drafts of statutes and ordinances to be introduced in legislative bodies. Some of these drafts were very good and deserved being published for a wider audience. The work was educational, rather novel, and it made legislative research and writing experience available to a larger group of students.

According to my count, the present issue of the *Harvard Journal on Legislation* is the sixty-seventh number of the *Journal*. At times its board has had difficulty getting issues out on time. And the financial problems of a student-edited law journal are considerable, and take a good deal of understanding from a hard-pressed dean. But, in my view, the *Journal* has been a great success. It has proved itself over the years and has developed increased stature. It also has helped by now to train several hundred students to have greater familiarity with statutes, and to accept them not only as wholly legitimate parts of our law, but as the normal and natural way in which our law develops.

---

[9]*Developments in the Law--Employment Discrimination and Title VII of the Civil Rights Act of 1964*, 84 HARV. L. REV. 1109 (1971).

It would serve no purpose to review the materials that have appeared in the *Journal*. They are readily available in most law libraries. They have included 150 major articles, an equal number of student notes and comments, an occasional draft of a statute by the Harvard Legislative Research Bureau, and reviews of and useful references to perhaps 200 books. The quality has been kept at a high level, and material in the *Journal* often has been cited by courts in cases where it is relevant. But that is not the real test of the *Journal*'s usefulness. The great opportunity for the *Journal* is to improve the content and drafting of statutes, particularly by finding ways to reduce the blinding complexity with which many of our statutory provisions are now obscured. We need, too, to find ways to temper the influence of special interest groups in the drafting and the enactment of statutes. We should consider readopting some procedures that were established in the nineteenth century to prevent important legislative matters from being "sprung": to make it impossible, for example, for substantive legislation to be added to appropriations bills or to be tacked as riders to other legislation with which it has no relation.[10] Legislation is surely here to stay, and it needs to be studied carefully. Ways need to be found to develop it as well as possible, and to keep it useful and coherent.

The *Journal* has made good progress down an important road. The path ahead of it is a challenging one. I am glad to salute to the *Journal* on its Twentieth Anniversary.

---

[10]Some other problems that might be explored include: the extent to which legislative provisions are adopted "in principle" by committees, including conference committees, and then are drafted by staff members without any member of the legislative body ever seeing the exact text of the language actually inserted in the bill (e.g., the adoption of the carry-over basis provision in the Tax Reform Act of 1976); and the extent to which committee reports, which often are relied upon heavily in the construction of statutes, actually are written by congressional staff members or by employees of the Executive Branch, perhaps even after the statute has been passed, and thus are never seen by the Members of Congress who enact the statute. It may be that statements on the floor of the House and the Senate by Members of Congress in charge of the bill are of far greater significance than final committee reports, particularly reports of conference committees.

STATUTE

A MODEL STATE LAW ON CIVIL COMMITMENT OF THE MENTALLY ILL

CLIFFORD D. STROMBERG*
ALAN A. STONE**

Involuntary commitment of mentally disordered persons has been the subject of unceasing litigation and legislative activity during the last decade. The problems of lax civil commitment laws and "warehousing" of patients in large institutions for long periods without effective treatment have been answered by reforms directed toward improving institutional conditions, transferring patients to community facilities, and regarding dangerousness as the only valid basis for commitment. Mental patients' "right to treatment," "right to refuse treatment," and "right to the least restrictive alternative" have dominated the debate and transformed state commitment laws.

In this Article, Mr. Stromberg and Dr. Stone argue that these necessary reforms have given rise to a second generation of problems as formidable as the first. These new problems include abandonment of severely mentally ill persons, custodial care without treatment for patients who refuse treatment, conversion of mental hospitals into jails, failure to protect society, frequent lawsuits against mental health professionals, and continuing confusion about the goals of civil commitment. Mr. Stromberg and Dr. Stone present here a Model Law to address these problems that was developed and approved by the American Psychiatric Association. The authors also provide a commentary in which they explain how each of the Model Law's provisions can help states begin to solve the emerging problems of the civil commitment system in the 1980's.

The laws governing civil commitment of persons afflicted with mental disorders have undergone drastic revisions during the past two decades. Prior to that, state laws generally had authorized commitment simply on the basis that a person was "mentally ill and in need of treatment." Due process was largely dispensed with on the ground that the goal of commitment was therapy, not punishment. Mental institutions during this period were a backwater of social policy, basically ignored by legislators, policymakers, and lawyers. Police officers, public hospitals, and families demanded that some place be found for dis-

*Arnold & Porter, Washington, D.C. B.A., Yale University, 1971; J.D., Harvard Law School, 1974. Mr. Stromberg was formerly Deputy Executive Secretary of the U.S. Department of Health and Human Services, and a consultant on health law and policy to the Federal Trade Commission.

**Touroff-Glueck Professor of Law and Psychiatry in the Faculty of Law and the Faculty of Medicine, Harvard University. B.A., Harvard University, 1950; M.D., Yale Medical School, 1955. Dr. Stone is a former President of the American Psychiatric Association.

ruptive mentally ill persons. As a result, states committed many people for long periods, often to large "mega-institutions" with inadequate medical staff, minimal treatment, and disgraceful living conditions.

Addressing the plight of mental patients became a major legal and social cause during the 1960's and 1970's. It became the vogue to assert that mental illness was largely a "myth" that was manufactured by psychiatrists, that mental patients were "prisoners of psychiatry," and that mental hospitals were all like that portrayed in *One Flew Over the Cuckoo's Nest.*[1] Legal activists saw mental patients as another class of people who, like blacks, had been denied basic civil rights. Legal reform therefore was directed towards removing psychiatrists' discretionary authority, restricting involuntary commitment to persons demonstrably dangerous to themselves or to others, and applying to civil commitment all the due process safeguards of the criminal justice system.[2]

In 1975 the Supreme Court addressed the civil commitment issue in *O'Connor v. Donaldson.*[3] Donaldson had been diagnosed as mentally ill and had been confined for fourteen years without ever receiving meaningful treatment. The Supreme Court ruled that "[a] finding of 'mental illness' alone cannot justify a State's locking up a person against his will and keeping him indefinitely in simple custodial confinement."[4] In succeeding years, many lower courts construed *Donaldson* to mean that civil commitment statutes were unconstitutional unless they required proof of dangerousness to self or others in every case.[5] The courts struck down laws permitting commitment simply based on a person's being "in need of treatment" or on com-

---

[1]K. KESEY, ONE FLEW OVER THE CUCKOO'S NEST (1962); *see also* B. ENNIS, PRISONERS OF PSYCHIATRY (1972); T. SZASZ, THE MANUFACTURE OF MADNESS (1970); T. SZASZ, THE MYTH OF MENTAL ILLNESS (1961).

[2]*See, e.g.*, Lessard v. Schmidt, 349 F. Supp. 1078 (E.D. Wis. 1972), *vacated and remanded*, 414 U.S. 473 (1974), *judgment reinstated*, 413 F. Supp. 1318 (E.D. Wis. 1976); A. STONE (WITH C. STROMBERG), MENTAL HEALTH AND LAW: A SYSTEM IN TRANSITION 47–59 (1976).

[3]422 U.S. 563 (1975)

[4]*Id.* at 575. The Court also held that "a State cannot constitutionally confine without more a nondangerous individual who is capable of surviving safely in freedom by himself or with the help of willing and responsible family members or friends." *Id.* at 576.

[5]*See, e.g.*, Suzuki v. Yuen, 617 F.2d 173 (9th Cir. 1980); Benham v. Edwards, 501 F. Supp. 1050 (N.D. Ga. 1980); Doe v. Gallinot, 486 F. Supp. 983 (C.D. Cal. 1979); Colyar v. Third Judicial Dist. Court, 469 F. Supp. 424 (D. Utah 1979); Bension v. Meredith, 455 F. Supp. 662 (D.D.C. 1978); Bethany v. Stubbs, 393 So. 2d 1351 (Miss. 1981).

mitment being "in the best interests of the patient."[6] The courts also began to rule that mental patients have a right to be placed in the "least restrictive alternative" facility or treatment program. During this period, a series of judicial orders and consent decrees in cases asserting mental patients' "right to treatment" transformed the conditions of care and treatment at many state hospitals.

These legal reforms, along with new legislation providing financial support for "deinstitutionalization," had a major impact on civil commitment and on the mentally ill population. The average number of persons subject to commitment in state and county mental hospitals per day declined from 560,000 in 1955 to 276,000 in 1972, and to about 138,000 in 1981.[7]

These reforms were necessary, but they bred a second generation of problems as serious as the first. One has only to ask: what happened to all the people who were not committed or who were released? Some, of course, never had any serious mental illness and should not have been hospitalized. A larger number, though seriously mentally ill, could benefit more from outpatient treatment while living in community facilities or with their families.[8] But many others were simply "trans-institutionalized," finding their way into jails, prisons, and nursing homes where therapy and custodial care were even worse than in mental hospitals.[9] Finally, the mental health system simply lost track of hundreds of thousands of other former patients.

---

[6]*See, e.g.*, Johnson v. Solomon, 484 F. Supp. 278 (D. Md. 1979); Goldy v. Beal, 429 F. Supp. 640 (M.D. Pa. 1976); Stamus v. Leonhardt, 414 F. Supp. 439 (S.D. Iowa 1976); Commonwealth *ex rel.* Finken v. Roop, 234 Pa. 155, 339 A.2d 764 (1975) (plurality opinion), *appeal dismissed*, 424 U.S. 960 (1976). *See generally* Stromberg, *Developments Concerning the Legal Criteria for Civil Commitment: Who Are We Looking For?*, in PSYCHIATRY 1982, 334 (L. Grinspoon ed. 1983).

[7]Eisenberg, *Psychiatric Intervention*, 229 SCI. AM. 117, 118 (1973); U.S. DEP'T OF HEALTH & HUMAN SERVICES, TOWARD A NATIONAL PLAN FOR THE CHRONICALLY MENTALLY ILL: REPORT TO THE SECRETARY BY THE DEPARTMENT OF HEALTH AND HUMAN SERVICES STEERING COMMITTEE ON THE CHRONICALLY MENTALLY ILL (1981) [hereinafter cited as HHS REPORT ON THE CHRONICALLY MENTALLY ILL]; Goldman, Adams & Taube, *Deinstitutionalization: The Data De-Mythologized*, 34 HOSP. & COMMUNITY PSYCHIATRY 129 (1983); Klerman, *National Trends in Hospitalization*, 30 HOSP. & COMMUNITY PSYCHIATRY 110 (1979).

[8]*See* Lamb & Goetzel, *The Long Term Patient in the Era of Community Treatment*, 34 ARCHIVES GEN. PSYCHIATRY 679 (1977); Langsley, Machotka & Flomenhaft, *Avoiding Mental Hospital Admission: A Follow-Up Study*, 127 AM. J. PSYCHIATRY 1391, 1394 (1971); Stein, Test & Marx, *Alternative to the Hospital: A Controlled Study*, 132 AM. J. PSYCHIATRY 517 (1975).

[9]*See* U.S. GEN. ACCOUNTING OFFICE, RETURNING THE MENTALLY DISABLED TO THE COMMUNITY: GOVERNMENT NEEDS TO DO MORE (1977); Bonavitz & Guy, *Impact of Restrictive Civil Commitment Procedures on a Prison Psychiatric Service*, 136 AM. J. PSYCHIATRY 1045 (1975); Haupt & Ehrlich, *The Impact of a New State Commitment*

The rate of readmission of former patients rose markedly during the 1970's.[10] Today's clientele includes large numbers of "new chronic" patients who live in the community, but constantly face "readmission, alcohol abuse, suicidal propensities and ineffectively treated schizophrenia."[11] In every city, many severely impaired mentally ill people live in "welfare hotels," flophouses, abandoned buildings, and alleys.[12] As one judge asked rhetorically, "[H]ow real is the promise of individual autonomy for a confused person set adrift in a hostile world?"[13]

This situation is a social disaster. To alleviate it will require an array of strategies, including providing funding for many new facilities and programs. Another key element will be fashioning revised commitment laws that address the new generation of mental health problems, without sacrificing the gains that have been made.[14] This will require transcending the polemical debate

Law on Psychiatric Patient Careers, 31 HOSP. & COMMUNITY PSYCHIATRY 745 (1980); Swank & Winer, *Occurrence of Psychiatric Disorder in a County Jail Population*, 133 AM. J. PSYCHIATRY 1331 (1976); Treffert, *Dying with Their Rights On*, 130 AM. J. PSYCHIATRY 1041 (1973); Warren, *New Forms of Social Control: The Myth of Deinstitutionalization*, 24 AM. BEHAV. SCIENTIST 724 (1981).

[10]*See* M. GREENBLATT, PSYCHOPOLITICS 162–63 (1978).

[11]Schwartz & Goldfinger, *The New Chronic Patient: Clinical Characteristics of An Emerging Subgroup*, 32 HOSP. & COMMUNITY PSYCHIATRY 470 (1981).

[12]An estimated 40,000 homeless mental patients live in New York City alone. Shipp, *Suit on Homeless Mental Patients Asks New York State for Housing*, N.Y. Times, May 21, 1982, at A1, col. 8; *see also* Arnhoff, *Social Consequences of Policy Toward Mental Illness*, 188 SCI. 1277 (1975); Bassuk & Gerson, *Deinstitutionalization and Mental Health Services*, 238 SCI. AM. 46 (1978); Kirk & Therrien, *Community Mental Health Myths and the Fate of Former Hospitalized Patients*, 38 PSYCHIATRY 209 (1975); Reich & Siegel, *The Emergence of the Bowery as a Psychiatric Dumping Ground*, 50 PSYCHIATRIC Q. 191 (1978); Scull, *A New Trade in Lunacy: The Recommitment of the Mental Patient*, 24 AM. BEHAV. SCIENTIST 741, 743–44 (1981). *See generally Community Support for Mental Patients: Hearings on Programs in the Community for the Chronically Mentally Ill Before the Subcomm. on Health and the Environment of the House Comm. on Interstate and Foreign Commerce*, 96th Cong., 1st Sess. (1979); *Mental Health Systems Act: Hearings on H.R. 4156 Before the Subcomm. on Health and the Environment of the House Comm. on Interstate and Foreign Commerce*, 96th Cong., 1st Sess. 65 (1979) (statement of Joseph A. Califano, Jr., Secretary, Department of Health, Education and Welfare (DHEW)) ("For many long-time residents of mental institutions, deinstitutionalization has meant being dumped into communities to live in poorly maintained boarding homes or single room occupancy hotels—often without adequate food or clothing.").

[13]Bazelon, *Institutionalization, Deinstitutionalization and the Adversary Process*, 75 COLUM. L. REV. 897 (1975).

[14]For accounts of the problems under various commitment laws, see, e.g., CALIFORNIA DEPARTMENT OF MENTAL HEALTH, AFTER A DECADE OF L-P-S—UNCERTAIN TIMES IN MENTAL HEALTH LAW (B. Chell ed. 1981); Landau, *Civil Commitment in Wisconsin: The Impact of* Lessard v. Schmidt, 1976 WIS. L. REV. 503; Miller & Fiddleman, *Involuntary Civil Commitment in North Carolina: The Result of the 1979 Statutory Changes*, 60 N.C.L. REV. 985 (1982); Rabin & Folks, *Dangerousness as the Criterion for Involuntary Hospitalization: A Time to Reassess*, 246 J. A.M.A. 990 (1981); Stier & Stoebe, *Involuntary Hospitalization of the Mentally Ill in Iowa: The Failure of the 1975 Legislation*, 64 IOWA L. REV. 1284 (1979).

between the proponents of medical discretion and the legal activists who see civil liberties as the ultimate good. In this context, the American Psychiatric Association (APA)[15] began to develop its own proposal for a humane state civil commitment law. In 1982, after an extensive drafting, consultation, and review process, the APA approved the Model Law that we present and interpret in this Article.[16]

The Model Law thus represents the views of our nation's psychiatrists, though like any document reached through both principled and political compromise, it may not be wholly satisfactory to any individual or group. Individuals and local associations of psychiatrists will, of course, differ on their views as to particular provisions. The Model Law will doubtless draw both support and criticism from other professions and the public. But it should not be dismissed as an effort simply to advance the narrow interests of psychiatrists. It does not seek to return complete discretion to psychiatrists or to dispense with countervailing legal safeguards. We hope that readers will agree that the Model Law represents a serious effort to come to terms with the competing interests that civil commitment must serve. Before examining each provision in detail, we discuss the reasoning that underlies the Model Law.

During the past century, states involuntarily confined mentally disordered persons to serve three social functions: protection of society ("police power"), psychiatric treatment for the patient's "own good" ("parens patriae"),[17] and custodial confine-

[15]The APA was founded in 1844 and is the nation's largest organization of doctors of medicine specializing in psychiatry. Its 26,000 members comprise more than three quarters of our nation's psychiatrists.

[16]Mr. Stromberg served as the APA's legal adviser throughout the process and was the chief draftsman of the Model Law. Dr. Stone was President of the APA when the Model Law was first commissioned and later chaired the APA's Council on Governmental Policy and Law that reviewed successive drafts.

APA District Branches and Area Councils around the country extensively reviewed the Model Law, which thus benefited from hundreds of psychiatrists' contributions. Although these people are too numerous to mention, the authors would like to acknowledge the work of at least the following: the members of the APA Council on Governmental Policy and Law (Richard Bonnie, Esq., Dr. Thomas Pfaehler, Dr. Shirley Phelps, Dr. Richard Rada, Dr. David Starrett, Dr. Laurence Tancredi, Dr. Hugo Van Dooren, and Dr. Andrew Watson), and the members of the APA Assembly Task Force on the Model Law (Dr. Jerome Beigler, Dr. Thomas Callahan, Dr. Merlin Johnson, Dr. Alfred Margulies, Dr. Ralph O'Connell, Dr. Pete Palasota, Dr. Roger Peele, and Dr. Thomas Pfaehler). Dr. Callahan also served ably as coordinator for the Assembly's review of the Model Law. Finally, a special acknowledgement is due Dr. Pfaehler, who orchestrated the APA's review of the Model Law. He skillfully resolved disputes and made his own substantial contributions to the document.

[17]Parens patriae is the power of the sovereign to act as "the general guardian of all infants, idiots and lunatics." Hawaii v. Standard Oil Co., 405 U.S. 251, 257 (1972)

ment (that is, simply caring for a patient's bodily needs). In recent years, the legal system, being concerned with the trade-off of individual liberty for social protection, often has acted as if the police power function of civil commitment were the only one. Psychiatrists and other health professionals, being concerned primarily with health care, instead have focused on the parens patriae function. Custodial care is no longer advanced itself as sufficient justification for commitment. Yet, ironically, during most of the last hundred years, custodial confinement has been the principal function actually served by civil commitment.[18] This fact has discredited psychiatry and the whole mental health system. The result has been a confused system in which none of the purposes of commitment—security, treatment, or custody—is adequately served.

The Model Law seeks to move beyond the narrow debate over the lawyers' versus the psychiatrists' conception of civil commitment. The Model Law seeks to put the *patients'* interests first. It therefore rejects the idea that dangerousness is the only valid basis for commitment, and instead makes the provision of treatment the indispensable element justifying commitment. Under the Model Law, *parens patriae commitment* requires that the patient suffers from a severe mental disorder, lacks "capacity to make a reasoned decision concerning treatment," is treatable, and is likely to harm himself or others. Significantly, with respect to the likelihood of harm, the Model Law expands current parens patriae commitment to include persons "likely to suffer substantial mental or physical deterioration," as well as those who are dangerous to themselves or unable to care for

(quoting 3 W. BLACKSTONE, COMMENTARIES *47); *see also* Johnson v. State, 18 N.J. 422, 114 A.2d 1 (1955); Curtis, *The Checkered Career of Parens Patriae: The State as Parent or Tyrant?*, 25 DE PAUL L. REV. 895 (1976); Custer, *The Origins of The Doctrine of Parens Patriae*, 27 EMORY L.J. 195 (1978).

[18]Even now, the question of mere custodial confinement unavoidably presents itself. Psychiatrists caring for chronically mentally ill patients often face a cruel choice between discharging patients even though no adequate alternative care facilities exist, and retaining in hospitals patients who no longer need such special, intensive treatment. These patients may meet the Supreme Court's *Donaldson* test—they can survive in freedom—but perhaps only barely so and for an uncertain period of time. *See Donaldson*, 422 U.S. at 576. It has been suggested that patients be allowed to "vote with their feet," that is, to stay or to leave the hospital as they wish. While this would resolve the psychiatrist's difficult choice, it would ignore the realities that confront the patient. Judicial attempts to address this problem have produced very limited success. *See, e.g.*, Dixon v. Weinberger, 405 F. Supp. 974 (D.D.C. 1975); Dixon v. Harris, C.A. No. 74-285, Sup. Ct. (D.D.C. 1980). Most jurisdictions have not provided the resources necessary to make community care a humane alternative to institutionalization for patients with recurring, serious impairment. Concern about this class of chronic patients was central to the debate over the Model Law.

their needs. Thus, it would permit commitment of many of those severely mentally ill people who, ignored by current commitment laws and abandoned by the mental health system, now roam the streets aimlessly and without hope.

For *police power commitment*, the Model Law also requires a severe mental disorder, clear and convincing proof of likely dangerous conduct, lack of capacity, and evidence that effective treatment will be provided. Many dangerous people cannot be treated effectively, so that present statutes confining them to mental hospitals merely achieve preventive detention under a therapeutic guise. This is both bad law and bad medicine. The Model Law allows emergency psychiatric intervention to evaluate such people, but does not allow long-term custodial confinement without treatment. The mental health system cannot solve all of society's problems arising from all forms of social deviance. If a violent person is not severely mentally ill or if treatment will not help him, the person should be dealt with by the criminal justice system or by other social institutions.

Proponents of diverting people from the criminal justice system argue that many people are better off in the mental health system. They regard the mental health system as a resource for solving criminal justice problems. In most cases, however, mental hospitals are not secure enough to protect society from dangerous persons without being restructured so drastically that psychiatry's therapeutic goals would be forsaken and all patients would find themselves in "hospitals" that strongly resemble jails. In addition, treatment constitutes the sole justification for a civil commitment process distinct from the criminal justice system. Absent that justification, confinement merely punishes or sequesters the patient. If society wishes to permit confinement on such a basis, it should reach that decision deliberately and candidly, and not obscure its decision with quasi-medical window dressing.

In sum, the Model Law expands the reach of the mental health system somewhat on the side of treating those who desperately need treatment, and contracts its reach on the side of those whom it can neither treat nor effectively sequester.

The Model Law also addresses the critical issues of a patient's "right to treatment" and "right to refuse treatment." These rights have been the subject of considerable constitutional litigation in recent years. Several courts have declared that involuntary patients have a right to refuse treatment because the laws under

which they were originally committed failed to require proof that they lacked the capacity to make an informed decision concerning treatment.[19] Under these decisions, when a patient exercises this right, the cruelest paradox in civil commitment results: a mentally ill patient is deprived of his liberty but cannot be treated.[20] This stalemate converts suffering mental patients into warehoused items, converts psychiatrists into jailors, and converts civil commitment into a charade.

The Model Law addresses this problem by requiring the court to find that the person to be involuntarily committed not only is severely mentally ill and treatable, but also "lacks the capacity to make an informed decision concerning treatment." Such an involuntary patient thus can be treated in accordance with law and good medical practice without his consent. Voluntary patients, as to whom there has been no adjudication of incompetency, retain the "right to refuse treatment," but may be discharged if they do so. The Model Law also ensures that every patient who is committed, whether voluntarily or involuntarily, has a "right to treatment" and a right to a "healthful and humane environment."

It has been argued that the past failings of civil commitment were largely attributable to use of a "medical model" in commitment rather than a "legal model."[21] The "medical model" conceives of commitment as a process for obtaining treatment for persons whom psychiatrists diagnose as mentally ill. Under this model, legal procedures should be secondary to therapeutic concerns. The "legal model" conceives of commitment as a deprivation of liberty in order to protect society from dangerous persons; accordingly, all of the criminal justice legal procedures should apply.

Some might argue that the Model Law unwisely resuscitates the medical model. Such a criticism would be amiss. In fact,

---

[19]*See* Rogers v. Okin, 478 F. Supp. 1342, 1364–65 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982); Rennie v. Klein, 462 F. Supp. 1131 (D.N.J. 1978), *supplemented*, 476 F. Supp. 1249 (D.N.J. 1979), *aff'd and modified*, 653 F.2d 836 (3d Cir. 1981), *vacated and remanded*, 102 S. Ct. 3506 (1982).

[20]*See* Appelbaum & Gutheil, *The Boston State Hospital Case: "Involuntary Mind Control," the Constitution, and the Right To Rot*, 137 AM. J. PSYCHIATRY 720 (1980); Treffert, *The Practical Limits of Patients' Rights*, in PSYCHIATRISTS AND THE LEGAL PROCESS: DIAGNOSIS AND DEBATE 227 (R. Bonnie ed. 1977).

[21]*See generally* D. MECHANIC, MENTAL HEALTH AND SOCIAL POLICY (2d ed. 1980); Hardisty, *Mental Illness: A Legal Fiction*, 48 WASH. U.L. REV. 735 (1973); Morse, *A Preference for Liberty: The Case Against Involuntary Commitment of the Mentally Disordered*, 70 CAL. L. REV. 54 (1982).

the Model Law wholly accepts neither the medical model nor the legal model, but draws from each. In our view, past failures have not been simply due to the use of a medical model, but to a lack of legal accountability in the civil commitment process. Confining people for security where there is no security, and for treatment without offering treatment, is a defect in practice, not in theoretical models.[22] Remedying this defect requires not only clarifying the basis for commitment, but also enforcing legal rights and providing adequate funds to ensure that the conditions of confinement accord with its theoretical purposes.

Unless one is willing to move toward a system of brazen preventive detention of anyone who is likely to cause harm to others, one must employ some medical concept of "mental disorder" as the threshold criterion for commitment. The Model Law tries to make this concept a clear, legally enforceable criterion, rather than a protean bit of jargon. It also requires proof that the person to be committed lacks capacity and is treatable. In addition, it guarantees a patient several legally enforceable rights. Advising patients, treatment staff, and the public that these rights must be respected helps ensure that civil commitment will actually advance therapeutic and protective goals.

Thus, the Model Law adopts elements from both the legal and the medical models. The proposal also requires specifying at the outset the precise purposes to be served by each commitment. Respondents in civil commitment should know why commitment is sought. Treatment staff need a clear sense of what they are trying to accomplish and the ability to communicate it to each patient. The Model Law therefore forces courts to state precisely the statutory basis for each commitment. It also contains different procedures for releasing persons committed under parens patriae and those committed under the police power. Because the decision to release a person committed on the basis that he was unable to care for his needs is essentially a psychiatric judgment, the psychiatrist may release such a patient directly. The decision to release a patient committed as dangerous, however, involves a balancing of broader social interests. The Model Law thus requires the treatment facility to apply to the court to authorize release of such a patient.

---

[22] *See* Stone, *Psychiatric Abuse and Legal Reforms: Two Ways To Make a Bad Situation Worse*, 5 INT'L J.L. & PSYCHIATRY 9 (1982).

Finally, the Model Law prescribes how long involuntary commitment should last. Modern psychiatric treatment at its best is geared to intensive, short-term therapy, not to long-term hospital care. Empirical data now indicate that most psychiatric inpatient cases can be resolved sufficiently to release the patient within thirty to ninety days.[23] Thus, the Model Law rejects open-ended commitment. It provides instead for initial emergency evaluation lasting up to fourteen days to be followed, if necessary, by a sequence of commitment periods lasting thirty and sixty days, which must be authorized in separate hearings. If at the end of this period a patient requires further treatment, a court may authorize repeated involuntary commitments lasting 180 days each.[24]

Dangerous persons present special problems. Most current commitment laws reflect the belief that psychiatrists can accurately predict whether particular mentally ill individuals will commit dangerous acts. But the empirical evidence demonstrates that psychiatrists cannot make such predictions accurately.[25] However inconvenient this is for the law, the prediction of violence will remain a rudimentary craft for the foreseeable future. Although at the time of commitment the patient's recent behavior may improve the accuracy of predictions about the need for hospitalization to avoid violence, at the time of discharge any such behavior is often too remote to be very meaningful. Currently, therefore, treatment staffs and the courts engage in the charade of justifying further periods of confinement

---

[23]*See, e.g.*, Gove & Fain, *A Comparison of Voluntary and Committed Mental Patients*, 34 ARCHIVES GEN. PSYCHIATRY 669, 673 (1977) (67% of involuntary patients discharged within 38 days); Tomellieri, Lakshminanayanan & Herjanic, *Who Are the Committed*, 165 J. NERVOUS & MENTAL DISEASE 288, 291 (1977) (63% of committed patients discharged within 90 days).

[24]Although the Model Law's final version does not reflect this, the authors believe that involuntary confinement for nondangerous persons should cease after emergency (14-day), 30-day, and 60-day commitment, i.e., after about 104 days. If a patient who is not dangerous to others has not benefited from treatment within that period and continues to object to hospitalization, we do not see the ethical or social mandate that justifies indefinite confinement. The provisions finally approved by the APA reflect a reluctance by psychiatrists to repudiate entirely custodial care, because few adequate alternatives now exist. Understandably, psychiatrists felt that discharging a possibly self-destructive patient would run counter to their ethical responsibilities of good care. Nevertheless, the authors of this Article would prefer to set a bright-line limit to parens patriae confinement of about 104 days. If the Constitution and state laws were held to permit mere custodial confinement, of which we have some doubts, we prefer a guardianship procedure that makes it explicit that psychiatric treatment is not the justification for loss of liberty, and that appropriately constrains the guardian's decisionmaking in ways that do not now exist in civil commitment.

[25]*See, e.g.*, J. MONAHAN, THE CLINICAL PREDICTION OF VIOLENT BEHAVIOR (1981); A. STONE (WITH C. STROMBERG) *supra* note 2, at 25–40.

for allegedly dangerous persons on the basis of guesswork predictions of violence and illusory promises of treatment.

Unlike most current laws, the APA proposal requires treatment staff to seek to discharge or to transfer a dangerous patient if no further substantial treatment purpose would be served by confinement. Because therapy constitutes the mental health system's primary function, police officers and courts should not use psychiatric facilities for preventive detention. Prosecutors and courts will be informed, however, before any person previously thought to be dangerous is released. If appropriate, criminal charges should be lodged and the criminal process should be activated before releasing such a person from the hospital. This will force the criminal justice system to take responsibility for such people and to develop decent, secure facilities where they can be handled. In deciding whether or not to discharge violent patients, courts will have to decide whether future confinement can be justified solely for treatment purposes, or whether criminal processes should be put in effect. Although some states may wish to add legislation permitting mere custodial care for long-term patients in secure facilities providing some psychiatric treatment, the Model Law does not sanction such efforts. Instead, the proposed legislation will remind treatment staffs that they are responsible for providing treatment, and patients will no longer be deceived as to why they are being confined.

We have provided here only a bare outline of the reasoning that went into the development of the Model Law. When the enterprise began, some members of the APA candidly expressed skepticism about the value of even trying to draft a model statute. Many psychiatrists believe that no matter how particular statutes are worded, judges and juries do as they wish and civil commitment remains a legal arabesque with only faint attention to psychiatric realities or to a patient's real interests. Many reformist lawyers believe that psychiatrists and judges disregard even the most precise statutes and routinely commit people based on their own biases and preconceptions.

After reviewing each state's civil commitment law, and the history of mental health legislation, litigation, and policy over the past two decades, we came to the opposite conclusion. In our view, the concepts and the precise language in state commitment laws matter a great deal. Often state laws define key terms such as "mentally ill," or "dangerous" so vaguely as to invite abuse in practice and invalidation in court. Patients' rights

are sometimes described so broadly as to be unenforceable, impossible to fulfill, or harmful to other patients' rights. The basis for continued commitment often is unspecified or confusing. Many current laws impose duties on treatment staff without fairly limiting their growing legal liability.

No statute alone could solve the medical, fiscal, and managerial problems of the mental health system. But a bad statute can mire the system in endless skirmishes, while a better statute can help the system to achieve at least a partial victory.

Thus, whatever contribution the Model Law may make to the theoretical debate, we offer it fundamentally as a practical effort to address practical problems. The subject of civil commitment has stimulated more than its share of philosophical debate. We, and the APA, feel that at some point one must cease cursing the errors of others and offer up one's own version of imperfect wisdom.[26] In this spirit, we submit the Model Law for consideration by legislators, lawyers, mental health professionals—and the public whom the civil commitment process serves.

The remainder of this article sets forth the Model Law approved by the APA, as well as the authors' commentary explaining each provision.[27]

MODEL STATE LAW ON CIVIL COMMITMENT OF THE MENTALLY ILL

TABLE OF CONTENTS

[26]For other attempts to design a model statute on some of these subjects, see, e.g., Public Health Service, *A Draft Act Governing Hospitalization of the Mentally Ill*, in THE MENTALLY DISABLED AND THE LAW (S. Brakel & R. Rock, revised ed. 1971); UNIF. HEALTH CARE CONSENT ACT (1982); Gauvey, Leviton, Shuger & Sykes, *Informed Consent and Substitute Consent to Health Care Procedures: A Proposal for State Legislation*, 15 HARV. J. ON LEGIS. 431 (1978); Mental Health Law Project, *Legal Issues in State Mental Health Care: Proposals for Change—Civil Commitment*, 3 MENTAL DISABILITY L. REP. 75 (1979).

[27]It must be stressed that the APA has approved only the text of the Model Law and not the authors' section-by-section commentary.

The APA did approve, in addition to the full Model Law, an abbreviated alternative. The APA recognized that some states might not wish to undertake a comprehensive revision of their civil commitment laws. They might, however, wish to adopt some of the key elements of the Model Law. For such states, the APA recommended adoption of the provision for commitment of a person likely to "suffer substantial mental or physical deterioration"; the requirement of an independent treatment facility determination prior to admission; the criteria for civil commitment set forth in subsection 6.C; and the provision in subsection 8.B that, because it is a prerequisite to commitment that the person lack capacity to make an informed decision concerning treatment, the consent to treatment of involuntary patients is not required.

Section 1. SHORT TITLE.

These provisions governing the psychiatric hospitalization of adults may be cited as Title I of the Mental Health Code.

COMMENTARY: The Model Law governs only the psychiatric hospitalization (civil commitment) of adults; "person" is defined as someone aged eighteen years or more. It does not address the special concerns that arise in the hospitalization of minors. The Model Law deals only with civil commitment; it does not deal with persons confined for forensic evaluation or other purposes in the criminal justice system. The Model Law is intended to be only one part of a state's mental health code. Thus it does not deal with many related subjects, such as administration of a mental health department, licensing of health professionals, or financing of facilities.

Section 2. LEGISLATIVE PURPOSES.

This Act is intended to achieve and shall be construed so as to promote these legislative purposes:

— To make psychiatric evaluation, care, and treatment available to persons who suffer from severe mental disorders and can benefit from treatment, and to encourage voluntary rather than involuntary admission whenever hospitalization is necessary;

— To safeguard the legal rights of patients in a manner that will advance and not impede the therapeutic and protective purposes of psychiatric hospitalization;

— To provide workable procedures for obtaining consent to and administering medications and other treatments;

— To provide legal immunity for reasonable, good-faith efforts to implement this Act, and legal penalties for knowing, willful efforts to subvert this Act; and

— To provide a statutory framework for the promulgation of regulations by the Department of Mental Health.

COMMENTARY: No comment is required.

Section 3. DEFINITIONS.

As used in this Act, the terms below shall have the meanings indicated:

COMMENTARY: Definitions used in the Model Law serve functional purposes specific to it. In other contexts, the same terms may acquire different meanings. This is particularly true of the clinical terms such as "aversive therapy" and "experimental treatment."

> **"*aversive therapy*" means any treatment or procedure that, because it is believed to be painful or physically uncomfortable to the patient, is administered in order to reduce the frequency or intensity of a behavior; except that aversive therapy does not refer to verbal therapies, seclusion or physical restraints used in conformity with subsection 10.F., or psychotropic medications which are not used for purposes of aversive conditioning.**

COMMENTARY: The Model Law defines "aversive therapy" because, as provided in subsection 8.C, such treatment may be administered only under strict regulations to be developed by each state's mental health department. The definition turns on the purpose for which the physician administers the therapy, not the patient's subjective reaction to it. The term applies to a therapy only if the physician employs it because he expects that it will be uncomfortable to the patient and that the discomfort will reduce the frequency or intensity of the patient's harmful behavior.[28]

The definition excludes purely verbal therapies because they are so nuanced and interactive that therapists often cannot predict the degree of discomfort they may cause a patient. Also, unlike aversive therapies, which use physical stimuli, verbal therapies have not been the subject of widespread public concern. Aversive therapy also does not include properly administered seclusion or physical restraints, because the therapeutic rationales for their use differ from those for aversive therapy.[29] Any use of seclusion or restraints not in conformity with the law or regulations required in subsection 8.C might, however, be deemed aversive therapy.

Finally, the definition of aversive therapy excludes psycho-

---

[28]*Cf.* N.M. STAT. ANN. § 43-1-3.A (1979). *See generally* APA, TASK FORCE REP. NO. 5, BEHAVIOR THERAPY IN PSYCHIATRY (1973); Goldiamond, *Singling Out Behavior Modification for Legal Regulation: Some Effects on Patient Care, Psychotherapy, and Research in General*, 17 ARIZ. L. REV. 105 (1975).

[29]*See* K. Tardiff, T. Gutheil, R. Liberman, J. Lion, P. Soloff, D. Gair, M. Mattson & D. Wexler, Progress Report from the Task Force on Seclusion and Restraint Submitted to the Council on Governmental Policy and Law of the American Psychiatric Association 47 (Sept. 24, 1982) (unpublished manuscript on file at HARV. J. ON LEGIS.) [hereinafter cited as Tardiff Progress Report].

tropic medications not used specifically for operant conditioning. Although a patient may on occasion find such medications unpleasant, when appropriately prescribed such medications are well tolerated, and the psychiatric justifications for their use are wholly unrelated to any aversive effect.[30]

> **"*consistent with the least restrictive alternative principle*" means that (1) each patient committed solely on the ground that he is likely to cause harm to himself or to suffer substantial mental or physical deterioration shall be placed in the most appropriate and therapeutic available setting, that is, a setting in which treatment provides the patient with a realistic opportunity to improve, and which is no more restrictive of his physical or social liberties than is believed conducive to the most effective treatment for the patient; and (2) each patient committed solely or in part on the ground that he is likely to cause harm to others shall be placed in a setting in which treatment is available and the risks of physical injury or property damage posed by such placement are warranted by the proposed plan of treatment.**

COMMENTARY: The Model Law ensures a patient the right to treatment "consistent with the least restrictive alternative principle," but defines that principle differently than it has sometimes been understood. The doctrine that mental patients have the right to be placed in the "least restrictive alternative" facility acquired considerable judicial acceptance during the early 1970's as the trend from "institutions" to "community placements" grew. Unfortunately, while the term acquired talismanic legal significance, its content and practical impact remained uncertain.

Originally, the concept of the "least restrictive alternative" meant simply that the government could not pursue its purposes "by means that broadly stifle fundamental personal liberties where [there are] less drastic means for achieving the same basic purpose."[31] Civil commitment was conceived as a system in which the government's purposes (treatment and prevention of harm) could be achieved in a hierarchical array of less-to-more-restrictive facilities.[32] This model ignored the fact that patients

---

[30]*See* R. BALDESSARINI, CHEMOTHERAPY IN PSYCHIATRY (1977); GROUP FOR THE ADVANCEMENT OF PSYCHIATRY, PHARMACOTHERAPY & PSYCHOTHERAPY, PARADOXES, PROBLEMS AND PROGRESS (1975); D. KLEIN & J. DAVIS, DIAGNOSIS AND DRUG TREATMENT OF PSYCHIATRIC DISORDERS: ADULTS AND CHILDREN (1980).

[31]Shelton v. Tucker, 364 U.S. 479, 488 (1960).

[32]Hoffman & Foust, *Least Restrictive Treatment of the Mentally Ill: A Doctrine in Search of its Senses*, 14 SAN DIEGO L. REV. 1100 (1977).

have interests not only in being unrestricted, but also in getting well.

For example, it may be less constraining for a patient to participate for thirty days in an effective therapeutic program in a "restrictive" institution than to be treated for a year in a less effective "community" program. The reverse also may be true. The point is that there is probably "no relationship between restrictiveness and [treatment] effectiveness, let alone a stable or singular one."[33] What is or is not in the patient's interests cannot be measured on a simplistic scale of restrictiveness.

When interpreting the "least restrictive alternative" doctrine, courts and legislatures generally recognize that it requires an accommodation of various concerns and does not imply an absolute right.[34] Confusion about its meaning, however, remains. In some cases, the doctrine has been construed to apply not only to the overall conditions of confinement, but to each treatment and therapy separately, even though this is impractical.[35]

---

[33]*Id.* at 1103–04.

[34]*See, e.g.*, Pennhurst v. Halderman, 451 U.S. 1 (1981); Covington v. Harris, 419 F.2d 617, 623 (D.C. Cir. 1969); Naughton v. Bevilacqua, 458 F. Supp. 610, 617 (D.R.I. 1978), *aff'd*, 605 F.2d 586 (1st Cir. 1979); Eubanks v. Clarke, 434 F. Supp. 1022, 1028 (E.D. Pa. 1977); Stamus v. Leonhardt, 414 F. Supp. 439, 446 (S.D. Iowa 1975); Dixon v. Weinberger, 405 F. Supp. 974 (D.D.C. 1975); Lynch v. Baxley, 386 F. Supp. 378, 392 (M.D. Ala. 1974); Davis v. Watkins, 384 F. Supp. 1196, 1203 (N.D. Ohio 1974); Welsch v. Likins, 373 F. Supp. 487, 502 (D. Minn. 1974) (decisionmakers must "make good faith attempts to place such persons in settings that will be suitable and appropriate to their mental and physical conditions while least restrictive of their liberties"), *aff'd in part and vacated in part*, 550 F.2d 1122 (8th Cir. 1977); Lessard v. Schmidt, 349 F. Supp. 1078 (E.D. Wis. 1972), *vacated on other grounds*, 414 U.S. 473 (1973); State v. Sanchez, 80 N.M. 438, 457 P.2d 370 (1969), *appeal dismissed*, 396 U.S. 276 (1970); Kesselbrenner v. Anonymous, 33 N.Y.2d 161, 165, 305 N.E.2d 903, 905, 350 N.Y.S.2d 889, 902 (1973); *see also* N.J. STAT. ANN. § 30:4-24.2(e)(2) (West 1981 & Supp. 1982–1983) (right to "the least restrictive conditions necessary to achieve the purposes of treatment"); PA. STAT. ANN. tit. 50, § 7302 (Purdon 1978) ("the least restrictions consistent with adequate treatment shall be employed").

[35]*See, e.g.*, Rennie v. Klein, 462 F. Supp. 1131 (D.N.J. 1978), *supplemented*, 476 F. Supp. 1294 (D.N.J. 1979), *aff'd and modified*, 653 F.2d 836 (3d Cir. 1981), *vacated and remanded*, 102 S. Ct. 3506 (1982); CAL. WELF. & INST. CODE § 5325.1(a) (West Supp. 1980); N.M. STAT. ANN. § 43-1-3(D) (1979). In our view, applying the "least restrictive alternative" doctrine to individual treatments would be a legal absurdity and a therapeutic nightmare.

> [T]he reactions to such treatments are highly individualized and vary markedly from patient to patient. It would not only be impractical for a court to attempt to review these treatment decisions under a "least restrictive alternative" approach, but it seems inevitable that such a review would interfere dramatically with the administration of needed medication and essential therapy. [It would] require each district court judge to exchange his robe for a medical gown, . . . becoming both a super-diagnostician and a super-physician for each institutionalized patient.

Rennie v. Klein, 653 F.2d at 859 (Garth, J., concurring in part and dissenting in part); *see also* Youngberg v. Romeo, 102 S. Ct. 2452, 2462 (1982).

Some lawyers interpret the court decisions to require community placements in virtually all cases,[36] even though commitment often serves protective as well as therapeutic purposes. Moreover, no decision has yet made clear what scale should be used to measure "restrictiveness."

The Model Law recognizes that the "idea of deinstitutionalization" is realistic only if the "community can provide the full range of patient services outside of institutional settings that are available inside the hospital."[37] Unfortunately, in many cases, communities cannot provide the necessary services, or community placements simply are not available.[38] Excessive deinstitutionalization based on the theory that it is always less restrictive too often has led to abandonment of patients to the inhospitable world of "welfare hotels" and alleys.[39]

The Model Law's definition of "consistent with the least restrictive alternative principle" reflects the various purposes civil commitment serves for different populations. A patient committed solely on a parens patriae basis (for his own good) must be placed in "the most appropriate and therapeutic available setting" that restricts his liberty no more "than is believed conducive to the most effective treatment for the patient." For such persons, treatment is the key goal. Restrictions on patient liberty should not be imposed unless they help to ensure effective treatment.

A patient committed on a police power basis (to protect society) must be placed where society will be adequately protected, and where any reduction in restrictiveness is "warranted by the proposed plan of treatment." Such patients must be

---

[36]*See* Brief for Respondents Terri Lee Halderman, et al., Halderman v. Pennhurst State School & Hosp., 612 F.2d 84 (3d Cir. 1979), *rev'd*, 451 U.S. 1 (1981) (on petitions for writ of certiorari); Chambers, *Alternatives to Civil Commitment of the Mentally Ill: Practical Guides and Constitutional Imperatives*, 70 MICH. L. REV. 1108, 1120 (1972).

[37]L. BACHRACH, DEINSTITUTIONALIZATION: AN ANALYTICAL REVIEW AND SOCIOLOGICAL PERSPECTIVE (1979); *see also* Braun, Kochansky, Shapiro, Greenberg, Gudeman, Johnson & Shore, *Overview: Deinstitutionalization of Psychiatric Patients, A Critical Review of Outcome Studies*, 138 AM. J. PSYCHIATRY 736 (1981).

[38]*See In re* Steinheiser, 424 A.2d 1006 (Pa. Commw. Ct. 1981); Hoffman & Foust, *supra* note 32; Note, *The Due Process of Community Treatment of the Mentally Ill: A Case Study*, 59 TEX. L. REV. 1481, 1502 (1981) [hereinafter cited as Note, *Community Treatment Case Study*].

[39]*See, e.g., Hearings on the Mental Health Systems Act Before the Subcomm. on Health and Scientific Research of the Senate Comm. on Labor and Human Resources*, 96th Cong., 1st Sess. 64-65 (1979) (remarks of Joseph A. Califano, Secretary, DHEW); Holden, *The Plight of the "Deinstitutionalized" Mental Patient*, 200 SCI. 1366 (1978); Talbott, *Deinstitutionalization: Avoiding the Disasters of the Past*, 30 HOSP. & COMMUNITY PSYCHIATRY 621 (1979).

afforded treatment, but not necessarily in "*the most* . . . therapeutic available setting," because security is also a concern. Thus, placement and treatment "consistent with the least restrictive alternative principle" is not an absolute right, but rather one part of the overall therapeutic and protective decision concerning hospitalization.[40]

> **"*court*" means the court or judicial officer designated under the laws of this State for the discharge of the functions described in this Act.**

COMMENTARY: Civil commitment involves important interests of both the state and private citizens. Commitment decisions therefore should be made by judicial rather than administrative officials. Whenever possible, these decisions should be made by a judge. If necessary, however, states may empower other judicial officers, such as magistrates.[41]

> **"*emergency situation*" means a situation in which the patient exhibits substantial behavior that is self-destructive or assaultive, threatens significant damage to the property of others, or indicates that the patient is suffering extreme anxiety amounting to panic or sudden exacerbation of his severe mental disorder.**

COMMENTARY: Under Section 7 of the Model Law, voluntary patients have the right to consent or to refuse to consent to any treatment, except in "emergency situations." These include situations in which the patient is physically self-abusive, assaults or threatens imminently to assault others, seriously panics, or otherwise undergoes a sudden and serious adverse change in his mental condition (imminent decompensation).

Several recent cases have attempted to define "emergency" in order to limit a patient's right to refuse treatment. They generally have recognized as an emergency only the likelihood

---

[40]A patient's right is not violated simply because he regards some other facility as "less restrictive," but only if the placement decision is a "substantial departure from accepted professional judgment." Rennie v. Klein, 653 F.2d 836, 855 (3d Cir. 1981), (Seitz and Aldisert, JJ., concurring) (quoting Romeo v. Youngberg, 644 F.2d 147 (3d Cir. 1980) (en banc)), *vacated and remanded*, 102 S. Ct. 3506 (1982). The Supreme Court, although vacating and remanding the Third Circuit's decision in *Romeo*, subscribed to this standard as applied to the use of restraints. *Romeo*, 102 S. Ct. 2452 (1982).

[41]*See* N.C. GEN. STAT. § 122-58.3 (1981) (magistrate or court clerk is authorized to perform certain key functions, such as ordering emergency evaluation); UTAH CODE ANN. § 64-7-35 (1981) (court may appoint any attorney as "mental health commissioner" who makes recommendations on commitment to the court). *But see* Dixon v. Commonwealth, 325 F. Supp. 966 (M.D. Pa. 1971) (striking down physician-controlled commitments).

of physical harm, and have ignored "psychiatric emergencies" arising from the clinical state of the patient.[42] The Model Law's definition, which recognizes such emergencies, is more flexible and realistic in light of the patient's conflicted state. Although one federal court has held that voluntary patients who refuse treatment should simply be discharged,[43] this response would be inappropriate for a severely disordered patient. Because a voluntary patient has consented to admission for treatment, the Model Law deems his consent sufficient for emergency treatment until the acute episode subsides. It would betray the therapeutic purposes of voluntary commitment, and needlessly endanger both patients and staff, if even emergency treatment were withheld upon the objection of an agitated, delusional patient. The Model Law, however, does not permit a treatment facility to vitiate a voluntary patient's general right to refuse treatment by routinely and inappropriately characterizing events as "emergency situations."

> **"*experimental treatment*" means any treatment other than one that is commonly accepted for treatment of the mental disorder involved (or is supported by widely accepted scientific studies) and is provided by a qualified health professional, if such treatment poses a significant risk of harm to the patient.**

COMMENTARY: Under subsection 8.C, the Model Law permits experimental treatments only under mental health department regulations. These regulations should include requirements such as informed consent of the patient, independent peer review of research protocols, assessment of the treatment's risk-to-benefit ratio, and adequate follow-up evaluation. In addition, subsection 9.C.2 requires that experimental treatments be administered in accordance with other applicable laws, such as federal regulations on research involving human subjects,[44] and any state laws governing experimental treatments.[45]

---

[42]*See, e.g.*, Rogers v. Okin, 478 F. Supp. 1342, 1364–65 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982); Rennie v. Klein, 462 F. Supp. 1131 (D.N.J. 1978), *supplemented*, 476 F. Supp. 1294 (D.N.J. 1979), *modified and remanded*, 653 F.2d 836 (3d Cir. 1981), *vacated and remanded*, 102 S. Ct. 3506 (1982). For general medical and legal definitions of "emergency," see 1977 ARK. ACTS 805 § 2(a); CAL. WELF. & INST. CODE § 5358 (West Supp. 1980); GA. CODE § 31-9-3 (1982); IDAHO CODE § 39-141(A) (1976); MISS. CODE ANN. § 41-41-7 (1966); 70 C.J.S. *Physicians & Surgeons* § 48 (1951).

[43]Rogers v. Okin, 478 F. Supp 1342, 1368 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982).

[44]*See, e.g.*, 45 C.F.R. § 46 (1982).

[45]A number of state laws regulate "experimental treatments." *See, e.g.*, DEL. CODE

The Model Law regulates experimental treatment only when it poses a "significant risk of harm" to the patient. If there is no real risk of harm (including the foregone opportunity for a clearly better therapy), then there is no need for restrictions beyond the usual requirements of good medical practice. For example, if a physician believes that playing classical music helps catatonic patients to come out of their withdrawal, this should not be subject to special legal regulation. If, however, he wishes to try a novel drug therapy that might have significant adverse side effects, special regulation is proper.

A treatment is deemed experimental if (1) it is not "commonly accepted" for the proposed use, and (2) it is not supported by "widely accepted" scientific studies. Under the first part of this standard, use of the drug thorazine to treat schizophrenia (which is commonly accepted) would not be experimental, while its use to treat alcohol abuse (which is not accepted) would be.[46] The second part of the standard recognizes that research published in respected journals may at some point justify employing a new therapy that has not yet become "commonly accepted." Were this not permitted, medical advances would be impossible. Finally, a therapy may be deemed experimental—even if its use per se is commonly accepted or supported by research—if it is administered by a type of professional who does not ordinarily administer it. Although such a procedure might be medically justified, it would be experimental if it posed a significant risk of harm to the patient.

> ***"informed consent to treatment"*** **means a knowing and voluntary decision to undergo treatment, evidenced in writing, and made by a person who has the capacity to make an informed decision, after staff of the treatment facility have explained to the person the nature and effects of the proposed treatment.**

COMMENTARY: The Model Law's provisions in subsection 7.A and Section 8 would require a psychiatrist to obtain his patient's

---

ANN. tit. 16, § 5161(a)(2)(d) (1974); ME. REV. STAT. ANN. tit. 34, § 2143.8(E) (1977); N.C. GEN. STAT. § 122-55.6 (1981); S.C. CODE ANN. § 44-23-1010 (Law. Co-op. 1974); S.D. CODIFIED LAWS ANN. § 27A-12-20 (1975).

[46]Some psychiatric researchers may consider the sharp distinction between "commonly accepted" and "experimental" somewhat artificial in light of the frequent gradual application of therapies to new medical problems. For example, recent research has suggested that drugs previously used only for depressive disorders also may be effective for certain phobias. *See* Sheehan, Ballenger & Jacobsen, *Treatment of Endogenous Anxiety With Phobic, Hysterical, and Hypochondriacal Symptoms*, 37 ARCHIVES GEN. PSYCHIATRY 51 (1980). Special review of such new clinical procedures is, however, appropriate.

informed consent before administering treatment, except where there is an emergency situation, or where the patient has been judicially found to lack "capacity to make an informed decision concerning treatment." The concept of informed consent has been the subject of extensive legal, medical, and philosophical debate. While some scholars assert that ethical treatment uniformly requires obtaining informed consent, others view the informed consent concept as an ideal that is seldom possible and even more rarely achieved. The theoretical justifications for requiring informed consent emphasize sometimes conflicting goals, including patient autonomy, maximizing good to the patient, respect for individuals, and society's own interests.[47]

At a practical level, the states are split between two standards for regulating what physicians must disclose to their patients to ensure informed consent. In some jurisdictions, a doctor need only impart that information which most doctors in the community would have disclosed in like circumstances.[48] Other jurisdictions, however, take a more "consumer-oriented" view, and require a physician to disclose "information which the patient has every right to expect." In these states, a "patient's right of self-decision shapes the boundaries of the duty to reveal."[49] The risks of therapy must be disclosed if a reasonable person in the position of the patient "would be likely to attach significance to the risk or cluster of risks in deciding whether or not to forego the proposed therapy."[50] Under either standard, the physician then must obtain the "informed," "knowing," and "understanding" consent of the patient.[51]

Unfortunately, neither formula works very well. Several studies have shown that many physicians do not provide a great

---

[47]*See generally* U.S. NAT'L COMM'N FOR THE PROTECTION OF HUMAN SUBJECTS OF BIOMEDICAL & BEHAVIORAL RESEARCH, THE BELMONT REPORT (1978); Mahler, Veatch & Sidel, *Ethical Issues in Informed Consent*, 247 J. A.M.A. 481 (1982).

[48]*See, e.g.*, Natanson v. Kline, 186 Kan. 393, 406, 350 P.2d 1093, 1106, *clarified*, 187 Kan. 186, 190–91, 354 P.2d 670, 673 (1960); R. ROSOFF, INFORMED CONSENT (1980).

[49]Canterbury v. Spence, 464 F.2d 772, 782, 786 (D.C. Cir.), *cert. denied*, 409 U.S. 1064 (1972).

[50]*Id.*; *see also* Cobbs v. Grant, 8 Cal. 3d 229, 241, 502 P.2d 1, 11, 104 Cal. Rptr. 506, 512 (1972). The physician should determine objectively what a reasonable person in the patient's position would want to know. *See* Sard v. Hardy, 281 Md. 432, 439, 379 A.2d 1014, 1022 (1977); Wilkinson v. Vesey, 110 R.I. 606, 624, 295 A.2d 676, 689 (1972). The physician also should explore whether a particular patient, subjectively, has other concerns that could reasonably be addressed. *See* Truman v. Thomas, 27 Cal. 3d 285, 291, 611 P.2d 902, 906, 165 Cal. Rptr. 308, 311 (1980).

[51]*See generally* Capron, *Informed Consent in Catastrophic Disease Research and Treatment*, 123 U. PA. L. REV. 340 (1974); Stromberg, *Health Law Comes of Age: Economics and Ethics In a Changing Industry*, 92 YALE L.J. ____ (1983) (in press).

amount of information bearing on therapeutic decisions—at least not in a way that patients comprehend.[52] If patients do comprehend this information, they retain little of the knowledge.[53] Still murkier is the nexus between comprehension and decisionmaking. Many studies suggest that even competent patients often do not engage in the rational weighing of risks, benefits, and alternatives that is assumed by legal models of informed consent.[54] It is unclear whether most patients really want to know all the risks and uncertainties of treatment,[55] whether they are frequently harmed by such disclosures,[56] and what the law concludes from this.[57]

---

[52]*See, e.g.*, Cassileth, Zupkis, Sutton-Smith & March, *Informed Consent—Why Are its Goals Imperfectly Realized?*, 302 NEW ENG. J. MED. 896 (1980); Garnham, *Some Observations on Informed Consent in Nontherapeutic Research*, 1 J. MED. ETHICS 138 (1975); Gruindner, *On the Readability of Surgical Consent Forms*, 302 NEW ENG. J. MED. 900 (1980); Morrow, *How Readable Are Subject Consent Forms*, 244 J. A.M.A. 56 (1980); Reicken & Ravich, *Informed Consent to Biomedical Research in Veterans Administration Hospitals*, 248 J. A.M.A. 344 (1982); U.S. Nat'l Comm'n for the Protection of Human Subjects of Biomedical & Behavioral Research, Institutional Review Boards, Report and Recommendations, 43 Fed. Reg. 56,174 (1978); *cf.* McCollum & Schwartz, *Pediatric Research Hospitalization: Its Meaning to Parents*, 3 PEDIATRIC RESEARCH 199 (1969).

[53]*See* Appelbaum, Mirkin & Bateman, *Empirical Assessment of Competency to Consent to Psychiatric Hospitalization*, 138 AM. J. PSYCHIATRY 1170 (1981); Leonard, Chase & Childs, *Genetic Counseling: A Consumer's View*, 287 NEW ENG. J. MED. 433 (1972); Robinson & Merav, *Informed Consent: Recall by Patients Tested Postoperatively*, 22 ANN. THORACIC SURGERY 209 (1976); Roth, Meisel & Lidz, *Tests of Competency to Consent to Treatment*, 134 AM. J. PSYCHIATRY 279 (1977). *But see* Woodward, *Informed Consent of Volunteers: A Direct Measurement of Comprehension and Retention of Information*, 27 CLINICAL RESEARCH 248 (1979).

[54]*See, e.g.*, Fellner & Marshall, *Twelve Kidney Donors*, 206 J. A.M.A. 2703 (1968); Kaltte, Lipscomb, Rozynko & Pugh, *Changing the Legal Status of Mental Hospital Patients*, 20 HOSP. & COMMUNITY PSYCHIATRY 199 (1969); McNeil, Pawker, Sox & Tvorsky, *On the Elicitation of Preferences for Alternative Therapies*, 306 NEW ENG. J. MED. 1259 (1982); Munetz, Roth & Cornes, *Tardive Dyskinesia and Informed Consent: Myths and Realities*, 10 BULL. AM. A. PSYCHIATRY & LAW 77 (1982); Park, Slaughter, Covi & Kriffin, *The Subjective Experience of the Research Patient: An Investigation of Psychiatric Outpatients' Reactions to the Research Treatment Situation*, 143 J. NERVOUS & MENTAL DISEASE 199 (1969).

[55]*Compare* Alfidi, *Informed Consent—A Study of Patient Reaction*, 216 J. A.M.A. 1325 (1971) [hereinafter cited as Alfidi, *Informed Consent*], *with* Alfidi, *Controversy, Alternatives and Decisions in Complying with the Legal Doctrine of Informed Consent*, 114 RADIOLOGY 231 (1975).

[56]*See* Alfidi, *Informed Consent*, *supra* note 55; Denney, Williamson & Renn, *Informed Consent: Emotional Responses of Patients*, 60 POST-GRADUATE MED. 205 (1976); Lankton, Batchelder & Ominsky, *Emotional Responses to Detailed Risk Disclosure for Anesthesia: A Prospective, Randomized Study*, 46 ANESTHESIOLOGY 294 (1977); Meisel & Roth, *What We Do and Do Not Know About Informed Consent*, 246 J. A.M.A. 2473 (1981); Rockwell & Pepitone-Rockwell, *The Emotional Impact of Surgery and the Value of Informed Consent*, 63 MED. CLINICS N. AM. 1341 (1979). Moreover, these studies generally do not include psychiatric patients, whose mental state is itself the major medical problem.

[57]*See, e.g.*, Nishi v. Hartwell, 52 Hawaii 188, 191–92, 473 P.2d 116, 119 (1970) ("A physician may withhold information of untoward consequences where full disclosure

Where mental patients are involved, there is even more reason to question whether all the elements of informed consent have been satisfied.[58] State legislatures often require physicians to obtain substituted or proxy consent for incompetent patients. Yet, "despite the importance of proxy decisionmaking, there has been hardly any empirical investigation of the many facets of the subject."[59] Amid such uncertainties, the most a Model Law can do is to identify the critical elements of informed consent that ought to be satisfied.

The Model Law recognizes four such elements of informed consent.[60] First, the consent must be "informed," that is, the treatment team must explain to the patient the nature and effects, including the appreciable risks, of the proposed treatment.[61] The scope of information disclosed as part of this process will depend on clinical and legal circumstances in each case. In particular, psychiatrists may consider the possibility of serious mental trauma if certain information is presented to an acutely mentally ill patient.[62] Ordinarily, however, the staff

---

would be detrimental to patient's best interest."); Scott v. Bradford, 606 P.2d 554 (Okla. 1979) (physician's therapeutic privilege permits nondisclosure of risks where "[full] disclosure would be detrimental to patient's total care and best interests"). *But see, e.g.*, Canterbury v. Spence, 464 F.2d 772, 789 (D.C. Cir.), *cert. denied*, 409 U.S. 1064 (1972) ("The privilege does not accept the paternalistic notion that the physician may remain silent simply because divulgence might prompt the patient to forego therapy the physician feels the patient really needs.").

[58]*See* Appelbaum, Mirkin & Bateman, *supra* note 53; Stone, *Informed Consent: Special Problems For Psychiatry*, 30 HOSP. & COMMUNITY PSYCHIATRY 321 (1979); *cf.* Soskis, *Schizophrenic and Medical Patients as Informed Drug Consumers*, 35 ARCHIVES GEN. PSYCHIATRY 645 (1978).

[59]Meisel & Roth, *supra* note 56, at 2475.

[60]Only a few states statutorily define "informed consent." *See* ARIZ REV. STAT. ANN. § 36-501(13) (1976); CAL. WELF. & INST. CODE § 5326.2 (West 1976 & Supp. 1982); COLO. REV. STAT. § 27-10.5(102)(2)(a) (1976); CONN. GEN. STAT. § 17-2060(h) (1974); IDAHO CODE § 39-4302 (1977); NEV. REV. STAT. § 433-484(1)(b) (1981); OHIO REV. CODE ANN. § 2317.54 (Page 1981); S.D. CODIFIED LAWS ANN. § 27A-1-2(8) (1976). *See generally* RESTATEMENT (SECOND) OF TORTS §§ 892A(2), 892B(3) (1972).

[61]Some would require more information to be disclosed in every case, including "the nature and seriousness of the patient's disorder," the "probable frequency and duration" of the treatments proposed, the "nature, degree, duration and the probability of side effects," whether "there exists a division of opinion as to the efficacy of the proposed treatment," and "the reasonable alternatives and why the physician is recommending this particular treatment." CAL. WELF. & INST. CODE § 5326.2 (West 1976 & Supp. 1982). The authors do not believe that such rigid requirements should be statutorily mandated, and most current state laws do not do so. *See generally* Burra, Kimberley & Miura, *Mental Competence and Consent to Treatment*, 25 CAN. J. PSYCHIATRY 251 (1980); Owens, *When is a Voluntary Commitment Really Voluntary*, 47 AM. J. ORTHOPSYCHIATRY 104 (1977).

[62]*See* Current Opinions of the Judicial Council of the American Medical Association § 8.07 (1982) (recognizing an exception to the duty to disclose "when risk disclosure poses such a serious psychological threat of detriment to the patient as to be medically contraindicated"). Such a "therapeutic privilege" should, however, be invoked only

should explain at least the basic nature, purpose, effects, and substantial risks of the treatment, as well as respond to a patient's questions.

Second, consent must be "knowing." This refers to the cognitive process of the patient. He must be able to understand the information provided and to engage in a rational decisionmaking process about whether or not to accept the treatment.[63]

Third, consent must be "voluntary." This refers to a patient's affective or volitional process. A patient should make his choice free of undue influence by physicians, family, clergy, or others.[64] If there is substantial reason to doubt the voluntariness of consent, the treatment team should seek a court determination of the patient's capacity to make treatment decisions.

Fourth, the person must be able to communicate his decision intelligibly. Thus, a person in a catatonic stupor or suffering from profound mania may not be able to consent to treatment.[65] The Model Law requires *written* consent to a course of therapy,

---

rarely, and the reasons for its application should in every instance be explained thoroughly in the patient's medical record.

[63]Among the phrases used to describe this capacity are judgment, reality-testing, rationality, and decisionmaking capacity. *See generally* Schwartz, *A Revised Checklist to Obtain Consent to Treatment with Medication*, 31 HOSP. & COMMUNITY PSYCHIATRY 765 (1980). This capacity also implies a lack of major thought disorders such as hallucinations, loose association, or delusions. *See* Burra, Kimberley & Miura, *supra* note 61. As one review of the literature said,

> [t]his [criterion] has been phrased in a variety of ways, asking that the subject "appreciate the consequences of giving or withholding consent," have "a sense of who he is and why he is agreeing," recognize, "in a mature fashion, the implications of alternative courses of action and appreciate both cognitively and affectively the nature of the thing to be decided," or "appreciate what is relevant to forming a judgment of the issue in question."

Appelbaum & Roth, *Competence to Consent to Research*, 39 ARCHIVES GEN. PSYCHIATRY 951, 955 (1982); *see also* Section 3 (defining "lacks capacity to make an informed decision concerning treatment").

[64]Voluntariness encompasses many subtle issues that are beyond the scope of this piece. Some argue, for example, that an institutionalized person, per se, cannot voluntarily consent. *See, e.g.*, Kaimowitz v. Michigan Dept. of Pub. Health, 13 CRIM. L. RPTR. (B.N.A.) 2452 (Cir. Ct. Wayne Co., Mich. Aug. 22, 1973) (*reprinted in* 42 U.S.L.W. 2063). Others strongly disagree. *See, e.g.*, Murphy, *Therapy and the Problem of Autonomous Consent*, 2 INT'L J.L. & PSYCHIATRY 415 (1979). Others sometimes distinguish between "assent," "consent," and "acquiescence." *See* NAT'L COMM'N FOR THE PROTECTION OF HUMAN SUBJECTS OF BIOMEDICAL & BEHAVIORAL RESEARCH, REPORT AND RECOMMENDATIONS ON RESEARCH INVOLVING THOSE INSTITUTIONALIZED AS MENTALLY INFIRM (1978). The notion of voluntariness also assumes the absence of a pathologic relationship such as excessive dependency or passivity, *see* Appelbaum & Roth, *supra* note 63, and the lack of any other pathology, such as total denial, *see, e.g.*, Faden & Faden, *False Belief and the Refusal of Medical Treatment*, 3 J. MED. ETHICS 133 (1977); Roth, Appelbaum, Sallee, Reynolds & Huber, *The Dilemma of Denial in the Assessment of Competency To Refuse Treatment*, 139 AM. J. PSYCHIATRY 910 (1982).

[65]*See generally* Appelbaum & Roth, *supra* note 63.

although not to every specific treatment or dosage, in order to protect both the patient and the treatment staff. As may be apparent from this brief description of the difficulties of the informed consent doctrine, clinicians and courts must take great care if any of the goals of informed consent is to be realized.

> ***"lacks capacity to make an informed decision concerning treatment"*** **means that the person, by reason of his mental disorder or condition, is unable, despite conscientious efforts at explanation, to understand basically the nature and effects of hospitalization or treatment or is unable to engage in a rational decisionmaking process regarding such hospitalization or treatment, as evidenced by inability to weigh the possible risks and benefits.**

COMMENTARY: The Model Law makes lack of "capacity to make an informed decision concerning treatment" a specific criterion for involuntary commitment, and also precludes accepting informed consent from an incompetent patient. Where the issue of capacity arises in a commitment proceeding, the court makes the determination, as provided for by subsection 6.C.4. Where the issue arises in connection with consent to hospitalization or treatment, if the physician decides that he cannot accept consent because the patient lacks capacity, then the court may resolve any dispute as to substituted consent.[66] The Model Law's term refers to an informed decision "concerning treatment," but the definition makes clear that it also refers to decisions regarding "hospitalization."

Before assessing a patient's capacity or lack thereof, the treatment team must conscientiously try to explain the nature and effects of the proposed hospitalization or treatment to the patient. When appropriate, the treatment staff may enlist the aid of a patient's family, friends, clergy, or others.

A person lacks capacity if, due to his mental disorder or condition, he cannot understand the basic nature and effects of the proposed hospitalization or treatment. A person does not lack capacity simply because he refuses treatment, which he might do, for example, because he is a Christian Scientist or is risk averse (being excessively so in the physician's opinion).[67]

---

[66]In this respect, assessing capacity requires both a legal and a medical judgment. *See generally* Appelbaum & Roth, *Clinical Issues in the Assessment of Competency*, 138 AM. J. PSYCHIATRY 1462 (1981).

[67]It is critical to avoid the trap of reasoning "backwards" from the conclusion that one does or does not agree with a patient's decisions to an inference that he is or is not competent to decide. *Id.*

Understanding requires a fundamental appreciation of those aspects of the proposed treatment that a reasonable person would find significant in decisionmaking. A patient need not understand every technical feature of a proposed therapy.

Even if an individual can understand the nature and effects of treatment, he lacks capacity if, due to his mental disorder or condition, he cannot engage in any rational decisionmaking process because, for example, he is unable to weigh the risks and benefits of the proposed therapy. The definition requires inability to engage in any rational process, not simply the one that the physician or court would employ. Rational modes of thinking may be unusual, eccentric, or even inconsistently related to reality. A patient's phobia, for example, might distort his apprehension or appreciation of particular facts without impairing his ability to reason concerning other facts or decisions.[68] Another patient's delusions, however, might broadly impair his ability to reason. An individual afflicted with a severe mental disorder may be unable to pay attention to and assimilate information, or his disorganized thoughts may preclude him from engaging in anything resembling a rational process.[69] Only this type of patient lacks capacity under the Model Law.[70]

> ***"likely to cause harm to himself or to suffer substantial mental or physical deterioration"*** **means that, as evidenced by recent behavior, the person (1) is likely in the near future to inflict substantial physical injury upon himself, or (2) is substantially unable to provide for some of his basic needs such as food, clothing, shelter, health, or safety, or (3) will if not treated suffer or continue to suffer severe and abnormal mental, emo-**

---

[68]*See* Appelbaum & Roth, *supra* note 63. A patient's ability to reason may be severely impaired by a variety of other conditions such as anxiety, panic, depression, euphoria, anger, or obsessive preoccupation.

[69]Several empirical studies conclude that a large percentage of current mental patients do lack capacity. *See* Appelbaum, Mirkin & Bateman, *supra* note 53; Munetz, Roth & Cornes, *supra* note 54; Olin & Olin, *Informed Consent in Voluntary Mental Hospital Admissions*, 132 AM. J. PSYCHIATRY 938 (1975); Palmer & Wohl, *Voluntary-Admission Forms: Does the Patient Know What He's Signing?*, 23 HOSP. & COMMUNITY PSYCHIATRY 250 (1972); Pryce, *Clinical Research upon Mentally Ill Subjects Who Cannot Give Informed Consent*, 133 BRIT. J. PSYCHIATRY 366 (1978); Soskis, *supra* note 58.

[70]The Model Law's definition is far more specific than the definition of competence as "capable" of making a responsible treatment decision that was struck down as unconstitutionally vague in Colyar v. Third Judicial Dist. Court, 469 F. Supp. 424 (D. Utah 1979). *Compare In re* B., 156 N.J. Super. 231, 234, 383 A.2d 760, 762 (1977) ("The court finds the patient's refusal to take Prolixin is not, however, based entirely on rational considerations, but reflects delusional thinking."), *with* Lane v. Candura, 6 Mass. App. Ct. 377, 376 N.E.2d 1232 (1978) (patient's irrational refusal to consider medical treatment does not establish incompetency), *and In re* Quackenbush, 156 N.J. Super. 282, 287, 383 A.2d 785, 788 (1978).

tional, or physical distress, and this distress is associated with significant impairment of judgment, reason, or behavior causing a substantial deterioration of his previous ability to function on his own.

COMMENTARY: In subsection 6.C, the Model Law authorizes some involuntary commitments based on the parens patriae doctrine, that is, when necessary for the patient's own good. Such commitments are justified only if it can be predicted that the person will soon harm himself, and that this likelihood is due to his severe mental disorder. This prediction of harm cannot be based solely on descriptions of the person's mental processes; it must be grounded in the patient's "recent behavior."[71] The harm must be more than possible; it must be probable or "likely."

The first of the three possible grounds for parens patriae commitment is that the person "is likely in the near future to inflict substantial physical injury upon himself." Ordinarily, this will mean a recent, credible threat of or attempt at self-mutilation or suicide, accompanied by a mental state indicating a likely recurrence.[72] The harm must be predicted to occur "in the near future"; this accords with the trend in state statutes[73] and court decisions[74] towards defining the time period more explicitly.

The second basis for parens patriae commitment is that the person "is substantially unable to provide for some of his basic needs." Many state laws denote such individuals as "gravely disabled."[75] Inability to provide for even one critical need may

---

[71]The courts are moving toward the view that commitment is not warranted simply because that person "might be dangerous if this pattern of thinking continued as it seemed to be developing." *In re* Conrad, 34 Or. App. 119, 119–120, 578 P.2d 1, 1–2 (1978); *see also In re* Chapman, 67 Ill. App. 3d 382, 385 N.E.2d 56 (1978); *In re* Hatley, 291 N.C. 693, 231 S.E.2d 633 (1977).

[72]*Cf. In re* Mendoza, 433 A.2d 1069, 1071–72 (D.C. 1981) (likely injury to one's career and reputation held not sufficient to justify confinement); N.C. GEN. STAT. § 122-58.2(1)(a) (1981); PA. STAT. ANN. tit. 50, § 7301 (Purdon 1969 & Supp. 1982–1983).

[73]*See, e.g.*, ALA. CODE § 22-52-1 (1975) ("real and present danger"); CAL. WELF. & INST. CODE §§ 5260, 5300 (West 1969) ("imminent" harm); GA. CODE § 88-501(v) (1982) (same); ILL. REV. STAT. ch. 91 1/2, § 1-119 (1981) ("in the near future"); MICH. COMP. LAWS § 330.1401 (1979) (same); MONT. CODE ANN. § 53-21-102(14) (1981) ("imminent" harm); N.C. GEN. STAT. ch. 122-58.2 (1981) ("within the near future"). *But see* CONN. GEN. STAT. § 17-176 (1975) (no standard for proximity of harm); IDAHO CODE § 66-317 (1980) (same); PA. STAT. ANN. tit. 50, § 7301 (Purdon 1969 & Supp. 1982–1983) ("reasonable probability that death, serious bodily injury, or serious physical debilitation would ensue within 30 days").

[74]*See, e.g.*, Suzuki v. Yuen, 617 F.2d 173 (9th Cir. 1980); Colyar v. Third Judicial Dist. Court, 469 F. Supp. 424 (D. Utah 1979). *But see* Hatcher v. Wachtel, 269 S.E.2d 849 (W. Va. 1980).

[75]*See, e.g.*, ARIZ REV. STAT. ANN. § 36-540(C) (1974 & Supp. 1981–1982); ARK.

justify commitment on this basis. For example, a severely mentally ill "bag lady" who scrounges sufficient food and clothes for her needs but who sleeps in city alleys and doorways might be committable because she cannot provide for her health or safety.

One serious concern that has arisen over use of such a "gravely disabled" concept is that it should not be used to commit people whose lifestyle simply offends the majority. Some courts have granted "gravely disabled" commitments based on what appears to be (from the limited data in the published opinions) tenuous grounds.[76] Other courts, however, have refused to order such commitments in a surprising number of cases.[77] Under subsection 6.C of the Model Law, a person is, in effect, considered gravely disabled only if his alleged inability to care for himself is due to a "severe mental disorder," rather than a choice of lifestyle or other factors. To be committable on this basis, he must also lack the capacity to make an informed decision concerning treatment. These factors should minimize abuse and make the Model Law constitutionally acceptable.[78]

The third alternative basis for parens patriae commitment is that, due to a "severe mental disorder," the person is likely to "suffer substantial mental or physical deterioration." This provision makes treatment available to a clearly defined group of severely mentally ill persons commonly excluded from the mental health system by current legal standards.[79] It applies only to persons who will suffer "severe and abnormal mental, emo-

---

STAT. ANN. § 59-1401(C) (1971 & Supp. 1979); CAL. WELF. & INST. CODE § 5008(h)(1) (West 1969 & Supp. 1982); COLO. REV. STAT. § 27-10(102)(5) (1976); CONN. GEN. STAT. § 17-176 (1975); IDAHO CODE § 66-329(k) (1980); *cf.* NEB. REV. STAT. § 83-1009 (1981).

[76]*See, e.g.*, Walker v. Dancer, 386 So. 2d 475 (Ala. Civ. App. 1980); Estate of Roulet, 23 Cal. 3d 219, 590 P.2d 1, 152 Cal. Rptr. 425 (1979); *In re* Paiz, 43 Colo. App. 352, 603 P.2d 976 (1979); *In re* Evans, 86 Ill. App. 3d 263, 408 N.E.2d 33 (1980); *In re* Janovitz, 82 Ill. App. 3d 916, 403 N.E.2d 583 (1980); *In re* Frick, 49 N.C. App. 273, 271 S.E.2d 84 (1980).

[77]*See, e.g.*, County Attorney v. Kaplan, 124 Ariz. 510, 605 P.2d 912 (1980); *In re* Linderman, 417 N.E.2d 1140 (Ind. 1981); *In re* Field, 120 N.H. 206, 412 A.2d 1032 (1980); Sheffel v. Sulikowski, 62 Ohio St. 2d 128, 403 N.E.2d 993 (1980); State *ex rel.* Pifer v. Pifer, 273 S.E.2d 69 (W. Va. 1980).

[78]In Doe v. Gallinot, 486 F. Supp. 983, 991 (C.D. Cal. 1979), the court conceded that "there could be some difference in interpretation of what constitutes basic need," but held a "gravely disabled" standard not unconstitutionally vague. *See also* Colorado v. Taylor, 618 P.2d 1127 (Colo. 1980). Some courts have read into the "gravely disabled" standard a "lack of capacity" requirement. *See, e.g.*, Walker v. Dancer, 386 So. 2d 475 (Ala. Civ. App. 1980); Northern v. State Dep't of Human Servs., 575 S.W.2d 946 (Tenn. 1978).

[79]*See supra* notes 8–14 and accompanying text.

tional, or physical distress."[80] The usual grief and depression over the death of a loved one clearly would not be "abnormal," and would therefore not justify commitment. Physical distress caused by a mental illness (for example, intractable pain, or delirium tremens (DTs) associated with an alcoholic brain syndrome) could suffice, but other physical distress would not.

Further, the distress must impair the person's reason or behavior so severely as to cause a major decline in his ability to function. This requirement suggests an acute episode or sudden collapse of mental state (decompensation). If a usually withdrawn and solitary person shuns society, it may be less solid evidence of a sudden change in mental condition than if a gregarious, well-adjusted person does so. The Model Law thus avoids judging individual lifestyles, but permits commitment of severely mentally ill individuals who are moving toward sudden collapse.[81]

> ***"likely to cause harm to others"*** **means that as evidenced by recent behavior causing, attempting, or threatening such harm, a person is likely in the near future to cause physical injury or physical abuse to another person or substantial damage to another person's property.**

COMMENTARY: In subsection 6.C, the Model Law authorizes involuntary commitment for the protection of society. Such police power commitments are predicated on "recent behavior causing, attempting, or threatening . . . harm" to others. They cannot be based simply on a diagnosis of a person's mental state, or on abstract verbalizations alone. As with involuntary parens patriae commitments, the potential for harm must arise from a "severe mental disorder" and the individual must meet the other criteria for commitment specified in subsection 6.C.

---

[80]A recent draft report on commitment in Pennsylvania stresses the need for such a provision. *See* Pepe, Evaluation of the Mental Health Procedures Act (Pa. Dep't Pub. Welfare 1981), *quoted in* Appelbaum, *Civil Commitment: Is the Pendulum Changing Direction?*, 33 HOSP. & COMMUNITY PSYCHIATRY 703 (1982); *see also* Durham & Pierce, *Beyond Deinstitutionalization: A Commitment Law in Evolution*, 33 HOSP. & COMMUNITY PSYCHIATRY 216 (1982) (describing Washington State's experience).

[81]This provision should not be as constitutionally vulnerable as the old "in need of treatment" statutes. Under the Model Law, commitment requires a severe mental disorder, treatability, lack of capacity, and clear and convincing proof of likely harm. It provides for a definite term of commitment and the right to treatment. The Supreme Court in O'Connor v. Donaldson, 422 U.S. 563 (1975), did not discuss the constitutionality of such a set of commitment criteria, but decisions since then indicate that this provision should be upheld. *See, e.g.*, Doe v. Gallinot, 486 F. Supp. 983 (C.D. Cal. 1979); Reynolds v. Sheldon, 404 F. Supp. 1004 (N.D. Tex. 1975); Colorado v. Taylor, 618 P.2d 1127 (Colo. 1980).

Studies have shown that psychiatric predictions of future dangerous conduct, like other clinical, sociological, and actuarial predictions, are not very reliable.[82] Because we lack better guidance, however, society has called upon psychiatrists and psychologists to make such predictions. To ensure some minimum degree of reliability, the Model Law requires that predictions of such dangerous conduct be based on recent behavior of the person.

How much threatening behavior must be observed to justify commitment? Continuing legal debate focuses on whether to require a recent "overt act." The "overt act" locution is unfortunate, because while the term has been borrowed from criminal law, its meaning has been considerably altered.[83] Nevertheless, several state laws require an "overt act, attempt, or threat" for commitment.[84] The courts have divided on the issue, but most cases indicate that an overt act is not always necessary.[85] As some decisions recognize, while "there are instances in which a psychiatrist can determine from a psychiatric clinical examination [alone] that a mentally ill person is reasonably likely to injure himself or another," such "cases may be relatively few."[86] The Model Law avoids confusion over the term "overt act" by

---

[82]*See, e.g.*, J. MONAHAN, *supra* note 25; A. STONE (WITH C. STROMBERG), *supra* note 2, at 25–40.

[83]In the law of criminal conspiracy an "overt act" means any act—even an innocuous act such as mailing a letter or making a telephone call—done for the purpose of advancing the conspiracy. *See* Braverman v. United States, 317 U.S. 49 (1942); W. LAFAVE & A. SCOTT, CRIMINAL LAW 477–78 (1972). In civil commitment cases, however, the courts sometimes use "overt act" to mean an action that itself imperils someone, rather than a mere threat.

[84]*See, e.g.*, ALA. CODE § 22-52-1 (1975); ARK. STAT. ANN. § 59-1401(A)(B) (1971 & Supp. 1979); GA. CODE § 37-3-1(12) (1982); HAWAII REV. STAT. § 334-1 (1976); OHIO REV. CODE ANN. § 5122.01 (Baldwin 1980); WIS. STAT. § 51.20 (1979–1980).

[85]Compare decisions requiring an overt act, such as Suzuki v. Alba, 438 F. Supp. 1106 (D. Hawaii 1977), *modified*, 617 F.2d 173 (9th Cir. 1980); Stamus v. Leonhardt, 414 F. Supp. 439 (S.D. Iowa 1976); Doremus v. Farrell, 407 F. Supp. 509 (D. Neb. 1975); Lessard v. Schmidt, 349 F. Supp. 1078 (E.D. Wis. 1972), *vacated and remanded on other grounds*, 414 U.S. 473 (1974), *judgment reinstated*, 413 F. Supp. 1318 (E.D. Wis. 1976), with cases holding that an overt act is not required, such as Colyer v. Third Judicial Dist. Court, 469 F. Supp. 424 (D. Utah 1979); United States *ex rel.* Mathews v. Nelson, 461 F. Supp. 707 (N.D. Ill. 1978); Bell v. Wayne County Gen. Hosp., 384 F. Supp. 1085 (E.D. Mich. 1974); People v. Howell, 196 Colo. 408, 586 P.2d 27 (1978) (en banc); *In re* Bumper, 441 A.2d 975 (D.C. 1982); *In re* Snowden, 423 A.2d 188 (D.C. 1980); People v. Sansone, 18 Ill. App. 3d 315, 309 N.E.2d 733 (1974); *In re* Sonsteng, 175 Mont. 307, 573 P.2d 1149 (1977); Scopes v. Shah, 59 A.D.2d 203, 398 N.Y.S.2d 911 (1977); *In re* Hernandez, 46 N.C. App. 265, 264 S.E.2d 780 (1980); *In re* Salem, 31 N.C. App. 57, 228 S.E.2d 649 (1976); Reed v. Texas, 622 S.W.2d 910 (Tex. Ct. App. 1981).

[86]United States *ex. rel.* Mathews v. Nelson, 461 F. Supp. 707, 711 (N.D. Ill. 1978); *cf. In re* Guffey, 54 N.C. App. 462, 283 S.E.2d 534 (1981) (holding that mere pugnaciousness does not justify commitment).

requiring prediction to be based on "recent behavior causing, attempting, or threatening" harm. A threat might be less than an "overt act" and yet be credible enough to warrant commitment.[87]

How recent must the predicate behavior have been? A few states require that such acts have occurred within a particular time period such as the past thirty days.[88] The courts have applied a somewhat more elastic measure.[89] In our view, the variety of factors potentially relevant to a prediction of harm calls for some flexibility,[90] and accordingly the Model Law refers simply to "recent behavior."

What kinds of predicted harm justify commitment? Likely physical violence to another or serious criminal activity clearly would suffice.[91] Although some current state laws require a prediction of "serious physical injury" or "serious bodily harm,"[92] the Model Law also protects against "physical abuse" short of "injury." Such "physical abuse" would include conduct of a patient who, though not violently assaultive, repeatedly pinches, kisses, or pats the breasts of women he encounters on the street, or of a patient who engages in pedophilia. Such behavior often causes victims recognizable harm, including mental trauma and consequent physical symptoms.

The Model Law does not, however, permit commitment on

---

[87]*See, e.g., In re* Oseing, 296 N.W.2d 797 (Iowa 1980); *In re* F.B., 615 P.2d 867, 869 (Mont. 1980); *In re* Goedert, 180 Mont. 484, 487, 591 P.2d 222, 224 (1979) ("While not every threat can be considered an overt act, the testimony and circumstances of this case [showed] that [a] threat to kill is a verbal act that falls within the definition of an 'overt act.'"); Hill v. County Bd. of Mental Health, 203 Neb. 610, 279 N.W.2d 838 (1979); Lux v. Mental Health Bd., 202 Neb. 106, 274 N.W.2d 141 (1979); N.C. STAT. ANN. § 122-58.2 (1981). *But see* Oregon v. Barker, 42 Or. App. 563, 600 P.2d 958 (1979).

[88]*See, e.g.*, ARIZ. REV. STAT. ANN. § 36-501(3) (West Supp. 1981–1982) (30 days for threats, 180 days for assaults); PA. STAT. ANN. tit. 50, § 7301 (Purdon Supp. 1982–1983) (30 days).

[89]*See, e.g., In re* Stephenson, 67 Ill. 2d 544, 367 N.E.2d 1273 (1977); Hill v. County Bd. of Mental Health, 203 Neb. 610, 615, 279 N.W.2d 838, 840 (1979) (court refused to "blindly disregard other incidents [several years before] which give some insight into the man's mental state"); *In re* Prime, 120 N.H. 849, 851, 424 A.2d 804, 806 (1980) (threat two and one-half months before held to be sufficiently recent in light of "consistent pattern" of such threats).

[90]These factors include the seriousness of the act, its nexus to the person's current mental state, and the chance that precipitating conditions will recur.

[91]*See In re* Snowden, 423 A.2d 188 (D.C. 1980) (threats and attempts to cause physical harm to others); *In re* Patterson, 90 Wash. 2d 144, 579 P.2d 1335 (1978) (en banc) (other felonies).

[92]*See, e.g.*, ARK. STAT. ANN. § 59-1401 (West Supp. 1979); CAL. WELF. & INST. CODE § 5300 (West 1972); GA. CODE. § 37-3-1(12) (1982); ILL. REV. STAT. ch. 91 1/2, § 1-119 (1981); MICH. COMP. LAWS § 330.1401 (1979); N.C. GEN. STAT. § 122.58.2 (1981).

the basis of acts that merely annoy, disgust, or cause mental upset to another person. We do not see how any standard that encompassed such acts could distinguish between truly harmful acts and behavior inherently protected by the "right to be different" that our judicial and political system cherishes.[93] Although "injury" and "abuse" on the one hand and mere "annoyance" on the other are vague guideposts, they chart a general path that can best be followed by judicial interpretation in the common law tradition.[94]

The Model Law also permits the risk of "substantial property damage" to justify commitment.[95] Personal liberties are generally more important than trivial property rights, such as that of having an untrampled flower bed. If, however, a patient checks into a hotel, spray paints the walls of his room, destroys the furniture, and threatens to do the same in the lobby, this may well constitute a threat of property damage sufficient to warrant commitment if the other criteria are satisfied.[96]

How soon must the predicted harm be thought likely to occur? The Model Law's "near future" standard is more elastic than "imminent harm,"[97] but is more definite than if no standard of proximity were stated.[98] Courts differ as to how proximate the predicted harm must be to justify commitment,[99] but the standard of "in the near future" should provide acceptable guidance. The time period permissibly encompassed also may be longer when the anticipated harm is greater and shorter when the harm is less.

---

[93]*See* N. KITTRIE, THE RIGHT TO BE DIFFERENT (1972); L. TRIBE, AMERICAN CONSTITUTIONAL LAW § 15-9 (1978).

[94]*See In re* Bumper, 441 A.2d 775 (D.C. 1982).

[95]Recently, a few states have deleted likely property damage as a separate basis for commitment. *See, e.g.*, KAN. STAT. ANN. § 59-2902 (Supp. 1980); *cf. In re* Gatson, 3 Kan. App. 2d 265, 593 P.2d 423 (1979). A federal court has invalidated the Hawaii law that authorized commitment based on damage to "any property," declaring that the Constitution requires some higher standard such as "substantial property damage" before a person can be deprived of his liberty. Suzuki v. Yuen, 617 F.2d 173, 176 (9th Cir. 1980).

[96]*Cf. In re* F.B., 615 P.2d 867 (Mont. 1980).

[97]*See, e.g.*, CAL. WELF. & INST. CODE §§ 5260, 5300 (West 1972); MONT. CODE ANN. § 53-21-102(14) (1981); *see also* ARIZ. REV. STAT. ANN. § 36-501(3) (West Supp. 1981–1982) (within 30 days).

[98]*See, e.g.*, CONN. GEN. STAT. § 17-176 (West Supp. 1982); IDAHO CODE § 66-317 (Supp. 1982); N.C. GEN. STAT. § 122-58.2 (1981).

[99]*Compare* Suzuki v. Yuen, 617 F.2d 173, 178 (9th Cir. 1980) ("imminent" harm required), *and* Colyar v. Third Judicial Dist. Court, 469 F. Supp. 424, 434 (D. Utah 1979) (same), *with* Hatcher v. Wachtel, 269 S.E.2d 849, 852 (W. Va. 1980) (rejecting "imminent harm" standard in favor of "substantial risk of harm in the foreseeable future" test).

How likely must the anticipated harm be? For most kinds of harm, it will mean "probable," or "more likely than not."[100] This standard, however, permits use of a sliding scale on which greater anticipated harm requires lesser certainty to justify commitment. Even a slight chance of mass murder within the next month surely meets the dangerousness criterion while a slight chance of slapping someone would not.

> **"*patient*" means any person receiving evaluation, care, or treatment under this Act, except that for purposes of the rights provided in Section 10, "patient" shall refer only to persons in residential treatment programs.**

COMMENTARY: Certain rights—such as the rights to visits and recreation provided in subsections 10.I and 10.J—by their nature apply only to patients in residential programs. Other rights—such as the right to treatment provided in Section 9 and subsections 10.C and 10.D—may be appropriate for those in outpatient settings as well. Although the Model Law does not prescribe how various patients' rights apply in out-patient programs, its definitions may be suggestive.

> **"*person*" means, for purposes of any provision of this Act authorizing the commitment or treatment of a "person," an individual aged eighteen years or older.**

COMMENTARY: The Model Law applies only to adults; involuntary hospitalization of minors involves many other issues that the Model Law does not address. The Model Law, like many existing state laws,[101] sets eighteen as the age of majority for purposes of commitment as an adult. Some states, however, may choose to set a different threshold age.

> **"*psychosurgery*" means any procedure that by direct access to the brain, removes, destroys, or interrupts the continuity of brain tissue which is histologically normal (as distinguished**

---

[100]The Model Law's standard of "likely" harm is more stringent than some current state laws. *See, e.g.*, ARIZ. REV. STAT. ANN. § 36-501(3) (West Supp. 1981–1982) (the person "constitutes a danger of [harm]"). It is less stringent than others. *See, e.g.*, ALA. CODE. § 22-52-1 (Supp. 1982) ("real and present threat of substantial harm"); MASS. GEN. LAWS ANN. ch. 123, § 1 (West Supp. 1982–1983) ("a very substantial risk"); PA. STAT. ANN. tit. 50, §§ 7301, 7304 (Purdon Supp. 1982) ("clear and present danger to others"). It is close to many state laws requiring "a substantial risk" of harm. *See, e.g.*, CONN. GEN. STAT. § 17-176 (1981); N.Y. MENTAL HYG. LAW § 9.41 (McKinney 1978); OHIO REV. CODE ANN. § 5122.01(B) (Page 1981). Other state laws similarly refer to a "probability" of harm. *See, e.g.*, GA. CODE § 88-501(v) (Supp. 1982); N.C. GEN. STAT. § 122-58-2 (1981).

[101]*See* AM. BAR ASS'N, GUARDIANSHIP AND CONSERVATORSHIP 43 (1979) (citing state statutes).

from normal in its physiological or psychological functioning) for the primary purpose of altering behavior or treating a mental disease or disorder. Psychosurgery includes the implantation of electrodes with such an effect and for such a purpose, with or without subsequent electrocoagulation. Psychosurgery does not include neurosurgical procedures designed to treat reliably diagnosed intractable physical pain or epilepsy.

COMMENTARY: Under subsection 8.C of the Model Law, psychosurgery is a special therapy that can be administered only pursuant to mental health department regulations.[102] A storm of legal, medical, ethical, and popular controversy has arisen over the use of psychosurgery.[103] Some argue that psychosurgery is unscientific, that it produces unpredictable results, that it violates the sacredness of the human mind, and that it has been misused by opponents of minorities.[104] Ill-considered suggestions that political violence might have a biological basis once fueled this debate.[105]

In the United States, physicians annually perform about four to five hundred psychosurgical procedures. Few if any are intended to control violence.[106] After reviewing the entire subject of psychosurgery, the National Commission for the Protection

---

[102]Psychosurgery really has very little relevance to the central concerns of the Model Law. It is specifically addressed only because so much controversy surrounds the use of psychosurgery. The authors believe that psychosurgery, as here defined, is rarely justified. Because there may not now exist sufficient knowledge about the brain to permit the responsible use of psychosurgery, a moratorium on its use might well be appropriate.

[103]*See, e.g.*, *Hearings on Psychosurgery Before the Subcomm. on Health of the Comm. on Labor and Public Welfare Meeting Jointly With the Subcomm. on Health and Hospitals of the Comm. on Veterans Affairs*, 93d Cong., 1st Sess. (1973); M. CRICHTON, THE TERMINAL MAN (1972); J. DELGADO, PHYSICAL CONTROL OF THE MIND: TOWARD A PSYCHOCIVILIZED SOCIETY (1969); V. MARK & F. ERVIN, VIOLENCE IN THE BRAIN (1970); NAT'L COMM'N FOR THE PROTECTION OF HUMAN SUBJECTS OF BIOMEDICAL & BEHAVIORAL RESEARCH, REPORT AND RECOMMENDATIONS ON PSYCHOSURGERY (1977) [hereinafter cited as PSYCHOSURGERY REPORT]; Breggin, *Psychosurgery for the Control of Violence: A Critical Review*, in NEURAL BASES OF VIOLENCE AND AGGRESSION 371 (W. Fields & W. Sweet ed. 1975); Geschwind, *Effects of Temporal Lobe Surgery on Behavior*, 289 NEW ENG. J. MED. 480 (1973); Knowles, *Beyond the Cuckoo's Nest: A Proposal for Federal Regulation of Psychosurgery*, 12 HARV. J. ON LEGIS. 610 (1975); Mason, *Brain Surgery to Control Behavior*, EBONY, Feb. 1, 1973; *The New Psychosurgery*, 226 J. A.M.A. 799 (1973) (editorial); Older, *Psychosurgery: Ethical Issues and A Proposal for Control*, 44 AM. J. ORTHOPSYCHIATRY 661 (1974); *Symposium: Psychosurgery*, 54 B.U.L. REV. 215 (1974).

[104]*See, e.g.*, Breggin, *The Return of Lobotomy and Psychosurgery*, *reprinted in* 118 CONG. REC. 5567 (1972); Chorover, *Big Brother and Psychotechnology*, PSYCHOLOGY TODAY, Oct. 1973, at 43. *See generally* Mark, *Psychosurgery Versus Anti-psychiatry*, 54 B.U.L. REV. 217 (1974).

[105]*See, e.g.*, Mark, Sweet & Ervin, *Role of Brain Disease in Riots and Urban Violence*, 201 J. A.M.A. 895 (1967).

[106]*See* PSYCHOSURGERY REPORT, *supra* note 103, at 28 & app. I-27.

of Human Subjects of Biomedical and Behavioral Research concluded that psychosurgery should be regulated, but not prohibited.[107]

Meanwhile, many states have acted to regulate psychosurgery.[108] At least one court declared that an involuntary patient cannot give truly informed consent to psychosurgery;[109] another court held that such a patient has a right to withhold his consent.[110] The Model Law does not attempt to resolve the profound issues concerning psychosurgery, but only to ensure that the practices will be subject to careful scrutiny. The Model Law reflects the view that psychosurgery may be an acceptable treatment of last resort in rare cases, and that it should not be entirely forbidden by law.[111]

Psychosurgery is defined broadly to include any procedure that removes, destroys, or interrupts the continuity of brain tissue.[112] This includes not only radical lobotomies such as were practiced in the 1950's, but also the newer stereotactic procedures that destroy minute amounts of brain tissue.[113] Psycho-

---

[107]PSYCHOSURGERY REPORT, *supra* note 103, at 57–72. The Commission concluded, inter alia, that (1) "[t]he published reports on the effects of psychosurgery are generally of limited usefulness," due to the lack of objective tests administered to the patients; (2) there is some reason to doubt the skill of some persons performing psychosurgery; (3) there is disagreement over which disorders psychosurgery is most likely to help; (4) the evidence contradicts the view that psychosurgery is used disproportionately on blacks and other minorities; and (5) several scientifically acceptable studies have reported high rates of patient satisfaction following psychosurgery. The Commission recommended regulations to ensure informed consent of the patient, institutional review board approval, review of certain procedures by a National Psychosurgery Advisory Board, and court approval for psychosurgery on any involuntarily committed patient. The DHEW issued draft regulations based on this report, but took the further step of absolutely barring any psychosurgery on involuntary patients. The regulations, however, were never finalized.

[108]*E.g.*, CAL. WELF. & INST. CODE § 5326.2–.95 (West Supp. 1982); KY. REV. STAT. § 202B.060 (1982); MASS. GEN. LAWS ANN. ch. 123, § 23 (West Supp. 1982–1983); N.J. STAT. ANN. § 30:4-24.2 (West 1981); OHIO REV. CODE ANN. §§ 5122.27.1, 5123.86 (Page 1981).

[109]Kaimowitz v. Michigan Dept. of Pub. Health, 13 CRIM. L. RPTR. (B.N.A.) 2452 (Cir. Ct. Wayne Co., Mich. Aug. 22, 1973) (*reprinted in* 42 U.S.L.W. 2063).

[110]Aden v. Younger, 57 Cal. App. 3d 662, 129 Cal. Rptr. 535 (1976).

[111]*See* Bridges, Goktepe & Maratos (with Browne & Young), *A Comparative Review of Patients with Obsessional Neurosis and with Depression Treated by Psychosurgery*, 123 BRIT. J. PSYCHIATRY 663 (1973); Sweet, *Treatment of Medically Intractable Mental Disease by Limited Frontal Leucotomy—Justifiable?*, 289 NEW ENG. J. MED. 1117 (1973).

[112]*Cf.* ARK. STAT. ANN. § 59-1401(f) (West Supp. 1979); CAL. WELF. & INST. CODE § 5325(g) (West Supp. 1981); CONN. GEN. STAT. § 17-206a(j) (1981); N.M. STAT. ANN. § 43-1-3(P) (1978); Bullock, *Neuroscientists on Psychosurgery*, 32 ARCHIVES NEUROLOGY 73 (1975) (psychosurgery is the "surgical removal or destruction of brain tissue in the absence of organic brain disease, with the primary intent of altering the behavior of the patient").

[113]*See* Kelly, Richardson & Mitchell-Heggs, *Stereotactic Limbic Leucotomy: Neurophysiological Aspects and Operative Technique*, 123 BRIT. J. PSYCHIATRY 133 (1973).

surgery, however, only encompasses surgery on tissue that is histologically normal (that is, normal at the cellular or tissue level as seen through a microscope). Thus, the Model Law does not restrict surgery on a cancerous brain tumor. Normality is measured at the cellular level because it would be circular to define it with respect to a tissue's chemical or psychological function.

A surgical procedure is psychosurgery only if its *primary* purpose is to treat a mental disease or to alter behavior.[114] Thus, if a surgeon would recommend operating on a brain tumor even absent any gross behavioral disturbance, the therapy would not be psychosurgery even though a secondary purpose might be to eliminate episodes of violence whose onset is associated with the spread of the tumor.[115]

Finally, the definition excludes neurosurgery to treat reliably diagnosed intractable pain or epilepsy. In such cases, there must be sufficient medical evidence to warrant intervention apart from any intention to alter behavior.[116]

> **"*severe mental disorder*" means an illness, disease, organic brain disorder, or other condition that (1) substantially impairs the person's thought, perception of reality, emotional process, or judgment or (2) substantially impairs behavior as manifested by recent disturbed behavior. Mental retardation, epilepsy, or other developmental disabilities do not, in themselves, constitute a severe mental disorder. [States may wish to provide by other provisions of law for persons whose use of or addiction to intoxicating substances warrants hospitalization.]**

COMMENTARY: The threshold criterion for commitment is whether the person's mental condition is serious enough that

---

[114]This accords with the definition in R. CAMPBELL, PSYCHIATRIC DICTIONARY 518 (5th ed. 1981). It differs from the approach used in several states, which turns on whether the purpose is somatic therapy *rather than* psychiatric therapy. *See, e.g.*, CAL. WELF. & INST. CODE § 5325(g) (West Supp. 1982); CONN. GEN. STAT. § 17-206a (1981); N.M. STAT. ANN. § 43-1-3(P) (1978).

[115]Many scholars support this conclusion. *See, e.g.*, PSYCHOSURGERY REPORT, *supra* note 103, at 57 & apps. I-9 to I-24.

[116]*See id.*, at 57 (psychosurgery "does not include (a) electric shock treatments, (b) surgery or other invasions of the brain designed to cure or ameliorate the effects of movement disorders (*e.g.*, epilepsy, parkinsonianism), and (c) excision of brain tumors"); *see also* TEMPORAL LOBE EPILEPSY (M. Baldwin & P. Bailey ed. 1958); Roberts & Valenshes, *Control of Pain Associated with Malignant Disease by Freezing: Cryolencotomy*, 37 CONN. MED. 184 (1973); Solber & Jannetta, *Central Pain and Central Therapy for Pain*, in CURRENT PROBLEMS IN SURGERY 59 (1973). New uses of psychosurgery may also be developed. *See, e.g., New Tissue will be Implanted in Brain to Treat Parkinson's Disease*, N.Y. Times, Nov. 24, 1981, at C1, col. 1.

commitment may be appropriate. Current state laws vary greatly in their definition of the necessary condition. While some are extremely precise,[117] others are unsuitably vague.[118]

The Model Law recognizes the serious social decision that underlies commitment. Only a person suffering from a "severe mental disorder" should lose his liberty. Indeed, commitments based on vague standards of impairment such as "mentally ill" or "in need of treatment" have been held unconstitutional.[119] Under the Model Law, "severe mental disorder" corresponds roughly to a psychotic disorder.[120] It must be stressed that the realities of a person's impairment, not diagnostic categories or nomenclature, determine the existence of a severe mental disorder. Thus, the issue is not the specific diagnosis alone, but whether a person's thought, perception, or other mental process is or is not substantially impaired.[121]

The class of severe mental disorders includes not only illness and disease, but also organic brain disorders such as toxic alcohol syndrome. Commitment based on the latter type of disorders requires a clinically observable physiological state (such as imminent onset of DTs) and not simply a behavior pattern (such as excessive drinking). Other conditions, such as physical trauma to the head, may constitute severe mental disorders if they substantially impair the person's mental faculties.

A severe mental disorder must generate effects sufficiently serious to meet one of two criteria. First, an illness, disease, organic brain disorder, or condition that "substantially impairs" a person's thought, perception, emotion, or judgment may jus-

---

[117]*See, e.g.*, IND. CODE § 16-14-9.1-1(a) (1976) ("a psychiatric disorder which substantially disturbs a person's thinking, feeling or behavior or impairs the person's ability to function"); UTAH CODE ANN. § 64-7-28 (Supp. 1981) ("a psychiatric disorder as defined by the current [APA] Diagnostic and Statistical Manual of Mental Disorders which substantially impairs a person's mental, emotional, behavioral or related functioning").

[118]*See, e.g.*, ARK. STAT. ANN. § 59-101 (1971) ("any person [who is] an idiot, lunatic, or of unsound mind"); COLO. REV. STAT. § 27-10-102 (1982) (person "of such a mental condition that he is in need of medical supervision, treatment, care or restraint"); TEX. REV. CIV. STAT. ANN. art. 5547-4(k) (Vernon Supp. 1982–1983) ("a person whose mental health is substantially impaired").

[119]*See, e.g.*, Goldy v. Beal, 429 F. Supp. 640 (M.D. Pa. 1976); Stamus v. Leonhardt, 414 F. Supp. 439 (S.D. Iowa 1976); Commonwealth *ex rel.* Finken v. Roop, 234 Pa. 155, 339 A.2d 764 (1975) (plurality opinion), *appeal dismissed*, 424 U.S. 960 (1976).

[120]Under exacerbating circumstances, other disorders could meet the standard. *See generally* APA, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS (3d ed. 1980).

[121]*See In re* D.W.H., 411 N.E.2d 721 (Ind. App. 1980) ("dissociative reaction" found to be a psychiatric disorder "only when it creates symptoms or dysfunctions or when it is maladapted"); *In re* Hatley, 291 N.C. 693, 231 S.E.2d 633 (1977).

tify commitment. Substantial impairment means more than transitory disorientation or an isolated phobia or sensitivity. It ordinarily entails broad deterioration in the structure, content, or integration of one of the cognitive or affective processes indicated. Because severe disability will result if any of these processes grossly malfunction, only one need be substantially impaired to justify commitment.[122]

A second criterion is that the illness, disease, organic brain disorder, or condition grossly impairs behavior. The disturbance must be gross, not minor or merely eccentric, and must be demonstrated by recent behavior.

The definition specifically states that "mental retardation, epilepsy, or other developmental disabilities do not, in themselves, constitute a severe mental disorder." Each of these conditions presents unique features that must be dealt with specially.[123] None of these conditions by itself constitutes a severe mental disorder. Their presence, however, does not preclude it; given clinical evidence of other disorders, a severe mental disorder may be found.[124]

Finally, the Model Law notes that states "may"—and the authors believe "should"—provide for persons with alcoholic or narcotic disorders by special laws.[125] Alcohol-related disorders, for example, range from purely behavioral syndromes without reliably diagnosed mental illness (such as drinking too much), to diagnosable mental disorders (such as alcoholism), to very dangerous toxic brain syndromes (such as DTs).[126] Often these are intertwined with other mental disorders. The definition does

---

[122]Severe mental disorders typically produce intense emotional distress in addition to impairing one or more faculties. Absent such distress, it is rarely appropriate to infer a substantial impairment of cognitive or affective processes. In our view, the existence of such suffering should be required as an element of parens patriae commitment to protect against excessive paternalism. *See* A. STONE (WITH C. STROMBERG), *supra* note 2, at 43–82.

[123]Many state laws recognize this principle. *See, e.g.*, N.D. CENT. CODE § 25-03.1-02 (1981) (mental illness does not refer to mental retardation, drug abuse, or alcoholism "per se" although persons suffering from these conditions may also be suffering from "mental illness").

[124]*See generally* Eaton & Menolascino, *Psychiatric Disorders in the Mentally Retarded: Types, Problems and Challenges*, 139 AM. J. PSYCHIATRY 1297 (1982).

[125]A number of state laws already do so. *See, e.g.*, ALASKA STAT. § 47.30.340(10) (1981); ARIZ. REV. STAT. ANN. § 36-501(17) (West Supp. 1981–1982); CAL. WELF. & INST. CODE §§ 5170, 5225, 5340 (West Supp. 1982); CONN. GEN. STAT. ANN. § 17-176 (West Supp. 1982); N.M. STAT. ANN. §§ 43-2-8, 43-2-22 (1979).

[126]*Cf. In re* Marquardt, 100 Ill. App. 3d 741, 427 N.E.2d 411 (1981) (where legislature intentionally did not define mental illness, court refused to conclude that drug abuse is a mental illness merely because of its inclusion in APA, *supra* note 120); *see generally* Vaillant, *Alcoholism in Drug Dependence*, in HARVARD GUIDE TO MODERN PSYCHIATRY (1978).

not preclude alcohol-related disorders from being considered as part of a composite diagnostic picture, nor does it make them sufficient alone to warrant commitment.

> **"*treatment facility*" means a community mental health facility, a general medical facility providing psychiatric services, or other psychiatric facility or program meeting applicable licensing standards, that has been approved for the provision of services under this Act by the Department of Mental Health; provided that no jail or other correctional facility shall be approved as a treatment facility for any persons other than those who could otherwise lawfully be detained there.**

COMMENTARY: The Model Law obligates treatment facilities to protect various rights of a patient.[127] A state's mental health department must approve such facilities in advance to ensure proper management. Treatment facilities may be state or county hospitals, community mental health centers, psychiatric wards of general hospitals, or other psychiatric facilities or programs.[128]

All treatment facilities must meet licensing and accreditation standards.[129] Where a facility specializes, the state should require evidence that the facility can adequately care for and treat the type of patients it serves. Finally, the definition makes clear that correctional facilities can be used only for treating individuals subject to the criminal process.

Section 4. EMERGENCY PSYCHIATRIC EVALUATION.

4.A. *Detention by a Police Officer.*

1. A police officer may take a person into custody, and transport the person to a treatment facility for emergency psychiatric evaluation if and only if:

a. the person would otherwise be subject to lawful arrest and the

[127]Prime among these are the rights to a humane environment and meaningful treatment. *See* Section 9 and subsections 10.B–C.

[128]According to the National Institutes of Mental Health (NIMH), of the approximately 301,000 psychiatric beds in all facilities in 1978, about 185,000 were in state and county mental hospitals; about 29,000 were in psychiatric units of nonfederal general hospitals; about 23,000 were in free-standing psychiatric facilities; about 34,000 were in Veterans' Administration hospitals; and the remainder were in other types of facilities. NIMH, *State and Regional Distribution of Psychiatric Beds*, in MENTAL HEALTH STATISTICAL NOTE 155 (M. Witkin ed. 1981).

[129]While the definition of "treatment facility" refers only to licensing requirements, as the text of subsection 9.A and the discussion of subsection 10.C indicate, all treatment facilities also should meet the standards of the Joint Commission on Accreditation of Hospitals (JCAH). Many psychiatrists opposed incorporating JCAH standards into the Model Law, because many large areas in this country have no accredited facilities. The authors believe that if a treatment facility cannot meet accreditation standards, there is serious doubt as to whether it can provide the quality of treatment necessary to justify parens patriae confinement "for the patient's own good."

police officer believes that the person is in need of emergency psychiatric treatment; or

b. the police officer has probable cause to believe that the person has attempted suicide within the last 48 hours; or

c. the police officer has probable cause to believe, based on his personal observation and investigation or based on the petition of any interested adult under subsection 4.C. and such corroboration as the police officer deems necessary in the circumstances, that the person is suffering from a severe mental disorder as a result of which he is likely to cause harm to himself or to others or is manifestly unable to care for some of his basic needs, and that immediate hospitalization is necessary to prevent harm to the person or to others; or

d. he is acting upon the certification of a licensed physician under subsection 4.B.

2. Any person taken into custody pursuant to this subsection shall be presented promptly to a treatment facility. Correctional facilities shall not be used as temporary shelter for such persons except for the protective custody of the person pending transportation to a treatment facility.

3. Upon or shortly after taking a person into custody, the police officer shall take reasonable precautions to safeguard and preserve the personal property of the person unless a guardian or responsible relative is able to do so. Upon presenting a person to a treatment facility, the police officer shall inform the staff in writing of the facts that caused him to take the person into custody, and shall specifically state whether the person is otherwise subject to arrest.

COMMENTARY: "It cannot be seriously doubted that the state may on occasion have a compelling interest in the emergency detention of those who threaten immediate and serious violence to themselves or others."[130] Civil commitment thus serves to increase social safety. Many also would argue that society has an ethical obligation to care for a severely mentally ill person until he can care for himself. In both these respects, police officers serve society by taking a person to a treatment facility if his behavior appears manifestly disordered.[131] But due in part to growing concern over possible tort liability, policemen increasingly hesitate to perform this function without explicit statutory authority. Such authority should be granted, but must be circumscribed to avoid giving police officers a "roving commission" to take into custody any person whose mental status they doubt.[132]

---

[130]Lynch v. Baxley, 386 F. Supp. 378, 387 (M.D. Ala. 1974).

[131]A recent NIMH study concluded that police officers generally discharge this responsibility in an "impressive" way, and that "in many cases, the officer is one of the few protectors a chronic patient may have." It found that officers often resolve problems on the spot and thereby avoid the need for hospitalization. *Cop on the Beat Is Unsung Friend of the Mentally Ill Citizen*, ADAMHA NEWS, Mar. 1982, at 6.

[132]*Compare* ARK. STAT. ANN. § 59-102 (1971) (duty of peace officers to "arrest" any

To accommodate these competing concerns, the Model Law authorizes police officers to initiate emergency psychiatric evaluation in only four situations. The first case arises where an officer could properly arrest the individual.[133] The authority to arrest ordinarily will be based on the officer's observation of the person's behavior, but it also may arise from warrants or court orders.

Second, a policeman may take a person to a treatment facility for evaluation if the officer has "probable cause" to believe that the person has attempted suicide in the past forty-eight hours.[134] Police officers understand the "probable cause" standard, which governs searches, seizures, and arrests. This provision establishes a similarly substantial prerequisite to police action leading to psychiatric confinement. Although "probable cause" does not require prima facie proof, generally a citizen's verbal report to the officer would not suffice unless the officer investigates and corroborates the alleged facts.[135]

"insane or drunken persons" who are "not in the care of some discreet person"), *and* Gross v. Pomerleau, 465 F. Supp. 1167 (D. Md. 1979) (striking down statute authorizing police officer to arrest any person who he "believes . . . to be mentally ill," because it failed to require an assessment of the likelihood of danger to self or others), *with* ILL. REV. STAT. ch. 91 1/2, § 3-606 (1981) (police officer may arrest person if, based on his personal observation, officer "has reasonable grounds to believe that the person is subject to involuntary admission and [is] in need of immediate hospitalization to protect such person or others from physical harm").

[133]One study found that police believed that they could have arrested about 30% of committed patients. Monahan, Caldeira & Friedlander, *Police and the Mentally Ill: A Comparison of Committed and Arrested Persons*, 2 INT'L J.L. & PSYCHIATRY 509, 513 (1979).

[134]Although suicide is not itself a crime, some state laws punish attempting such an act. W. LAFAVE & A. SCOTT, *supra* note 83, at 69. Our legal system reflects society's desire to prevent suicides. *See generally* N. ST. JOHN-STEVAS, LIFE, DEATH & THE LAW 232–61 (1961); G. WILLIAMS, THE SANCTITY OF LIFE AND THE CRIMINAL LAW 248–310 (1957).

[135]Under established criminal law doctrine, probable cause exists when the facts within the knowledge of the officer or of which he has reasonably trustworthy information are in themselves sufficient to warrant a man of reasonable caution in the belief that a particular crime has been committed by a particular person. *See, e.g.*, United States v. Hill, 500 F.2d 315 (5th Cir. 1974); State v. Kolb, 239 N.W.2d 815 (N.D. 1976). Thus it is less stringent a test than "more-likely-than-not"; "the words 'reasonable cause' are perhaps close to what is meant." United States v. Melvin, 596 F.2d 492, 495 (1st Cir.), *cert. denied*, 444 U.S. 837 (1979); *see also In re* Paiz, 603 P.2d 976 (Colo. Ct. App. 1979).

By requiring probable cause, the Model Law may set a higher standard than is now used by some states. *See, e.g.*, N.Y. MENTAL HYG. LAW § 9.41 (McKinney Supp. 1982–1983) (officer may take into custody "any person who appears to be mentally ill" and likely to harm someone); TEX. REV. CIV. STAT. ANN. art. 5547-27 (Vernon Supp. 1982–1983) (officer "has reason to believe and does believe"). It is roughly equivalent to "reasonable grounds" or "reasonable cause." *See* CONN. GEN. STAT. § 17-183(a)–(b) (1981); N.M. STAT. ANN. § 43-1-10 (1978). "Probable cause" is now used in a number of state laws. *See, e.g.*, CAL. WELF. & INST. CODE § 5202 (West 1982); COLO. REV. STAT. § 27-10-106(4) (1982); CONN. GEN. STAT. ANN. § 17-183(b) (West 1982); UTAH CODE § 64-7-34 (1953).

This provision requires that the attempted suicide occurred within the past forty-eight hours. Although numerically arbitrary, the time frame reflects the clinical observation that many acute suicidal episodes subside within two days. Nothing in the Model Law would prevent a family member who discovers a suicide attempt after forty-eight hours and fears a recurrence from seeking emergency evaluation under other provisions of Section 4 of the Model Law. But a suicide attempt taking place long ago should not be the basis alone for an "emergency" commitment.

Third, where a policeman's direct investigation, or his corroboration of a complaint, establishes probable cause that a person is commitable, the officer may seek emergency evaluation for that person. The officer himself must investigate, verify the facts, and conclude that there is probable cause as to mental illness,[136] likelihood of harm,[137] and the need for immediate confinement. This provision would not authorize "preventive detention" of a potentially dangerous person who is not mentally ill, or of a mentally ill person who is not likely to cause harm to himself or to others, or of a person where there is time to file a petition for thirty-day commitment before any harm is likely to occur.

Fourth, as provided in subsection 4.B, a policeman may act pursuant to a physician's certification.

Paragraph 4.A.2 requires that any person taken into custody must be "promptly" presented to a treatment facility. Ordinarily, this will mean immediately. But the exigencies of law enforcement and the management of hospitals—including transportation problems, shift changes, and possible difficulties with night admissions—may justify some flexibility. Such delay should rarely amount to more than a few hours.

Subsection 3 requires a police officer to take "reasonable precautions" to secure the personal property of a person taken into custody unless a guardian or responsible relative will do so. For example, the officer should lock the person's home and car. The officer need take only reasonable steps, however, and not every conceivable precaution. Preferably a spouse or other

[136]Police officers are not qualified to make definitive psychiatric judgments. They are able, however, to ascertain from all the circumstances whether a person's disordered behavior might be due to a mental disorder so that he should be evaluated professionally rather than simply placed in a jail.

[137]*See* Section 3 (defining "likely to cause harm to himself or to suffer substantial mental or physical deterioration" and "likely to cause harm to others").

responsible relative should perform these functions.[138] Upon presenting a person to a treatment facility, the police officer should (orally) explain why he took the person into custody and state whether the person lawfully could have been arrested.[139] This ensures that the treatment facility knows the available facts and knows whether the person should be released directly or returned to police custody if it does not admit him.[140]

4.B. ***Certification by a Licensed Physician.***

A person may be taken into custody by a police officer, or accepted by an ambulance service, and transported and presented to a treatment facility for emergency psychiatric evaluation, when a licensed physician certifies in writing that he has examined the patient in the last 72 hours or has ongoing medical responsibility for the person and has knowledge of his current condition, and that on such basis he has probable cause to believe that such person is suffering from a severe mental disorder as a result of which: he lacks capacity to make an informed decision concerning treatment; and he is (1) likely to cause harm to himself or to suffer substantial mental or physical deterioration, or (2) likely to cause harm to others; and immediate hospitalization is necessary to prevent such harm.

COMMENTARY: This subsection permits an ambulance service to accept a patient, and a treatment facility to admit him involuntarily, if a licensed physician[141] certifies that he has probable cause to believe that the person meets the standards for emergency evaluation. The physician must affirm his knowledge of the patient's current condition based on either a recent examination or ongoing contact with and medical responsibility for

---

[138]*Cf.* ARIZ. REV. STAT. ANN. § 36-508 (1974); CAL. WELF. & INST. CODE § 5156 (West 1972); COLO. REV. STAT. § 27-10-106(8) (1982).

[139]*Cf.* OHIO REV. CODE ANN. § 5122.10 (Page 1981) (requires police to state facts in writing).

[140]*See* paragraph 4.D.3.

[141]In this and several other Model Law provisions, the APA took the position that only a physician—not other health professionals such as a clinical psychologist—should have clinical responsibility for the decision to hospitalize a patient. The training and abilities of psychiatrists and psychologists do overlap, although most patients who meet the criteria for commitment under the statute will require medication, a traditional domain of physicians. In connection with the Model Law, the authors take no position on the merits of the debate between psychiatrists and psychologists on this issue.

Several states grant other professionals the same power to initiate commitment as physicians. *See, e.g.*, ARK. STAT. ANN. § 59-407 (1971) ("licensed psychologist, social worker, or psychiatric nurse"); CAL. WELF. & INST. CODE § 5150 (West 1972) ("professional person designated by the county"); COLO. REV. STAT. § 27-10-105 (1982) ("a professional person"); ILL. REV. STAT. ch. 91 1/2, § 3-702 (1981) ("psychiatrist . . . or clinical psychologist"); OHIO REV. CODE ANN. § 5122.10 (Page 1981) ("psychiatrist, licensed clinical psychologist, or licensed physician").

the patient.[142] The physician need have only "probable cause" for his judgment because his certification mandates only evaluation; he need not make a definitive diagnosis. Grievance review procedures, possible lawsuits, and the Model Law's sanctions provided in subsection 18.B should be sufficient to deter frivolous certifications.[143]

4.C. *Petition by Any Interested Adult.*

Any interested adult may petition for, or present a person for, emergency psychiatric evaluation by alleging based on personal observation that he has probable cause to believe that such person is suffering from a severe mental disorder as the result of which: he is likely to cause harm to himself or to others or is manifestly unable to care for some of his basic needs; and immediate hospitalization is necessary to prevent harm to the person or to others.

COMMENTARY: The Model Law embodies the belief that care and treatment should be made available to severely mentally ill persons without unnecessary procedural hurdles.[144] Accordingly, any "interested adult" may initiate the process to determine a person's need for emergency evaluation. Several state laws so provide, while others allow petitions only from the person's family.[145] Under the Model Law, "interested" adults may include family members, close friends, neighbors, coworkers, and others who know the allegedly disturbed person in a substantial way. These individuals may have a real interest in securing treatment for him. "Interested" adults should not be construed to include shopkeepers, bus drivers, or others who have had only a passing encounter with the person. In view of the embarassment and trauma which can be caused by having one's mental condition questioned, police departments, treat-

---

[142]*Cf.* ALASKA STAT. § 47.30.020 (1979) (licensed physician must have examined the patient within 15 days); UTAH CODE ANN. § 64-7-34 (1953 & Supp. 1981) (examination within 3 days).

[143]This provision intentionally overlaps with subsection 4.C, which authorizes any interested adult to present a person for emergency evaluation without getting a physician's certification. Subsection 4.B is needed because many people would be too uncertain of their judgment, or of their legal right to petition for commitment, to act independently. They might, however, be willing to ask a family physician to assess whether their concerns about a family member's mental condition are warranted.

[144]*See supra* notes 8–13 & 21–22 and accompanying text.

[145]*Cf.* MD. HEALTH-GEN. CODE ANN. § 10-622 (1982) ("any interested person"); N.Y. MENTAL HYG. LAW §§ 9.27, 9.45 (McKinney Supp. 1982–1983) (parent, spouse, child, physician, health officers, peace officer, or police officer); OHIO REV. CODE ANN. § 5122.11 (Page 1981) (any person); PA. STAT. ANN. tit. 50, § 7304 (Purdon 1969) ("any responsible party"); UTAH CODE ANN. § 64-7-34 (1953 & Supp. 1981) ("any responsible person").

ment facilities, and courts must take care to ensure that "interested" adults do not abuse this power to instigate commitment proceedings. In subsection 18.B, the Model Law imposes penalties for extreme abuses.

This provision permits taking only a nonprotesting person to a treatment facility for evaluation; the Model Law does not authorize a private citizen to use force or to take another person into custody against his will.[146] Under subsection 4.B, an interested adult could invoke the assistance of a police officer, but only if the officer was presented with evidence sufficient to persuade him to act on his own authority.

4.D. ***Treatment Facility Determination.***

1. Upon the presentation of a person to a treatment facility pursuant to this Section 4, the facility shall accept the person and shall promptly examine him to determine whether he meets the criteria for emergency evaluation and treatment set forth in paragraph 2.

2. The person shall be admitted for emergency evaluation and treatment only if the examining psychiatrist determines that there is probable cause to believe that the person suffers from a severe mental disorder as the result of which: he lacks capacity to make an informed decision concerning treatment; and he is (a) likely to cause harm to himself or to suffer substantial mental or physical deterioration, or (b) likely to cause harm to others; and immediate hospitalization is necessary to prevent such harm.

3. If the examining psychiatrist determines that there is not probable cause to believe that the person meets the criteria for emergency evaluation and treatment, the person shall be released. If a person was presented to the treatment facility by a police officer and was otherwise subject to lawful arrest, he shall remain under the custody of police officers.

COMMENTARY: When a person is presented to a treatment facility, the staff shall immediately "accept" him, that is, take medical responsibility for him. He shall, however, be "admitted" and confined for emergency evaluation only if the examining psychiatrist finds probable cause to believe that the person meets the commitment criteria.[147] While the Model Law creates

---

[146]*But cf.* ALASKA STAT. § 47.30.030(a) (1979) (after certification by physician, any person may take the person into custody); ARK. STAT. ANN. § 59-1406 (Supp. 1979) (any person may act in the first instance).

[147]The Model Law thus differs from state laws that provide for admission first, and for an examination only thereafter. *See, e.g.*, ALASKA STAT. § 47.30.040 (1979); COLO. REV. STAT. § 27-10-105(4) (1982); OHIO REV. CODE ANN. § 5122.10 (Page 1981). Other states take an approach similar to the Model Law. *See, e.g.*, ARIZ. REV. STAT. ANN. § 36-526 (1974); PA. STAT. ANN. tit. 50, § 7302 (Purdon 1969).

broad avenues by which a person's need for treatment may be evaluated, this independent examination requirement provides a check against unwarranted deprivations of the person's liberty. The examining psychiatrist must act as an independent decisionmaker; he must determine if probable cause exists based on his examination of the patient, and not simply based on the presenting party's allegations. This provision further requires that if a person who could have been arrested on criminal charges is not admitted to the treatment facility, he should be returned to police custody.

4.E. ***Advice of Rights.***

The treatment facility shall advise any person admitted for emergency evaluation and treatment of the purposes and possible duration of emergency evaluation and of his rights under this Act as soon after admission as his medical condition permits.

COMMENTARY: In initiating treatment of severely mentally disturbed persons, the patient's interest in learning his rights, though important, may not be paramount. Harm could result from administering a long description of legal rights to an agitated, frightened patient who, for the next few hours, needs a different kind of therapeutic interaction. Thus, the Model Law requires telling a patient his rights "as soon after admission as his medical condition permits."[148] Often this may be done shortly after admission, and in almost all cases it should occur within four to twelve hours. This advice should be viewed as a therapeutic exchange, an effort to explain to the patient what is occurring and why. The dialogue should not degenerate into a routinized incantation read from a card, like the "Miranda" warnings given to persons taken into criminal custody.

4.F. ***Hearing on Emergency Evaluation.***

1. Each person who is admitted to a treatment facility shall receive a preliminary hearing before the court within five business days of admission or be discharged, unless he has, after consultation with counsel, executed a written waiver of such hearing. The hearing shall be informal

[148] *Cf.* CONN. GEN. STAT. § 17-183(c) (1981) (patient to be informed of his rights "promptly"); ILL. REV. STAT. ch. 91 1/2, § 2-200 (1981) (immediately after admission, or "as soon thereafter as the condition of the recipient [of services] permits"); MD. HEALTH-GEN. CODE ANN. § 10-631 (1982) (within 12 hours of admission); N.M. STAT. ANN. § 43-1-10(E) (1978) ("upon arrival," with no exceptions). *See generally* Am. Hosp. Ass'n, A Patient's Bill of Rights ¶ 12 (1972), *reprinted in* K. COUNTRYMAN & A. GEKAS, DEVELOPMENT AND IMPLEMENTATION OF A PATIENT'S BILL OF RIGHTS IN HOSPITALS (1980) [hereinafter cited as A Patient's Bill of Rights].

and subject to such rules as the court sets consistent with fundamental fairness.

2. The court shall determine at the close of the hearing, or within five business days of the patient's admission, whether he should be discharged. A patient shall then be discharged, unless the court determines that there is probable cause to believe that he satisfies the criteria for thirty-day commitment provided in Section 6, and unless within two business days of the court's decision a petition for such commitment is filed with the court.

COMMENTARY: Under established law, a person confined for emergency mental evaluation "must be given a hearing within a reasonable period of time to test whether the determination is based upon probable cause to believe that confinement is necessary under constitutionally proper standards for commitment."[149] Since the Supreme Court's decision in *O'Connor v. Donaldson*,[150] the judiciary generally has required that an initial hearing be held within five to seven days after admission.[151] Recently revised state laws tend to employ similar time periods.[152]

---

[149]Lynch v. Baxley, 386 F. Supp. 378, 388 (M.D. Ala. 1974); *see also* Logan v. Arafeh, 346 F. Supp. 1265, 1268 (D. Conn. 1972), *aff'd sub nom.* Briggs v. Arafeh, 411 U.S. 911 (1973) (Emergency commitment "without prior notice and hearing does not offend the due process clause provided that [a hearing is available] within a reasonable period of time."); Anderson v. Solomon, 315 F. Supp. 1192, 1194 (D. Md. 1970) ("Due process requires that a hearing be held at some reasonable point in time before a person can be indeterminately committed.").

[150]422 U.S. 563 (1975). Prior to *Donaldson*, several courts had approved detention without a commitment hearing for periods as long as 20 to 45 days, usually on the ground that as a benevolent therapeutic process civil commitment was not subject to strict due process constraints. *See, e.g.*, Coll v. Hyland, 411 F. Supp. 905 (D.N.J. 1976); Logan v. Arafeh, 346 F. Supp. 1265 (D. Conn. 1972), *aff'd sub nom.* Briggs v. Arafeh, 411 U.S. 911 (1973). *Donaldson*, however, undercut this rationale by declaring that mental illness alone does not provide a justification for open-ended periods of confinement.

[151]*See, e.g.*, Doe v. Gallinot, 486 F. Supp. 983 (C.D. Cal. 1979) (17 days is not acceptable; 72 hours plus sufficient time to arrange a hearing, for a total of less than seven days, is permissible); Suzuki v. Quisenberry, 411 F. Supp. 1113 (S.D. Iowa 1976) (a few days); Doremus v. Farrell, 407 F. Supp. 509 (D. Neb. 1975) (five days); Bartley v. Kremens, 402 F. Supp. 1039 (E.D. Pa. 1975) (72 hours); Lynch v. Baxley, 386 F. Supp. 378 (M.D. Ala. 1974) (seven days); Bell v. Wayne County Gen. Hosp., 384 F. Supp. 1085 (E.D. Mich. 1974) (five days); Lessard v. Schmidt, 349 F. Supp. 1078 (E.D. Wis. 1972), *vacated and remanded*, 414 U.S. 473 (1974), *judgment reinstated*, 413 F. Supp. 1318 (E.D. Wis. 1976) (48 hours); State *ex rel.* Doe v. Madonna, 295 N.W.2d 356 (Minn. 1980) (72 hours).

[152]*See, e.g.*, MD. HEALTH-GEN. CODE ANN. §§ 10-620 to -628 (1982) (a hearing is required within four days if a police officer or physician initiates emergency evaluation, within one day if another person does so); N.M. STAT. ANN. § 43-1-11(A) (1978) (seven days); OHIO REV. CODE ANN. § 5122.141 (Page 1981) (three court days, or up to 10 court days for good cause shown); PA. STAT. ANN. tit. 50, § 7303 (Purdon 1969) (five days). *But see* CONN. GEN. STAT. § 17-183 (1981) (up to 15 days emergency commitment upon a physician's certificate alone); N.C. GEN. STAT. ANN. § 122.58-7(a) (1981) (10 days).

The Model Law provides that a hearing must be held within five business days.[153] A person could not be committed for even this brief period unless his behavior created probable cause to believe that he is committable and an independent psychiatric evaluation by the treatment staff indicates that he needs emergency evaluation. These findings justify the relatively short confinement period. Because many acute psychiatric episodes subside within one to four days,[154] if a hearing were required too soon, many patients would be retained unnecessarily for longer-term commitment periods. With a five-day period for emergency psychiatric treatment, however, many patients will improve sufficiently to be released.

The Model Law's "probable cause" hearing is not a full adversary procedure with formal testimony by witnesses and presentation of documentary evidence. Instead, in this informal inquiry the court should engage in dialogue with the parties to learn the key facts bearing on probable cause. Strict rules of evidence need not be applied. The court must release the person unless there is "probable cause" to believe that commitment is warranted.[155] The judge may take time to consider written materials and render his decision (provided he does so within five business days of the patient's admission).

4.G. ***Duration of Emergency Evaluation and Treatment.***

The period of emergency evaluation and treatment shall in no case exceed fourteen days.

COMMENTARY: Every involuntary patient receives a preliminary hearing within five business days of admission (a maximum of seven days); if the court finds the requisite probable cause, a petition for thirty-day commitment must be filed within two business days.[156] The patient then gets a hearing on thirty-day

---

[153]It provides that such a hearing must be held in every case, unless the respondent, after actually consulting with counsel, waives the hearing. This is stricter than the practice in states that provide a hearing only upon a patient's specific request. *See, e.g.*, N.Y. MENTAL HYG. LAW § 9-31 (McKinney 1978 & Supp. 1982–1983).

[154]*See generally* M. Lipton & G. Burnett, *Pharmacological Treatment of Schizophrenia*, in DISORDERS OF THE SCHIZOPHRENIA SYNDROME (L. Bellak ed. 1979).

[155]*See* Doe v. Gallinot, 486 F. Supp. 983, 994 (C.D. Cal. 1979) ("[A] probable cause hearing is not one where the State must defend a final judgment of [committability] with medical and psychiatric findings . . . . [It] requires only that the state show probable cause for confinement."); Stamus v. Leonhardt, 414 F. Supp. 439, 446 (S.D. Iowa 1975); *cf.* OHIO REV. CODE ANN. §§ 5122-05 to -06 (Page 1981). This is a lower standard than the "clear and convincing" evidence required in subsection 6.D for 30-day commitment.

[156]*See* subsection 4.F.

commitment within three business days (five days) of the filing of the petition.[157] Thus, the maximum period of emergency evaluation is fourteen days.[158]

Section 5. VOLUNTARY ADMISSION.

5.A. *Admission.*

1. A treatment facility may admit a person if after examining the patient a psychiatrist [or: "a physician"][159] on the staff or with privileges at the treatment facility believes the person is mentally ill and in need of hospitalization and if the person gives written consent to admission. Prior to such admission, the person shall be advised orally and given a written statement of his rights under this Act, provided that, if his condition upon admission makes such advice infeasible and the medical reasons are entered in the record, such advice may be deferred until the patient's medical condition permits, for not more than forty-eight hours. Each patient shall be asked to sign an acknowledgement that he has been so advised and has consented to voluntary admission for treatment.

2. Initial consent to voluntary admission for treatment shall be valid for sixty days. Thereafter, a patient may remain at the treatment facility for periods of up to one hundred eighty days each upon a signed consent executed after the patient has had an opportunity to consider with such persons as he wishes his need for continued hospitalization and treatment.

3. If the responsible psychiatrist [or: "the responsible physician"][160] has substantial reason to believe that a person seeking to admit himself or to consent to further hospitalization lacks capacity to make an informed decision concerning treatment, he shall obtain, in addition to the consent of the patient, the informed consent of the patient's next of kin or guardian. The responsible psychiatrist [or: "the responsible physician"][161] shall renew his effort to obtain the informed consent of the patient if the patient regains the capacity to make an informed decision concerning treatment.

COMMENTARY: The Model Law seeks to encourage knowing and voluntary admissions. A person applying for such admission must be examined by a psychiatrist to ensure that the person is

---

[157]*See* subsection 6.D.

[158]*Cf.* State *ex. rel.* Lockman v. Gerhardstein, 107 Wis. 2d 325, 320 N.W.2d 27 (1982); State v. Hungerford, 84 Wis. 2d 236, 267 N.W.2d 258 (1978) (14 days is maximum permissible period before full due process hearing); WIS. STAT. § 51.20(7)(c) (1979–1980).

[159]Optional provision. In this and other provisions, the Model Law provides an option as to whether a "psychiatrist" or a "physician" performs certain functions. The APA believes that only a trained psychiatrist should provide psychiatric treatment. Many areas of the country, however, lack psychiatrists to staff treatment facilities. Therefore, the Model Law leaves this to each state to decide.

[160]Optional provision.

[161]Optional provision.

an appropriate patient. For example, a physically violent, agitated patient might not be allowed voluntary status.[162] Ordinarily, the treatment staff next should explain to the patient the consequences of voluntary admission and his rights, and the patient could then sign a consent form.[163] A voluntary patient might, however, be so panicky, catatonic, or delusional that extracting written consent at that time would be impossible or harmful to him. Thus, if the examining physician perceives such a situation and writes the reasons in the patient's record, the Model Law grants the physician up to forty-eight hours to advise the patient of his rights and to obtain consent. This provision applies, however, only to patients who genuinely appear to be nonprotesting.

A patient must periodically renew his consent to voluntary hospitalization.[164] This requirement both ensures that the patient desires continued hospitalization and protects the treatment facility against charges that it is relying on an outdated consent. The patient also must be allowed to discuss with anyone he wishes whether he should remain hospitalized.[165] In particular, a patient's access to counsel at all should times be protected.

The admitting psychiatrist must initially judge whether the person seeking voluntary admission lacks the capacity to make an informed decision concerning hospitalization or treatment.[166] If the patient lacks such capacity and admission is advisable, then, as an ethical matter, the next best available consent to voluntary admission—that of the next of kin or guardian—also should be obtained.[167] The only alternative would be to refuse

[162]*Cf.* CONN. GEN. STAT. § 17-187(b) (1981) (permitting voluntary status "if the superintendent deems such person clinically suitable for such admission"). Also, the general provision concerning voluntary admissions should not be construed to preclude "informal admissions," where, in some states, an individual may enter a treatment facility voluntarily without a formal record being kept. *See, e.g.*, ILL. REV. STAT. ch. 91 1/2, § 3-300 (1981) (psychiatrist must decide that the "person [is] clinically suitable for admission upon an informal basis"); N.Y. MENTAL HYG. LAW §§ 9.15, 9.17 (McKinney 1978).

[163]Many writers have noted the dubieties inherent in "voluntary" admissions. *See, e.g.*, Gilboy & Schmidt, *"Voluntary" Hospitalization of the Mentally Ill*, 66 NW. U.L. REV. 429 (1971); Olin & Olin, *supra* note 69; Palmer & Wohl, *supra* note 69. Nevertheless, treatment staff should strive to make the process of obtaining informed consent a meaningful interaction so that voluntary admission occurs only when, given the patient's limitations, that genuinely appears to be his wish.

[164]*See* paragraph 5.A.2; *cf.* ILL. REV. STAT. ch. 91 1/2, § 3-404 (1981) (requires reaffirmation of voluntary status after 30 days, and then every 60 days thereafter); N.Y. MENTAL HYG. LAW § 9.25 (McKinney 1978) (requires yearly review of the patient's consent and of his suitability for treatment).

[165]*See* paragraph 5.A.2.

[166]*See* Section 3 (defining "lacks capacity to make an informed decision concerning treatment").

[167]A major issue that we only note here is the "good faith" of the substituted consent

voluntary admissions and to require that the patient be involuntarily committed.[168] The Model Law rejects this approach. It recognizes that patients do not fit neatly into "voluntary" and "involuntary" categories. Many are nonprotesting, ambivalent, or confused. Because there may be therapeutic value to being a voluntary patient, a patient who lacks capacity and appears to be genuinely nonprotesting should be permitted voluntary status if his family appears to be acting in his interests in giving substituted consent. The psychiatrist must renew efforts to obtain the patient's consent if he appears to regain his capacity for reasoned decisionmaking.

5.B. ***Discharge or Petition for Thirty-Day Commitment.***

Any patient who is voluntarily admitted to a treatment facility shall be discharged within five business days of his written request for discharge (and any patient who indicates his desire to be discharged but is unable to write shall be helped to put his request in writing), unless a petition for thirty-day commitment is filed within that period by the treatment facility or the patient's next of kin or guardian.

COMMENTARY: A truly voluntary patient has the right to be discharged upon request. For valid medical reasons, however, such discharge may not be immediately appropriate. Medication may have to be adjusted; arrangements may need to be made with family members or alternative care facilities; or the treatment facility may believe that the person should continue to be confined—on an involuntary basis. An imminently suicidal patient, for example, should not be released. The Model Law thus requires that a voluntary patient be discharged within five business days of his request, unless a commitment petition is filed.[169]

---

by the next of kin or guardian. By referring to the "next of kin or guardian," the Model Law merely incorporates by reference whatever relationships and procedures for substituted consent the state law currently recognizes. In the mental health context, next of kin are not always objective decisionmakers who advance only the patient's interests. Some will have hostile or otherwise charged motives for committing family members. In other cases, they may be benevolent, but the psychological toll of a disturbed household member may have pushed them to a decision that is not in the patient's best interests. The examining psychiatrist and the court should, therefore, be vigilant to ensure that voluntary admission on the basis of substituted consent is really in the patient's best interests. Where sharply conflicting motives are present, the psychiatrist may decline to accept the substituted consent of the next of kin or guardian, and seek court appointment of another decisionmaker for the patient.

[168]Some states permit voluntary admissions only of competent patients. *See, e.g.*, Pima County v. Superior Ct., 26 Ariz. App. 85, 546 P.2d 354 (1976); *In re* Hop, 29 Cal. 3d 20, 623 P.2d 282, 171 Cal. Rptr. 652 (1981); MD. HEALTH-GEN. CODE ANN. §§ 10-603, 609 (1982). *But see* MASS. GEN. LAWS ANN. ch. 123, § 10(a) (West Supp. 1982).

[169]*Cf.* ARIZ. REV. STAT. ANN. § 36-519 (1974) (release required within one day, plus weekends and holidays); CONN. GEN. STAT. § 17-187(a) (1981) (five days plus weekends); DEL. CODE ANN. tit. 16, § 5123(e) (1975) (five days); MASS. GEN. LAWS ANN.

The request for discharge must be in writing so that the treatment facility need not initiate extensive discharge procedures whenever a patient's verbal statement might be construed to mean that he wants to leave. If a patient clearly states a desire to be discharged, he should immediately be assisted to put it in writing. Either the treatment facility or the next of kin or guardian may file the petition for thirty-day commitment. The treatment facility knows the patient's recent mental condition and behavior. The family, however, by virtue of prior experiences with the patient, recent visits, and other information, may perceive aspects of his condition that the treatment staff do not detect, and may identify other valid reasons why the patient should not be returned home. If a petition for involuntary commitment is filed, then the patient can be confined until the hearing, which under paragraph 6.D.1 must occur within three business days.

5.C. ***Conversion from Involuntary to Voluntary Status.***

A patient who is subject to involuntary hospitalization pursuant to Sections 4, 6, or 11 of this Act may at any time convert to voluntary status if the responsible psychiatrist [or: "the responsible physician"][170] agrees that such conversion is made in good faith and that the patient is an appropriate patient for voluntary hospitalization.

COMMENTARY: The Model Law encourages involuntary patients to convert to voluntary status. Psychiatrists generally agree that therapeutic efforts are probably enhanced when the patient voluntarily cooperates in the treatment program.[171] Because some patients may seek conversion simply in order to request immediate discharge, a psychiatrist must approve each conversion by confirming that it appears to be sought in good faith and that the patient is appropriate for voluntary status. The latter criterion might, for example, exclude placing a homicidal patient on a ward with depressed voluntary patients.[172]

---

ch. 123, § 11 (West Supp. 1982) (three days); MONT. CODE ANN. § 53-21-111 (1981) (five days); N.Y. MENTAL HYG. LAW § 9.13 (McKinney 1978 & Supp. 1982–1983) (three days); OHIO REV. CODE ANN. § 5122.03 (Page 1981) (three "court days"); PA. STAT. ANN. tit. 50, § 7703 (Purdon 1969) (three days); TENN. CODE ANN. § 33-601 (1977) (two days); UTAH CODE ANN. § 64-7-31 (1953) (two days plus weekends). Only two states permit a period of more than five days. *See* OKLA. STAT. tit. 43A, § 53 (1981) (15 days); S.C. CODE ANN. § 44-17-330 (Law. Co-op. 1976) (15 days). In light of these laws, the Model Law's "five business days" is probably a bit too long.

[170]Optional provision.

[171]*See generally* T. GUTHEIL & P. APPELBAUM, HANDBOOK OF PSYCHIATRY AND LAW 48–49 (1982).

[172]This provision accords with statutes and court decisions in states that have ad-

Section 6. THIRTY-DAY COMMITMENT.

6.A. *Petition.*

1. Persons who are present at a treatment facility under voluntary admission but have requested discharge, and persons present at a treatment facility for emergency psychiatric evaluation, may be committed involuntarily for a period of up to thirty days upon a petition filed by the treatment facility or by the next of kin or guardian; and other persons may be so committed upon a petition filed by any interested adult. The petition shall allege that such person meets the criteria set forth in subsection 6.C. The petition shall set forth the facts supporting the allegations and, in the case of petitions filed by a treatment facility, shall describe why the patient requires treatment. The petition shall be filed with the court, which shall have copies promptly served upon the patient, the next of kin or guardian, and the patient's attorney if known.

2. The copies of the petition served by the court shall be accompanied by a notice advising of the person's rights concerning the proceeding.

COMMENTARY: The process leading to thirty-day commitment begins with the filing of a petition. If a patient is currently hospitalized, either the treatment facility, which knows his present condition, or his next of kin or guardian, may file the petition. As to a person who is not currently hospitalized, any interested person can file a petition.[173] Petitions filed by the treatment facility must include, in addition to the other information specified, a statement identifying the treatment to be provided.[174] This ensures that commitment will serve some significant therapeutic purpose. To protect individual liberty, subsection 18.B severely sanctions submission of frivolous or fraudulent petitions. The court must serve the petition on all appropriate parties.[175]

6.B. *Summons for Evaluation; Psychiatric Report.*

1. Upon the filing of a petition for thirty-day commitment of a person who is not currently under emergency evaluation or voluntary admission, the court shall issue a summons to the person to submit to an examination

---

dressed the issue. *See, e.g.*, Von Luce v. Rankin, 267 Ark. 34, 588 S.W.2d 445 (1979); *In re* Paiz, 603 P.2d 976 (Colo. Ct. App. 1979); *In re* Byrd, 68 Ill. App. 3d 849, 386 N.E.2d 385 (1979); ARK. STAT. ANN. § 59-1412 (Supp. 1979); N.Y. MENTAL HYG. LAW § 9.23 (McKinney 1978).

[173]*Cf.* ALASKA STAT. § 47.30.020 (1979) ("an interested party"); ARIZ. REV. STAT. ANN. § 36-520 (1974) ("any responsible individual"); ARK. STAT. ANN. § 59-1404 (Supp. 1979) ("any person"); CAL. WELF. & INST. CODE § 5201 (West 1972) ("any individual"); COLO. REV. STAT. § 27-10-106(2) (1982) ("any individual"); OHIO REV. CODE ANN. § 5122.11 (Page 1981) ("any person").

[174]The petition also should specify all persons who uniquely know the facts. *See, e.g.*, ARK. STAT. ANN. § 59-1404 (Supp. 1979).

[175]*See* Moss v. State, 539 S.W.2d 936 (Tex. Civ. App. 1976).

(on an outpatient basis) conducted by a psychiatrist at a treatment facility or a private psychiatrist. The examining psychiatrist shall promptly prepare a report on his examination and file it with the court. The court shall have copies promptly served upon the patient, the next of kin or guardian, and the patient's attorney if known.

2. A person served with a summons to submit to a psychiatric examination may in lieu of such examination submit a report of a psychiatrist stating that he has recently examined the person or has ongoing medical responsibility for the person and knowledge of his current condition, and that in his opinion the person does not meet the criteria for involuntary commitment. The petition for commitment may then be dismissed by the court or continued.

COMMENTARY: To ensure adequate information at the thirty-day commitment hearing, the subject of a commitment petition must submit to a psychiatric examination.[176] The state must pay for the examination and may arrange for it to be conducted by a psychiatrist in a treatment facility or by a private psychiatrist. To protect against unwarranted, meddlesome, or ill-motivated petitions, however, the subject may instead be examined by a psychiatrist of his own choosing.[177] Ordinarily the judge should dismiss a petition if the psychiatrist reports the individual to be noncommitable. If the report clearly lacks credibility, however, the judge may continue the petition.

6.C. *Criteria for Thirty-Day Commitment.*

A person may be involuntarily committed for a period of up to thirty days if, after the hearing provided in subsection 6.D., the court determines on the basis of clear and convincing evidence that:

1. the person is suffering from a severe mental disorder; and

2. there is a reasonable prospect that his disorder is treatable at or through the facility to which he is to be committed, and such commitment would be consistent with the least restrictive alternative principle; and

3. the person either refuses or is unable to consent to voluntary admission for treatment; and

4. the person lacks capacity to make an informed decision concerning treatment; and

5. as the result of the severe mental disorder, the person is (a) likely to cause harm to himself or to suffer substantial mental or physical deterioration, or (b) likely to cause harm to others.

---

[176]*Cf.* ALASKA STAT. § 47.30.070(c) (1979); ARIZ. REV. STAT. ANN. 36-520-522 (1974); ILL. REV. STAT. ch. 91 1/2, § 3-703 (1981); OHIO REV. CODE ANN. § 5122.14 (Page 1981); UTAH CODE ANN. § 64-7-36 (1953). *But see* PA. STAT. ANN. tit. 50, § 7304 (Purdon 1969) (examination only if requested by the respondent).

[177]This provision creates a possibility of persons "shopping" for a favorable psychiatric opinion. There does not appear to be any way to preclude this entirely. Professional ethics and the possibility of legal liability, however, make it likely that virtually all psychiatrists will undertake such evaluations in a responsible fashion.

COMMENTARY: This subsection, which is the heart of the Model Law, sets forth criteria for thirty-day commitment and for subsequent recommitments.[178] The criteria must be established by "clear and convincing evidence".[179] The Supreme Court held in *Addington v. Texas* that the Constitution requires a standard at least this strict to be met before a person is deprived of his freedom through civil commitment.[180] Many states, in fact, employ a still higher standard of proof.[181]

The Model Law adopts the "clear and convincing evidence" standard because it best balances society's interest in being protected with the individual's interest in liberty, because a higher standard would be impractical when proving the inherently elusive and judgmental elements of mental illness and likelihood of harm,[182] and because most states employ a com-

---

[178]*See* subsection 11.A (60-day recommitment) and subsection 11.B (180-day recommitments). For an extensive discussion of recent trends in state laws on the criteria for commitment, see Stromberg, *supra* note 6, at 334.

[179]The standard of proof allocates risks of error between litigants, and instructs "the fact finder concerning the degree of confidence our society thinks he should have in the correctness of factual conclusions for a particular type of adjudication." *In re* Winship, 397 U.S. 358, 370 (1970) (Harlan, J., concurring). It is generally believed that the "preponderance of the evidence" standard used in civil trials represents 51% certainty, and the "beyond a reasonable doubt" standard used in criminal cases represents about 90% or 95% certainty. Clear and convincing evidence would then mean something like 75% certainty. *Cf. In re* Sedillo, 84 N.M. 10, 12, 498 P.2d 1353, 1355 (1972) ("For evidence to be clear and convincing, it must instantly tilt the scales in the affirmative when weighed against the evidence in opposition and the fact finder's mind is left with an abiding conviction that the evidence is true."). *See generally* United States v. Fatico, 458 F. Supp. 388, 410 (E.D.N.Y. 1978); McCauliff, *Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?*, 35 VAND. L. REV. 1293, 1324–29 (1982).

[180]441 U.S. 418 (1978). Although the Court conceded that no empirical studies compared how the various standards of proof were actually applied, it nevertheless concluded that "[i]ncreasing the burden of proof is one way to impress the fact finder with the importance of the decision and thereby perhaps reduce the chances that inappropriate commitments will be ordered." *Id.* at 427.

[181]*See, e.g.*, Superintendent of Worcester State Hosp. v. Hagberg, 374 Mass. 271, 372 N.E.2d 242 (1978); Proctor v. Butler, 117 N.H. 927, 380 A.2d 673 (1977); HAWAII REV. STAT. § 334-60(b)(4)(I) (1976 & Supp. 1981); *cf.* Estate of Roulet, 23 Cal. 3d 219, 590 P.2d 1, 152 Cal. Rptr. 425 (1979).

[182]

> [T]he initial inquiry in a civil commitment proceeding is very different from the central issue in . . . a criminal prosecution. In the latter [case] the basic issue is a straight-forward factual question—did the accused commit the act alleged? . . . Whether the individual is mentally ill and dangerous to either himself or others and is in need of confined therapy turns on the *meaning* of the facts which must be interpreted by expert psychiatrists and psychologists. *Given the lack of certainty and the fallibility of psychiatric diagnosis, there is a serious question as to whether a state could ever prove beyond a reasonable doubt that an individual is both mentally ill and likely to be dangerous.*

Addington v. Texas, 441 U.S. at 429 (citations omitted) (emphasis added); *see also* Bartley v. Kremens, 402 F. Supp. 1038, 1052–53 (E.D. Pa. 1975), *vacated*, 431 U.S. 119

parable standard.[183]

All five of the commitment criteria set forth in subsection 6.C must be satisfied to justify involuntary commitment. The first criterion requires that the person is suffering from a "severe mental disorder."

The second criterion has two parts. First, the petitioner must establish a reasonable expectation that the person's condition is treatable. A few states have moved to make treatability a criterion of commitment,[184] but most have not. The Model Law reflects the view that "[w]ithout some form of treatment, the state's justification for acting as parens patriae becomes a nullity."[185] Even under the police power, if the mental health system offers only preventive confinement, then the criminal system should take responsibility for the person.[186] The petitioner need not prove that treatment will cure the patient, but appropriate therapy must be available—and not just in theory, but "at" or "through" (that is, arranged by) the facility where the patient will be placed.

Second, the commitment must be consistent with the "least restrictive alternative principle." About twenty states impose a similar statutory requirement;[187] about fifteen more identify the principle as a factor to be considered.[188] At present, however,

---

(1977), *reaff'd sub nom.* Institutionalized Juveniles v. Secretary of Pub. Welfare, 459 F. Supp. 30 (E.D. Pa. 1978) (clear and convincing evidence is the highest standard that is realistic in light of the "relatively undeveloped state of psychiatry as a predictive science"), *rev'd and remanded*, 442 U.S. 640 (1979); Lynch v. Baxley, 386 F. Supp. 378, 393 n.12 (M.D. Ala. 1974); People v. Taylor, 618 P.2d 1127, 1136 (Colo. 1980); State *ex rel.* Hawks v. Lazaro, 157 W. Va. 417, 444, 202 S.E.2d 109, 126–27 (1974).

[183]*See* Addington v. Texas, 441 U.S. at 431–32 nn.6–8 (citing statutes). See also cases since *Addington*, such as State *ex rel.* Doe v. Madonna, 295 N.W.2d 356, 363 n.11 (Minn. 1980); *In re* Carter, 102 Misc. 2d 867, 424 N.Y.S.2d 833 (Sup. Ct. 1980).

[184]*See* N.M. STAT. ANN. § 43-1-11(C)(2) (1978); OHIO REV. CODE ANN. § 5122.01(B)(4) (Page 1981). *But see In re* Oseing, 296 N.W.2d 797, 799 (Iowa 1980) (treatability need not be shown in initial commitment of dangerous person).

[185]*In re* Ballay, 482 F.2d 648, 659 (D.C. Cir. 1973); *cf.* Jackson v. Indiana, 406 U.S. 715 (1972) (invalidating an open-ended confinement of a nontreatable, retarded, deaf-mute who was—and would probably always remain—incompetent to stand trial).

[186]*See infra* note 201.

[187]*See, e.g.*, ALA. CODE § 22-52-10(5) (Supp. 1982); CONN. GEN. STAT. § 17-178(c) (1981); MD. HEALTH-GEN. CODE ANN. § 10-617(a)(4) (1982); N.M. STAT. ANN. § 43-1-11(c)(3) (1978); N.D. CENT. CODE § 25-03.1-21 (Supp. 1981); UTAH CODE ANN. § 64-7-36(10)(d) (Supp. 1981); *cf.* IND. CODE § 16-14-9.1-10(d) (Supp. 1982) (requiring alternatives to hospitalization without explicitly adopting the least restrictive alternative approach); MICH. COMP. LAWS §§ 330.1441 - .1469 (1979) (same).

[188]*See, e.g.*, Eubanks v. Clarke, 434 F. Supp. 1022, 1027–28 (E.D. Pa. 1977); Stamus v. Leonhardt, 413 F. Supp. 439, 452–53 (S.D. Iowa 1976); Lynch v. Baxley, 386 F. Supp. 378, 392 (M.D. Ala. 1974); Davis v. Watkins, 384 F. Supp. 1196, 1203 (N.D. Ohio 1974); Welsch v. Likins, 373 F. Supp. 487, 501–02 (D. Minn. 1974), *vacated in part on other grounds*, 550 F.2d 1122 (8th Cir. 1977); Lessard v. Schmidt, 349 F. Supp. 1078, 1095–

this requirement is more aspirational than real. As one study concluded, "[t]he least restrictive alternative is a nice phrase—unfortunately, few if any exist."[189] Nevertheless, the Model Law includes this criterion so that judges—and legislators—will confront the need for appropriate placement alternatives.[190]

The third criterion for involuntary commitment is that the person is unwilling to consent to voluntary treatment.[191] Although the threat of involuntary commitment doubtless affects the "voluntariness" of some admissions, the practical and therapeutic advantages to voluntary admission make consent to admission a meaningful patient right.[192]

The fourth criterion requires that a committed person "lacks capacity to make an informed decision concerning treatment." Including this criterion in the Model Law represents a major policy judgment. Although a few recently enacted state statutes seem to require lack of capacity for some commitments,[193] and certain courts have enunciated similar principles,[194] the Model Law's across-the-board requirement is somewhat novel.

---

96 (E.D. Wis. 1972), *vacated on other grounds*, 414 U.S. 473 (1974); *In re* Bumper, 441 A.2d 975, 977 (D.C. 1982).

[189]Hoffman & Foust, *supra* note 32, at 1126 (1977).

[190]For an account of Michigan's extensive efforts to ensure meaningful evaluation of alternative placements, see Note, *Community Treatment Case Study*, *supra* note 38.

[191]This accords with statutes in several states. *See, e.g.*, CAL. WELF. & INST. CODE §§ 5250(b), 5260(b) (West 1969); COLO. REV. STAT. § 27-10-107(1)(b) (1982); CONN. GEN. STAT. § 17-178(e) (1981); MD. HEALTH-GEN. CODE ANN. § 10-617(a)(3) (1982); MASS. GEN. LAWS ANN. ch. 123, § 12(e) (West Supp. 1982–1983); OR. REV. STAT. § 426.130 (1981). Some other states enforce this requirement as a matter of judicial practice. *See, e.g.*, People *ex rel.* Henderson, 44 Colo. App. 102, 610 P.2d 1350 (1980); People v. Hill, 72 Ill. App. 3d 638, 391 N.E.2d 51 (1979); *In re* Byrd, 68 Ill. App. 3d 849, 854, 386 N.E.2d 385, 388 (1979); *In re* Farrow, 41 N.C. App. 680, 686–87, 255 S.E.2d 777, 782 (1979).

[192]*See, e.g.*, Sisneros v. District Court, 199 Colo. 179, 181–82, 606 P.2d 55, 57 (1980); Goedecke v. State Dep't of Institutions, 198 Colo. 407, 411, 603 P.2d 123, 125 (1979). A court can, however, refuse to approve "voluntary" admission if it is simply a ruse to permit the patient to be discharged upon his request at any time. *See In re* Byrd, 68 Ill. App. 3d 849, 854, 386 N.E.2d 385, 389 (1979).

[193]*See, e.g.*, COLO. REV. STAT. § 27-10-102(5) (1982) ("gravely disabled" means that the person is unable to care for his basic needs and "lacks the capacity to understand [that] this is so"); HAWAII REV. STAT. § 334-1 (1976) ("dangerous to self" means likely to injure oneself "together with incompetence to determine whether treatment . . . is appropriate"); KAN. STAT. ANN. § 59-2902(1)(a) (Supp. 1977) ("lacks sufficient understanding or capacity to make responsible decisions with respect to his or her need for treatment"); UTAH CODE ANN. § 64-7-36(10)(c) (Supp. 1981) ("lacks the ability to engage in a rational decisionmaking process . . . as demonstrated by evidence of inability to weigh the possible costs and benefits of treatment"); WYO. STAT. ANN. § 25-10-101(viii) (Michie 1982) (mentally ill person is one who, inter alia, "cannot comprehend the need for or purposes of treatment").

[194]Colyar v. Third Judicial Dist. Court, 469 F. Supp. 424, 434 (D. Utah 1979) (only persons "incapable of making a rational treatment decision" may be committed under the parens patriae doctrine); Lynch v. Baxley, 386 F. Supp. 387, 391 (M.D. Ala. 1974);

We believe that this requirement is warranted for several reasons. Parens patriae commitments cannot be justified by mental illness unless the person also lacks capacity. Many mental disorders do not impair a person's ability to assess his desire for hospitalization or a particular therapy.[195] Capacity is not unitary; it reflects a set of contextual competences varying with the choices posed and other circumstances.[196] If a competent person refuses treatment, by what social ethic can the state substitute another judgment? This has never been answered successfully, and so the Model Law does not permit such substituted judgment. In addition, the Model Law would limit police power commitments to those who lack capacity. The criminal justice system should bear the primary responsibility for protecting society against dangerous but competent persons; such people do not usually belong in mental hospitals.

Moreover, unless lack of capacity is made a threshold criterion of all involuntary commitments, there is no way to decide which patients retain the right to refuse treatment.[197] Several courts have held that an involuntarily committed patient retained such a right precisely because the statute under which he was committed did not require that he lack capacity.[198] In this situation, the treatment facility could neither release nor treat the patient. Although the empirical evidence is sketchy, it suggests that half to two-thirds of those persons now committed would be found to lack capacity to make informed decisions concerning treatment.[199]

Lessard v. Schmidt, 349 F. Supp. 1078, 1094 (E.D. Wis. 1972), *vacated on other grounds*, 414 U.S. 473 (1974); Commonwealth *ex rel.* Finken v. Roop, 233 Pa. Super. 762, 339 A.2d 764, 776–77 (1975).

[195]C. LIDZ, A. MEISEL, E. ZERUBAVEL, M. ASHLEY, R. SESTAK & L. ROTH, INFORMED CONSENT? AN EMPIRICAL STUDY OF DECISION MAKING IN PSYCHIATRY (1983).

[196]*Id.*

[197]In theory, one could argue that commitment on the basis of mental illness and dangerous acts alone ought to justify a certain period of involuntary medication or treatment, but several courts have rejected this view. *See* subsection 8.B and accompanying discussion; *see also* Stone, *The Right to Refuse Treatment: Why Psychiatrists Should and Can Make It Work*, 38 ARCHIVES GEN. PSYCHIATRY 358 (1981).

[198]*See, e.g.*, Rogers v. Okin, 478 F. Supp. 1342, 1361 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982); Rennie v. Klein, 462 F. Supp. 1131 (D.N.J. 1978), *supplemented*, 476 F. Supp. 1294 (D.N.J. 1979), *aff'd and modified*, 653 F.2d 836 (3d Cir. 1981).

[199]*See* Grossman & Summers, *A Study of the Capacity of Schizophrenic Patients to Give Informed Consent*, 31 HOSP. & COMMUNITY PSYCHIATRY 205 (1980); Roth & Appelbaum, *What We Do and Do Not Know About Treatment Refusals in Mental Institutions*, in REFUSING TREATMENT IN MENTAL HEALTH INSTITUTIONS: VALUES IN CONFLICT 179, 187–88 (A. Doudera & J. Swazey ed. 1982).

Fifth, the Model Law requires that an individual be "likely to cause harm to himself or to suffer substantial mental or physical deterioration," or be "likely to cause harm to others," *as the result of* a severe mental disorder. For example, a non-mentally ill person who plans to commence a hunger strike to make a political statement would not meet this criterion, because any likely harm would be due to a reasoned decision, not a severe mental disorder.

In sum, the criteria in the Model Law are in some respects broader than those in some current state laws (for example in providing for commitment of persons whose mental state is likely to deteriorate), while in other respects the criteria are stricter (for example in permitting commitment only of persons who lack capacity). We believe that relatively few persons who would now be committable as dangerous would not be committable under the Model Law,[200] and that many severely disordered people who are not now committable as gravely disabled could be committed under the Model Law. But however the criteria for commitment were defined, it would be necessary to come to grips with what happens to those who do not meet the criteria. This problem involves a host of troubling practical, social, and ethical issues, which we do not believe the Model Law—or any state law—can itself solve. Nevertheless, we discuss in the margin the disposition of various classes of people who would fail to meet one or more of the Model Law's commitment criteria.[201]

---

[200]The authors have previously articulated a set of commitment criteria that, although different in some important respects, are similar to the set proposed by the Model Law. *See* A. STONE (WITH C. STROMBERG), *supra* note 2, at 66–79. One study concluded that 86% of the patients committed under California's "dangerousness" standard also would be committable under these criteria. *See* Monahan, Ruggiero & Friedlander, *Stone-Roth Model of Civil Commitment and the California Dangerousness Standard: Operational Comparison*, 39 ARCHIVES GEN. PSYCHIATRY 1267 (1982).

[201]*Some Major Categories of Persons Who Do Not Meet the Criteria for Involuntary Commitment.*

1. *Does NOT suffer from a "severe mental disorder."* Such persons should not be committable, because involuntary hospitalization has less to offer them, and because permitting commitment based on any lower standard of mental illness permits too much ambiguity and discretion. Although some patients who do not suffer from a "severe mental disorder" could benefit from treatment, use of this standard will help to prevent abuses of civil commitment. A person who is likely to injure himself or others for reasons other than a "severe mental disorder"—such as hostile feelings, a love of danger, or some political purpose—cannot be subjected to preventive confinement on that ground alone. Serious mental impairment, not simply social deviance, should trigger the commitment process.

2. *Does NOT Lack Capacity to Make a Reasoned Decision Concerning Treatment.* Improving institutional conditions, enforcing patients' rights, and establishing a thera-

6.D. ***Hearing on Thirty-Day Commitment.***

COMMENTARY: This subsection describes a patient's rights at his thirty-day commitment hearing, as well as in the subsequent hearings provided for in Section 11. The hearing on thirty-day commitment is mandatory and may not be waived. This ensures that a neutral party will independently evaluate the need for hospitalization. In this respect, the Model Law differs from several states' laws that authorize such a hearing only if the respondent requests it.[202]

6.D.1. Every person as to whom a petition for thirty-day commitment has been filed shall be notified by the court sufficiently in advance to be able to prepare for the hearing and shall receive a prompt hearing. For persons confined for emergency psychiatric evaluation or currently under voluntary admission, this hearing shall take place within three business days of the filing of the petition.

COMMENTARY: This paragraph provides for notice so that the respondent can prepare for the hearing by arranging to be examined by his own psychiatrist and by conferring with counsel.[203]

---

peutic relationship all may encourage voluntary admissions. But if a patient who has the capacity to make an informed decision concerning his need for hospitalization or treatment still refuses to accept treatment voluntarily, there is no legal or moral ground for confining him "for his own good." If the police power goal of preventive detention is the only justification, then the person should be processed through the criminal system and afforded all its due process protections. Otherwise, criminal sanctions for proven past conduct are simply replaced by civil sanctions based on unreliable predictions of future conduct. Moreover, the fact that psychiatric treatment may be indicated does not mean that a mental hospital is required. If a person has committed physical acts sufficient to provide a solid basis for predicting future harm, he probably can be prosecuted criminally. If he has not committed such an offense, a serious question arises as to how we are so sure he is dangerous. In sum, the mental health system cannot solve all problems, and the problem of the dangerous but competent patient who needs treatment is one that other social institutions must address.

3. *Is NOT Likely to Cause Harm to Himself or Others or to Deteriorate.* If a person is not likely to harm himself or others, the law traditionally has taken the view that there is no justification for involuntary confinement. In practice, however, many people have been deemed "likely to harm themselves" because of their general mental condition. This gave rise to statutes, which have since been upheld as constitutional, explicitly providing for commitment on the basis of "grave disability." The Model Law does so as well, and adds the alternative of imminent mental or physical deterioration due to a severe mental disorder. If a person does not meet even this standard, he should not be committed.

4. *Is NOT Treatable.* If there is not even a "reasonable prospect" that a person may be treatable, confining him "for treatment" on a parens patriae basis is a charade. Alternatively, if confinement is to be permitted simply for police power preventive detention, the society should candidly and soberly make that decision through new provisions of law, which would then be subject to direct constitutional tests in the courts. But it should not surreptitiously reach the same result by pretending that it is engaging in therapy. The Model Law rejects mere custodial confinement without treatment as a justification for civil commitment.

[202]*See, e.g.*, CONN. GEN. STAT. § 17-183(d) (Supp. 1982).

[203]Due process requires fair notice. *See In re* Gault, 387 U.S. 1, 33 (1967); Stamus v.

Persons not currently confined, that is, whose condition has not required emergency evaluation, are entitled to a "prompt" hearing. Ordinarily, this means within five to ten days,[204] although the court may permit an adjournment upon the request of the respondent if appropriate. To protect persons currently confined, however, the Model Law seeks to minimize the delay before their rights are formally adjudicated. Thus, they must receive a hearing within three business days of the filing of the petition.

6.D.2. The respondent shall be present at the hearing unless the court finds (a) that he has knowingly and voluntarily waived such right after consulting with counsel, or (b) that because his behavior at the hearing is so disruptive, it cannot reasonably continue in his presence. Hearings shall be held in the treatment facility whenever feasible given the other functions of the court.

COMMENTARY: The Model Law protects the important right to be present at one's commitment hearing.[205] Only two exceptions are allowed. First, a patient may knowingly and voluntarily waive his right, after actually consulting with counsel. This exception should be narrowly construed,[206] because presence at the hearing is a historically valued safeguard against arbitrary decisions. But if a patient believes that the hearing may cause him unwanted trauma, he should not be forced to attend. Second, the patient may not disrupt the proceedings so seriously that they cannot reasonably continue. A mere prediction of disruptive behavior does not justify excluding the respondent. The patient must be given the chance to attend, and may be excluded only if he becomes so disruptive that the court deems

---

Leonhardt, 414 F. Supp. 439, 446 (S.D. Iowa 1976); Suzuki v. Quisenberry, 411 F. Supp. 1113, 1127 (D. Hawaii 1976); Lynch v. Baxley, 386 F. Supp. 378, 388 (M.D. Ala. 1974).

[204] *Cf.* CONN. GEN. STAT. § 17-178(a) (1981) (ten days); ILL. REV. STAT. ch. 91 1/2, § 3-706 (1981) (five days plus weekends); N.Y. MENTAL HYG. LAW. § 9.31(c) (McKinney 1978) (five days); OHIO REV. CODE ANN. § 5122.141 (Baldwin 1980) (30 to 45 days from initial emergency confinement); UTAH CODE ANN. § 64-7-36(8) (Supp. 1981) (ten court days).

[205] *See* Suzuki v. Quisenberry, 411 F. Supp. 1113, 1129 (D. Hawaii 1976); Lynch v. Baxley, 386 F. Supp. 378, 388 (M.D. Ala. 1974); Bell v. Wayne County Gen. Hosp., 384 F. Supp. 1085, 1094 (E.D. Mich. 1974); *In re* Watson, 91 Cal. App. 3d 455, 460, 154 Cal. Rptr. 151, 155 (1979); ARK. STAT. ANN. § 59-1408 (1979); CONN. GEN. STAT. § 17-178(f) (1981); ILL. REV. STAT. ch. 91 1/2, § 3-806 (1981); OHIO REV. CODE ANN. § 5122.15(A)(2) (Baldwin Supp. 1981).

[206] *See* French v. Blackburn, 428 F. Supp. 1351, 1357–58 (M.D.N.C. 1977), *aff'd mem.*, 443 U.S. 901 (1979); Stamus v. Leonhardt, 414 F. Supp. 439, 447 (S.D. Iowa 1976); Doremus v. Farrell, 407 F. Supp. 509, 515 (D. Neb. 1975); Bartley v. Kremens, 402 F. Supp. 1038, 1051 (E.D. Pa. 1975), *vacated*, 431 U.S. 119 (1977), *reaff'd sub nom.* Institutionalized Juveniles v. Secretary of Pub. Welfare, 459 F. Supp. 30 (E.D. Pa. 1978), *rev'd and remanded*, 442 U.S. 640 (1979). *See generally* Johnson v. Zerbst, 304 U.S. 458 (1938); N.M. STAT. ANN. § 43-1-11(B) (1978).

him to have waived his right.[207] Finally, this provision encourages courts to hold hearings concerning inpatients at the treatment facilities. This minimizes disruption to patients and to treatment staff alike.[208]

6.D.3. Any respondent who is unable to pay for counsel shall have the right to be provided with counsel to prepare for and represent him at the hearing. [Any respondent who is unable to pay for an examination for purposes of the hearing shall have the right to be provided with one examination by a licensed psychiatrist, at the expense of the (state or local government).][209]

COMMENTARY: This paragraph preserves the constitutional right to counsel at civil commitment hearings and the right of indigents to appointed counsel.[210] It also requires that counsel be provided far enough in advance to help prepare for the hearing.[211] The Model Law leaves decisions about waivers of this right and pro se representation to the court's discretion. A court should be reluctant to find such a waiver, and be careful to ensure that the respondent can effectively present his case.[212]

---

[207]This is similar to the procedure in criminal cases. *See* Illinois v. Allen, 397 U.S. 337, 342 (1962).

[208]*Cf.* CONN. GEN. STAT. § 17-177(a) (1981).

[209]Optional provision.

[210]*See, e.g.*, Sarzen v. Gaughan, 489 F.2d 1076, 1085-(1st Cir. 1973); Stamus v. Leonhardt, 414 F. Supp. 439, 448 (S.D. Iowa 1976); Suzuki v. Quisenberry, 411 F. Supp. 1113, 1129 (D. Hawaii 1976); Lynch v. Baxley, 386 F. Supp. 378, 389 (M.D. Ala. 1974); Bell v. Wayne County Gen. Hosp., 384 F. Supp. 1085, 1092–93 (E.D. Mich. 1974); Lessard v. Schmidt, 349 F. Supp. 1078, 1097 (E.D. Wis. 1972), *vacated on other grounds*, 414 U.S. 473 (1974); *In re* Tuntland, 71 Ill. App. 3d 523, 390 N.E.2d 11 (1979); *In re* Collman, 9 Or. App. 476, 482–84, 497 P.2d 1233, 1236–37 (1972); Quesnell v. State, 83 Wash. 2d 224, 233–35, 517 P.2d 568, 574–75 (1974); State *ex rel.* Hawks v. Lazaro, 157 W. Va. 417, 439–40, 202 S.E.2d 109, 124 (1974); *see also* CONN. GEN. STAT. § 17-178(a) (1981); ILL. REV. STAT. ch. 91 1/2, § 3-805 (1981); MICH. COMP. LAWS § 330.1454(2) (1979); OHIO REV. CODE ANN. § 5122.15 (Baldwin Supp. 1981).

[211]Some courts have held that this is legally required. *See* Lynch v. Baxley, 386 F. Supp. 378, 389 (M.D. Ala. 1974); *In re* Fisher, 39 Ohio St. 2d 71, 313 N.E.2d 851 (1974); *see also* 42 U.S.C. § 9501(L)–(M) (Supp. V 1981); N.M. STAT. ANN. § 43-1-4 (1978). Some studies suggest that counsel frequently are poorly prepared to represent respondents. *See, e.g.*, Hiday, *The Attorney's Role in Involuntary Civil Commitment*, 60 N.C.L. REV. 1027, 1030 (1982). The Model Law does not go as far as N.M. STAT. ANN. § 43-1-4, which provides that patients "shall be entitled to obtain the advice of counsel at any time regarding their status under the Code," apparently at state expense. Ideally, however, adequate legal services should be available.

[212]In *In re* Tuntland, 71 Ill. App. 3d 523, 390 N.E.2d 11 (1979), the court permitted respondent to refuse appointed counsel and proceed pro se, but ordered counsel to remain available to advise respondent as needed. *See also In re* Hop, 29 Cal. 3d 82, 623 P.2d 282, 171 Cal. Rptr. 721 (1981) (on the validity of waivers). Several commentators have argued that adequate representation by counsel is critical to the outcome of commitment hearings. *See, e.g.*, Andalman & Chambers, *Effective Counsel for Persons Facing Civil Commitment: A Survey, a Polemic and a Proposal*, 45 MISS. L.J. 43 (1974); Cohen, *The Function of the Attorney and the Commitment of the Mentally Ill*, 44 TEX. L. REV. 424 (1966); Hiday, *The Role of Counsel in Civil Commitment: Changes, Effects,*

This paragraph also includes an optional provision by which the state might decide to provide indigent respondents one free psychiatric examination to help them in presenting their case. Only a few states now provide such a right.[213]

6.D.4. The District Attorney or County Counsel shall represent the interests of the State at the hearing. [If the District Attorney or County Counsel fails to proceed with the commitment, the next of kin or a petitioning party may retain counsel to do so in his stead, and the reasonable costs of such counsel shall be paid by the (state or local government).][214]

COMMENTARY: The state's interests, like those of the respondent, must be well represented at the commitment hearing. Unfortunately, many overburdened District Attorneys and County Counsels attach a low priority to this function, so that at many hearings no one adequately represents the public. This provision specifically mandates that the state's interests be represented.[215] The optional provision would permit the next of kin or a petitioning party to retain private counsel to pursue with the commitment if the District Attorney or County Counsel failed to do so.[216] Family members or other petitioning parties may have profound interests in commitment proceedings, and should not be precluded from asserting those interests because of the state's inaction. Compensation to such privately retained counsel would be limited to "reasonable" costs, perhaps ac-

---

Determinants, 5 J. PSYCHIATRY & L. 551 (1977); Litwack, *The Role of Counsel in Civil Commitment Proceedings: Emerging Problems*, 62 CALIF. L. REV. 816 (1974); Zander, *Civil Commitment in Wisconsin: The Impact of* Lessard v. Schmidt, 1976 WIS. L. REV. 503, 513–17; Special Project, *Involuntary Hospitalization of the Mentally Ill in Iowa: The Failure of the 1975 Legislation*, 64 IOWA L. REV. 1284 (1979); Note, *The Role of Counsel in the Civil Commitment Process: A Theoretical Framework*, 84 YALE L.J. 1540 (1975).

[213]*See, e.g.*, OHIO REV. CODE §§ 5122.05(c)(2), 5122.15(A)(4) (Baldwin 1980); *cf.* Suzuki v. Quisenberry, 411 F. Supp. 1113, 1129 (D. Hawaii 1970) ("The current state of the law does not require provision in all cases for investigators and expert witnesses at public expense to assist in the defense of an indigent person subjected to a civil commitment proceeding, but it does not take much foresight to anticipate a trend in this direction.").

[214]Optional provision.

[215]*Cf.* N.C. GEN. STAT. § 122.58.24 (1981) (creating full-time Associate Attorney General positions in each of the state hospitals to represent the state's interests). "It was felt that the full-time respondent attorneys in the state hospitals had an unfair advantage over their part-time adversaries, resulting in release of many respondents who might have been found to satisfy the criteria for commitment if both sides had been adequately represented." Miller & Fiddleman, *supra* note 14, at 996. The results of commitment hearings before and after full-time counsel were appointed apparently confirmed this belief. *Id.* at 1004. *See generally* Stone, *The Myth of Advocacy*, 30 HOSP. & COMMUNITY PSYCHIATRY 819 (1979).

[216]*Cf. In re* Kossow, 393 A.2d 97 (D.C. 1978).

cording to a schedule like that used for appointed defense counsel under the Federal Criminal Justice Act.[217]

6.D.5. The rules governing evidentiary and procedural matters at hearings under this Act shall be applied so as to facilitate informal, efficient presentation of all relevant, probative evidence and resolution of issues with due regard to the interests of all parties. Hearsay evidence may be received, and experts and other witnesses may, consistent with law, testify to any relevant and probative facts at the discretion of the court.

COMMENTARY: This paragraph encourages states to permit flexible, pragmatic procedures at commitment hearings, without necessarily adhering to all the rules of evidence applicable in criminal trials.[218] The criminal law evidentiary and procedural rules have evolved over hundreds of years. Although technical refinements continue, there is a relatively stable social consensus supporting the Fifth Amendment privilege, hearsay rule, and spousal privilege in criminal cases. The ritualistic importing of these rules into the civil commitment process in recent years, however, has ignored the fact that civil commitment has different goals, different substantive standards, and different results than the criminal process. Except for constitutionally required due process safeguards, the evidentiary and procedural rules used in criminal cases should be evaluated with an open mind as to their propriety in civil commitment hearings.

For example, hearsay evidence may be especially necessary in some civil commitment cases because the facts sought to be established include the elusive datum of mental status, not just physical events as in most criminal trials.[219] Especially where the issues are tried before a judge, who can weigh the probative value of the evidence, such hearsay need not be excluded.[220] It also may be proper to afford a broader ambit to psychiatric opinion evidence. Similarly, the interests that underlie the priv-

---

[217]18 U.S.C. § 3006A(d) (1976).

[218]*Cf.* UTAH CODE ANN. § 64-7-36(9) (Supp. 1981). *But cf.* N.C. GEN. STAT. § 122.58.7(i) (1981).

[219]Courts generally do not exclude as hearsay an expert's testimony based on reports of the type usually relied upon (e.g., medical charts, nurses' oral reports). *See* Jenkins v. United States, 307 F.2d 637 (D.C. Cir. 1962) (en banc); People v. Ward, 61 Ill. 2d 559, 338 N.E.2d 171 (1975); *In re* Dean, 94 N.M. 45, 607 P.2d 132 (1980); *cf.* FED. R. EVID. 703. Some have criticized the application of this rule to psychiatric testimony. *See, e.g.*, Note, *Hearsay Bases of Psychiatric Opinion Testimony: A Critique of Federal Rule of Evidence 703*, 51 S. CAL. L. REV. 129 (1977).

[220]*Cf. In re* Germich, 103 Ill. App. 3d 626, 431 N.E.2d 1092 (1981); Ritterbusch v. Speakes, 602 S.W.2d 937 (Mo. Ct. App. 1980). *But see In re* Hutchison, 279 Pa. Super. 401, 412–13, 421 A.2d 261, 266 (1980).

ilege against spousal testimony in criminal cases may not be served by invoking such a privilege in civil commitment.[221] There may be proper exceptions to other privileges as well, such as the psychotherapist-patient privilege.[222]

The Model Law contemplates that the court may receive testimony in narrative form, and talk informally with the parties to ascertain the basic facts. This provision does not, however, dispense with appropriate limits on the admissibility of evidence. Such evidence must still be relevant, probative, and presented in a manner reflecting "due regard to the interests of all parties."

6.D.6. Patients shall not have a "right to remain silent" at a psychiatric examination or hearing conducted pursuant to this Act; provided that no patient shall be held civilly or criminally liable for not speaking or testifying. Any information obtained from or disclosed by the patient during the course of evaluation or treatment is admissible in any hearing provided in this Act without regard to whether it would otherwise be privileged; provided that no disclosure made by the patient during the course of evaluation or treatment or in any proceeding conducted under this Act, and no opinion testimony based on such disclosures, may be admitted against the patient on the issue of guilt in a criminal proceeding unless he places his mental condition in issue in such proceeding, and unless the disclosure or opinion is relevant to such an issue raised by him.

COMMENTARY: This paragraph specifically provides that patients do not have a Fifth Amendment right to remain silent at a psychiatric examination conducted pursuant to the Model Law. Historically, the Fifth Amendment bar against compulsory self-incrimination has applied only where the testimony might lead to criminal liability.[223] Continuing legal controversy surrounds the question of whether this doctrine should be applied in the civil commitment context. Some courts, stressing that

---

[221] *See* Commonwealth *ex rel.* Platt v. Platt, 266 Pa. Super. 276, 282–83, 404 A.2d 410, 413–14 (1979).

[222] *See, e.g., In re* Field, 120 N.H. 206, 412 A.2d 1032 (1980); *In re* Farrow, 41 N.C. App. 680, 255 S.E.2d 777 (1979); *In re* Winstead, 67 Ohio App. 2d 111, 425 N.E.2d 943 (1980); Jones v. State, 610 S.W.2d 535 (Tex. Civ. App. 1980).

[223] The nature of the proceedings in which the privilege is invoked is not decisive; what is decisive is the possible liability that may accrue. Thus,

> [t]he privilege can be claimed in any proceeding, be it criminal or civil, administrative or judicial, investigatory or adjudicatory . . . . [I]t protects any disclosures which the witness may reasonably apprehend could be used in a criminal prosecution or which could lead to other evidence that might be so used.

In re Gault, 387 U.S. 1, 47–48 (1967) (quoting Murphy v. Waterfront Comm'n, 378 U.S. 52, 94 (1964)) (emphasis omitted).

involuntary commitment deprives people of their liberty, have ruled that the Fifth Amendment does apply.[224] A larger and growing number of courts, however, have held that the Fifth Amendment privilege should not apply.[225]

Many cases that discuss whether the privilege should apply focus on the metaphysical issue of whether commitment is "essentially" a civil or criminal proceeding. This analysis misses the point. Advising a patient—who may well have an emergency psychiatric condition—at the beginning of an interview about his right to remain silent would be fundamentally inconsistent with the therapeutic purposes of the process. It might bewilder and alarm the patient. It might make it impossible to ascertain the patient's mental state, thereby preventing the assessment both of his need for treatment and his potential dangerousness.

Commitment decisions would be transformed into judgments based solely on overt acts, becoming virtually indistinguishable from decisions made in the criminal process. Granting the Fifth Amendment privilege would establish a bootless procedural right—in some cases forcing the state to instigate criminal proceedings and in many cases depriving seriously mentally ill persons of needed treatment.[226] Moreover, as several writers

---

[224]*See, e.g.*, Tyars v. Finner, 518 F. Supp. 502, 509–10 (C.D. Cal. 1981); Lynch v. Baxley, 386 F. Supp. 378, 394 (M.D. Ala. 1974); Lessard v. Schmidt, 349 F. Supp. 1078, 1100–02 (E.D. Wis. 1972), *vacated on other grounds*, 414 U.S. 423 (1974); *cf. In re* Gault, 387 U.S. 1, 42–57 (1967) (privilege applicable to juvenile offense proceedings). Certain state laws provide a similar privilege. *See, e.g.*, ARK. STAT. ANN. § 59-1408(5) (Supp. 1979); ILL. REV. STAT. ch. 91 1/2, § 3-208 (1981); OHIO REV. STAT. § 5122.15(A)(12) (Baldwin 1980).

[225]*See, e.g.*, French v. Blackburn, 428 F. Supp. 1351, 1358–59 (M.D.N.C. 1977), *aff'd mem.*, 443 U.S. 901 (1979); Cramer v. Tyars, 23 Cal. 3d 131, 588 P.2d 793, 151 Cal. Rptr. 653 (1979); Mitchell v. County of Los Angeles, 114 Cal. App. 3d 606, 170 Cal. Rptr. 759 (1980); People v. Taylor, 618 P.2d 1127, 1140 (Colo. 1980); *In re* Nolan, 66 Ill. App. 3d 744, 384 N.E.2d 134 (1978); *In re* Field, 120 N.H. 206, 412 A.2d 1032 (1980); *In re* Winstead, 67 Ohio App. 2d 111, 425 N.E.2d 943 (1980); *In re* Matthews, 46 Or. App. 757, 613 P.2d 88 (1980), *cert. denied*, 450 U.S. 1040 (1981); State *ex rel.* Ellenwood, 567 S.W.2d 251, 253 (Tex. Civ. App. 1978); State *ex rel.* Hawks v. Lazaro, 157 W. Va. 417, 443–44, 202 S.E.2d 109, 126 (1974). Also, it is generally held that the Fifth Amendment does not bar a compulsory psychiatric examination. *See, e.g.*, Shifflett v. Commonwealth, 221 Va. 760, 769–70, 274 S.E.2d 305, 311 (1981).

[226]As one court has observed, imposing such a privilege might make many proper commitments impossible:

> To recognize the same rights against self-incrimination which would be required in a criminal proceeding would make almost any commitment impossible, and would make the procedures so burdensome that medical conclusions obtained through examination [in emergency detention situations] would be inadmissible .... It would do a great disservice to individuals to make the procedural requirements so cumbersome that suicidal and maniacal individuals could never be hospitalized until they had injured themselves or others.

State *ex rel.* Hawks v. Lazaro, 157 W. Va. 417, 443–44, 202 S.E.2d 109, 126 (1974);

have concluded, applying the Fifth Amemendment privilege to civil commitment does not even make sense as a matter of legal theory.[227]

This paragraph also provides a countervailing protection for the patient. No disclosure made by the patient, and no opinion testimony based upon such disclosure, may be admitted against him on the issue of guilt in a criminal proceeding unless he places his mental condition in issue.[228] Such information could be adduced on issues other than guilt, such as competency, or if the defendant raises the insanity defense.[229]

This provision also establishes that the physician-patient privilege does not preclude disclosure in a commitment hearing of information obtained from a patient during evaluation or treatment. The family's and the state's interest in seeing that a person who meets the commitment criteria receives treatment outweighs any competing privacy interest of the patient.[230] It would

---

accord Suzuki v. Yuen, 617 F.2d 173, 177–78 (9th Cir. 1980); French v. Blackburn, 428 F. Supp. 1351, 1358–59 (M.D.N.C. 1977), *aff'd mem.*, 443 U.S. 901 (1979); State *ex rel.* Ellenwood, 567 S.W.2d 251, 253 (Tex. Civ. App. 1978).

[227]"Recognition of [an] individual's right to remain silent would . . . seriously impair the state's ability to achieve the valid objectives of civil commitment." Note, *Developments in the Law—Civil Commitment of the Mentally Ill*, 87 HARV. L. REV. 1190, 1307 (1974). *See also* Aronson, *Should the Privilege Against Self-Incrimination Apply to Compelled Psychiatric Examinations?*, 26 STAN. L. REV. 55 (1973); Bonnie & Slobogin, *The Role of Mental Health Professionals in the Criminal Process: The Case for Informed Speculation*, 66 VA. L. REV. 427, 497–502 (1980); Danforth, *Death Knell for Pretrial Mental Examination? Privilege Against Self-Incrimination*, 19 RUTGERS L. REV. 489 (1965); Lefelt, *Pretrial Mental Examinations: Compelled Cooperation and the Fifth Amendment*, 10 AM. CRIM. L. REV. 431 (1972).

[228]*See* Hughes v. Mathews, 576 F.2d 1250 (7th Cir.), *cert. dismissed*, 439 U.S. 801 (1978). *But see* Steele v. State, 97 Wis. 2d 72, 294 N.W.2d 2 (1980) (no psychiatric opinion testimony is admissible on guilt in bifurcated proceeding). *See generally* Note, *Restricting the Admission of Psychiatric Testimony on a Defendant's Mental State: Wisconsin's* Steele *Curtain*, 1981 WIS. L. REV. 733.

[229]We stress that this Model Law does not deal with the situations akin to that in Estelle v. Smith, 451 U.S. 454 (1981), a case presenting the most repugnant abuses of psychiatry, in which the Supreme Court held that the Fifth Amendment barred psychiatric testimony at a capital sentencing hearing where the defendant was compelled to submit to a psychiatric examination and was not told that he had a right to remain silent. *Estelle* may suggest that if, under the Model Law, no such "Miranda" warning were given, then the results of the examination could not be introduced in a subsequent sentencing hearing.

[230]While virtually every state has either a general physician-patient privilege or a specific psychiatrist-patient privilege, the scope of such privileges is on the wane. *See* Camperlengo v. Blum, 56 N.Y.2d 251, 436 N.E.2d 1299, 451 N.Y.S.2d 697 (1982); *In re* Farrow, 41 N.C. App. 680, 255 S.E.2d 777 (1979); *In re* Winstead, 67 Ohio App. 2d 111, 425 N.E.2d 943 (1980); CAL. EVID. CODE § 1010–1026 (West 1966 & Supp. 1982–1983); D.C. CODE ANN. §§ 2-1704.16, 14-307 (1981); ILL. REV. STAT. ch. 111, § 5306 (1981). For a criticism of the need for the psychiatrist-patient privilege, see Shuman & Weiner, *The Privilege Study: An Empirical Examination of the Psychotherapist-Patient Privilege*, 60 N.C.L. REV. 893, 895 (1982).

confound the commitment process to require a psychiatrist to examine an emergency patient and recommend whether he should be committed, but then bar the psychiatrist from revealing his results. Other state laws will govern whether a similar rule applies to a privately retained psychiatrist who previously examined the patient not then "under this Act."[231]

6.D.7. The hearing shall be closed to the public, unless the respondent requests that it be open, or the court determines for other good cause that the hearing should be open. The court shall keep a complete record, written or recorded, of every hearing.

COMMENTARY: The Model Law begins with the assumption that most commitment hearings should be closed to the public. This is to maximize candor from the respondent's family and other witnesses, and to minimize embarrassment to the respondent. It is also important, however, that no suspicion of secrecy or taint of unfairness surround the civil commitment process. The Model Law accordingly recognizes the respondent's right to request an open hearing. In addition, the court may order an open hearing where an overriding public interest in a celebrated case or other factors warrant it. On the other hand, as is recognized in juvenile and certain other proceedings, countervailing considerations may dictate closing a particular hearing even over the respondent's objection. For example, a woman may feel strongly that her husband needs to be committed, but may be unwilling to proceed if she is required to give public testimony about intimate details. The authors believe that this provision should be construed to permit some judicial discretion in such cases.[232]

In all cases, the court must keep a complete record of the hearing.[233] The less costly option of a taped record may be preferred in some states.

---

[231]A Miranda-type warning may be appropriate in such cases. *See* ILL. REV. STAT. ch. 91 1/2, § 3-208 (1981).

[232]*Cf.* TEX. REV. CIV. STAT. ANN. art. 5547-49(e) (Vernon Supp. 1982–1983) (so long as the patient consents, "the court may exclude all persons not having a legitimate interest in the proceedings"); UTAH CODE ANN. § 64-7-36(9) (Supp. 1981) (court can exclude on its own motion all persons "not necessary" to the proceedings). *Compare* PA. STAT. ANN. tit. 50, § 7304(e)(7) (Purdon 1969 & Supp. 1981) (hearing is public unless respondent asks that it be closed), *with* OHIO REV. CODE ANN. § 5122.15(A)(5)–(6) (Baldwin 1980) (hearing is closed unless respondent requests it to be open, but in the event of a closed hearing the court may admit certain individuals).

[233]Due process requires this. *See* Suzuki v. Quisenberry, 411 F. Supp. 1113, 1133 (D. Hawaii 1976); Schuman v. State, 358 So. 2d 1333 (Fla. 1978); *see also* CONN. GEN. STAT. § 17-178(c) (1981); ILL. REV. STAT. ch. 91 1/2, § 3-817 (1981).

6.D.8. At the conclusion of the hearing, or within one business day thereafter, the court shall make its findings, including specific findings as to whether the commitment is warranted because the person is (a) likely to cause harm to others, or (b) likely to cause harm to himself or to suffer substantial mental or physical deterioration, or (c) both (a) and (b). As to any person found likely to cause harm to himself or to suffer substantial mental or physical deterioration, the court shall further make findings as to whether commitment is warranted because the person (a) is likely in the near future to inflict substantial physical injury upon himself, or (b) is substantially unable to provide for some of his basic needs, such as food, clothing, shelter, health, or safety, or (c) will, if not treated, suffer severe and abnormal mental, emotional, or physical distress and this distress is associated with significant impairment of judgment, reason, or behavior causing a substantial deterioration of his previous ability to function on his own.

COMMENTARY: Hearings on thirty-day commitments are held before a court, not a jury. (See also the discussion of hearings on recommitment under Section 11.) This provision requires the court to make specific findings as to the basis for each commitment. The Model Law would permit the court to consider the matter overnight or over a weekend and to study written materials, so long as it renders its decision within one business day of the end of the hearing.[234] Much confusion in the civil commitment process and in the treatment of institutional populations arises because most state statutes do not require such an articulation of the reasons for particular commitments. This omission confounds later efforts to decide whether restraints or seclusion may be warranted, how long a period of commitment is required, and what must be demonstrated in order to justify discharge. The Model Law therefore requires that the basis for each commitment be made clear. This also is necessary because under the Model Law discharge procedures differ for police power and parens patriae commitments.[235]

6.D.9. The court shall enter an order discharging the person unless it finds by clear and convincing evidence that the person satisfies all of the criteria for commitment in subsection 6.C., in which event it shall enter an order committing the person for evaluation and treatment for a period of "up to thirty days." If at any time during thirty-day (or any subsequent) commitment a patient is absent without permission, the order of commitment constitutes a continuing authorization to the treatment facility and to any police officer to procure his return.

---

[234]*See* PA. STAT. ANN. tit. 50, § 7304(e)(7) (Purdon 1969 & Supp. 1981) (decision of court must be made within 48 hours).

[235]*See* Section 12.

COMMENTARY: This provision supplies clear statutory authority to police officers to return to a treatment facility a person who leaves without authorization.

Section 7. INFORMED CONSENT TO MEDICATION OR OTHER TREATMENT—VOLUNTARY PATIENTS.

7.A. *Informed Consent.*

Except in an emergency situation, a treatment facility shall, prior to beginning any course of medication or other treatment for a patient who is subject to voluntary admission under Section 5, obtain informed consent to treatment. If the patient does not lack capacity to make an informed decision concerning treatment, the consent shall be his own. If he does lack such capacity, the consent shall be that of his next of kin or guardian, provided that such a patient may receive appropriate medications or other treatments, except as limited by subsection 8.C., until such time as the consent or refusal to consent of such next of kin or guardian can be obtained.

COMMENTARY: This section requires the informed consent of a voluntary patient before treatment, and confers on him the right to refuse treatment "[e]xcept in an emergency situation."[236] A full informed-consent protocol should be conducted before beginning each new "course of medication or other treatment;" it is not required before administering each dosage of medication. The patient, however, may withdraw his consent at any time. In order to avoid the charade of obtaining "informed consent" from a voluntary patient who lacks capacity, this section also provides that for such patients, the informed consent of the next of kin or guardian shall be obtained.[237] Thus, the best feasible informed consent is required.

Initially, the treatment staff must decide whether the patient lacks the capacity to give his informed consent to treatment. If so, and if the patient objects to substituted consent by the next of kin or guardian, or if there appears to be some conflict of interest, the treatment facility may apply to the court for a prompt decision.

7.B. *Revocation of Consent.*

A voluntary patient (or the next of kin or guardian who consented to treatment on his behalf) may revoke consent to treatment at any time

[236]*Accord* CONN. GEN. STAT. § 17-206(a) (1981); N.J. STAT. ANN. § 30:4-24.2(d)(1) (West 1981); WIS. STAT. § 51.61(6) (1979–1980); *see also* Rogers v. Okin, 478 F. Supp. 1342, 1368 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982).

[237]See the discussion of subsection 5.A. concerning the need to ensure that the next of kin or guardian is acting in the patient's best interests.

by a reasonably clear statement in writing (and patients who indicate a desire to revoke consent but are unable to write shall be helped to put their statement in writing). If such consent is revoked, the treatment shall be promptly discontinued, provided that a course of treatment may be concluded or phased out where necessary to avoid the harmful effects of abrupt withdrawal.

COMMENTARY: A voluntary patient (or the next of kin or guardian who has been acting as the substitute decisionmaker) may revoke his prior consent to treatment at any time. The Model Law requires written revocation, to avoid confusion and litigation over whether a patient's casual remark should be construed as a demurrer to further treatment. For example, even a voluntary patient should not be able to revoke his consent to a medication or therapy session simply because he wishes to sleep late one day, or to skip a group therapy session. The patient should not be allowed to waste the scarce medical resources available. On the other hand, if a patient verbally expresses his firm desire that treatment stop, he should be helped to put this desire in writing. Under most circumstances, treatment should terminate "promptly." However, because the abrupt withdrawal of medication may cause harm to the patient and others, certain treatments may be phased out rather than terminated immediately.[238]

7.C. *Refusal to Consent.*

Except in an emergency situation, any voluntary patient (himself or through his next of kin or guardian) shall have the right to refuse any and all medications or other treatments. If appropriate medications or treatments are refused, the facility may then discharge the patient, and shall not be liable in any respect for such action.

COMMENTARY: In all situations other than emergencies, a voluntary patient may refuse any medication or other treatment, and the facility must respect his refusal. In an emergency situation, however, even a voluntary patient may be treated without his consent.[239]

If a voluntary patient refuses all appropriate treatments, however, the hospital is often faced with a practical problem: it can

[238]The legal justification can be found in the original "contract" for voluntary commitment, provided that the patient was advised of this procedure. *Cf.* Rogers v. Okin, 478 F. Supp. 1342, 1367–68 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982).

[239]*See, e.g.*, Rennie v. Klein, 653 F.2d 836 (3d Cir. 1981), *vacated and remanded*, 102 S. Ct. 3506 (1982); Rogers v. Okin, 634 F.2d 650, 659–60 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982).

neither treat the patient nor release him. A voluntary patient should not be permitted to dictate the form of treatment to staff, or to fill a scarce hospital bed without receiving treatment.[240] If a patient refuses a particular treatment, the medical staff should explore alternative therapies with him. If he also refuses these alternatives and the staff concludes that confinement would serve no further purpose, the Model Law authorizes the facility to discharge the patient without incurring legal liability.

Section 8. Informed Consent to Medication or Other Treatment—Involuntary Patients.

8.A. *Consent During Emergency Evaluation.*

Following admission and during the period of emergency evaluation provided in Section 4, the treatment facility may administer medications or other treatments to a patient, except as limited by subsection 8.C., consistent with good medical practice and without the informed consent of the patient or his next of kin or guardian. However, prior to administering any such medication or other treatment, the staff shall explain the purposes, nature, and effects of the treatment and shall request the patient's consent to it, unless the responsible psychiatrist [or: "the responsible physician"][241] determines that the patient's condition makes doing so infeasible or harmful to him and enters the reasons for not doing so in the record.

Commentary: The existence of a medical emergency, the action of the persons initiating emergency evaluation, and the decision of the independent psychiatrist who admitted the patient, together provide a sufficient legal justification for treating a patient during emergency evaluation without his consent.[242] The special therapies subject to regulation under subsection 8.C are excluded, however.[243] Although a patient is presumed legally

---

[240]*See* Rogers v. Okin, 634 F.2d at 661:

> [T]he district court in effect found that Massachusetts citizens have a constitutional right upon voluntary admittance to state facilities to dictate to the hospital staff the treatment that they are given. The district court cited no authority for this finding, and we know of none . . . . The statute does not guarantee voluntary patients the treatment of their choice. Instead it offers a treatment regimen that state doctors and staff determine is best, and if the patient thinks otherwise, he can leave.

[241]Optional provision.

[242]*See* Rogers v. Okin, 478 F. Supp. 1342, 1362 (D. Mass. 1979), *aff'd*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982); Rennie v. Klein, 462 F. Supp. 1131 (D.N.J. 1978), *supplemented*, 476 F. Supp. 1294 (D.N.J. 1979), *modified and remanded*, 653 F.2d 836 (3d Cir. 1981), *vacated and remanded*, 102 S. Ct. 3506 (1982); Cal. Welf. & Inst. Code § 5358.2 (West Supp. 1982); Idaho Code § 39-143 (1977); La. Rev. Stat. Ann. § 40:1299.54 (West 1977 & Supp. 1983); 70 C.J.S. *Physicians & Surgeons* § 48 (1951).

[243]*Accord* Ark. Stat. Ann. § 59-1415 (1971) (also excluding long-lasting drugs such

competent, he is de facto incompetent during the emergency period.[244] This "emergency" exception to the informed-consent doctrine has long been recognized. It was not challenged even in the "right to refuse treatment" cases, although these cases do debate what constitutes an "emergency."[245] The treatment staff must describe the emergency treatment to the patient and the reasons for employing it, unless doing so is infeasible (for example, a violent, out-of-control patient), or harmful to the patient (for example, a psychotic, agitated patient).[246] The reasons for dispensing with such an explanation must be noted in the chart to ensure that this exception does not swallow the rule.

8.B. ***Consent During Thirty-Day or Subsequent Commitments.***

Since it is a prerequisite to involuntary commitment that the person lacks capacity to make an informed decision concerning treatment, the treatment facility shall be authorized to administer medications or other treatments to such persons, except as limited by Section 8.C., consistent with good medical practice and without their consent. Although consent to treatment is not required, during the course of treatment the responsible psychiatrist [or: "the responsible physician"][247] shall consult with the patient and his next of kin or guardian and give consideration to the views they express concerning treatment and any alternatives.

COMMENTARY: Unlike a voluntary patient, an involuntary patient may not refuse treatment. In recent years, the right to refuse treatment has become perhaps the most incendiary issue in all of mental health law. To some civil-liberties lawyers, it partially substitutes for the outright abolition of involuntary commitment that they seek. Others champion the right because they suspect that certain therapies (such as electroconvulsive therapy (ECT) or behavior modification) are overused, that an-

---

as fluphenazine decanoate); CONN. GEN. STAT. § 17-206d (1981); MICH. COMP. LAWS § 330.1716 (1979); N.M. STAT. ANN. § 43-1-15(F) (1978); N.C. GEN. STAT. § 122-55.6 (1981 & Supp. 1981); VT. STAT. ANN. tit. 18, § 7708 (1968); WASH. REV. CODE § 71.05.370(7) (1981).

[244]*Cf. In re* President & Directors of Georgetown College, 331 F.2d 1000 (D.C. Cir. 1964), *cert. denied*, 377 U.S. 978 (1964) (refusal of de facto incompetent patient need not be accepted); Demers v. Gerety, 85 N.M. 641, 615 P.2d 645 (Ct. App. 1973) (assent of de facto incompetent person may not be relied on).

[245]*See* Rogers v. Okin, 634 F.2d at 654; Rennie v. Klein, 653 F.2d at 847. *See generally* Meisel, *The Exceptions to the Informed Consent Doctrine: Striking a Balance Between Competing Values in Medical Decisionmaking*, 1979 WIS. L. REV. 413, 434–38.

[246]*See* Appelbaum & Roth, *supra* note 66, at 1463 ("The seemingly neutral words of the informed consent form . . . can, in fact, be highly charged for the patient . . . due to the patient's unique matrix of previous experiences . . . . [T]he patient may [then] be forced to revert to more primitive, even psychotic, levels of defense for coping with it.").

[247]Optional provision.

tipsychotic medications have an unfavorable risk-to-benefit ratio for most patients, and that less restrictive modes of therapy are usually available.[248] Some judges seem to view this right as a patient's protection against poor institutional conditions.[249] But the major issue in the "right to refuse treatment" cases "lies deeper than questions of physical harm allegedly caused by these treatments."[250] The issue is the allocation of decisionmaking power to mental patients rather than to their physicians. The problem thus involves "preserving the freedom, autonomy, and integrity of individuals in a philosophical sense."[251]

So conceived, it was virtually inevitable that the courts would recognize a right to refuse treatment. A mental patient has undeniable liberty and privacy interests in not being forced to take powerful, mind-affecting drugs. Because several of the cases raising the issue involved extreme therapies administered in repellant circumstances, the courts came to believe that the right to refuse treatment could help protect patients against poor treatment conditions. Unfortunately, several recent "right to refuse treatment" cases have left a confusing residue. Two key cases deserve detailed discussion.

In *Rennie v. Klein*,[252] a schizophrenic, depressed, involuntary patient asserted that his constitutional right to privacy was violated when he was forced to take antipsychotic drugs. Swayed in part by the inadequacies of the treatment staff and descriptions of the "harsh side effects" of such drugs, the district court ruled that "individual autonomy demands" that hospitals respect a patient's "qualified right" to refuse medication.[253] Such a right might be affected by other considerations, the court said, including (1) the patient's potential to physically threaten injury

---

[248]*See, e.g.*, Plotkin, *Limiting the Therapeutic Orgy: Mental Patients' Right to Refuse Treatment*, 72 NW. U.L. REV. 461 (1977). Plotkin's evaluation of the side effects of antipsychotic medications are distorted and misleading in several respects, however. *See* Amicus Brief for the APA, Rogers v. Mills, 102 S. Ct. 2442 (1982).

[249]*See* Gutheil, *The Right to Refuse Treatment*, in PSYCHIATRY 1982, 379 (L. Grinspoon ed. 1983); Stone, *supra* note 197.

[250]Comment, *Madness and Medicine: The Forcible Administration of Psychotropic Drugs*, 1980 WIS. L. REV. 497, 500.

[251]*Id.* at 502.

[252]462 F. Supp. 1131 (D.N.J. 1978), *supplemented*, 476 F. Supp. 1294 (D.N.J. 1979), *modified and remanded*, 653 F.2d 836 (3d Cir. 1981), *vacated and remanded*, 102 S. Ct. 3506 (1982).

[253]*Id.* at 1145. The psychotropic drugs referred to included prolixin, thorazine, mellaril, and haldol. *See generally* Note, Rennie v. Klein: *Constitutional Right of Privacy Protects a Mental Patient's Refusal of Psychotropic Medication*, 57 N.C.L. REV. 1482 (1979); Comment, *Mental Health--Restraint Treatment in Institutions--Right to Refuse Treatment Based on Right of Privacy*, 55 N.D.L. REV. 503 (1979).

to another, (2) his competence to decide on treatment, (3) the availability of less restrictive treatment, and (4) the risk of permanent side effects.[254] In a supplementary decision, the court prescribed a complex procedure that hospitals must follow when patients refuse treatment, including referral to a "patient advocate" for review and possible referral for a hearing before an "independent psychiatrist."[255]

The appellate court upheld the right to refuse treatment but modified the procedures. The Third Circuit Court of Appeals agreed that involuntary patients "retain a constitutional right to refuse antipsychotic drugs that may have permanently disabling side effects."[256] It added that the "state may override that right when the patient is a danger to himself or others"[257] or when required for "needed care or administrative concerns,"[258] but that "in non-emergency situations [it] must first provide procedural due process."[259] The court made a rather vague ruling that required the therapists to engage in a "conscious weighing of the constitutional liberty interest in any determination of proper treatment alternatives," and to consider whether a drug's "cost benefit ratios, weighed from the patient's standpoint, are unacceptable."[260] While conceding that "due process procedures must provide an opportunity for the exercise of professional judgment,"[261] the court left the treatment staff with essentially no standards for its conduct, except the ad hoc decisions later made by judges and juries.

In modifying the procedures required by the district court, the Third Circuit recognized that "the adversary process implicit in the District Court's order is ill-suited to the type of medical determination that must be made."[262]It said that the New Jersey administrative rules were adequate in providing for independent review of the advisability of treatment by the medical director or by another physician independent of the treatment team.[263]

---

[254]462 F. Supp. at 1145–48.
[255]476 F. Supp. at 1314.
[256]653 F.2d at 838.
[257]*Id.*
[258]*Id.* at 845.
[259]*Id.* at 838.
[260]*Id.* at 847.
[261]*Id.* at 848.
[262]"The weeks or months that a patient spends in these institutions should provide a more accurate and reliable basis for the staff's judgment as to whether the patient poses a danger to himself or others or whether he is capable of making a rational treatment decision." *Id.* at 850; *accord* Parham v. J.R., 442 U.S. 584, 607 (1979).
[263]653 F.2d at 848.

The second pivotal "right to refuse treatment" case was *Rogers v. Okin*.[264] In this case, the district court again noted the serious side effects of psychotropic drugs, and the fact that under Massachusetts law "commitment to Boston State [Hospital], even on an involuntary basis, is not an adjudication of incompetence."[265] It concluded that:

> The committed patient has a right to be wrong in his analysis of [treatment] information . . . as long as the consequences of such error do not pose a danger of physical harm to himself, fellow patients, or hospital staff. And so, while the state may have an obligation to make treatment available, and a legitimate interest in providing such treatment, a competent patient has a fundamental right to decide to be left alone, absent an emergency.[266]

The First Circuit affirmed in part and reversed in part. It recognized the constitutional right of every person to be left alone. The appellate court observed, however, that the state's police power justifies forcible medication of a patient in an emergency "whether or not that person has been adjudicated incompetent."[267] It rejected the district court's "unitary standard" for defining emergencies (that is, that violence was likely unless the patient was forcibly medicated). Instead, it said that

> The professional judgment call required in . . . determining whether a patient should be subjected to forcible administering of antipsychotic drugs demands an *individualized* estimation of the possibility and type of violence, the likely effects of particular drugs on a particular individual, and an appraisal of alternate, less restrictive courses of action.[268]

In summarizing its holding, the First Circuit said that (1) under the parens patriae power "medication cannot be forcibly administered solely for treatment absent a finding of incompetency,"[269] and (2) under the police power, the "decision must be the result of a determination that the need to prevent

---

[264]478 F. Supp. 1342 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982).

[265]*Id.* at 1359–61.

[266]*Id.* at 1367.

[267]634 F.2d at 654, 656.

[268]*Id.* at 655–56.

[269]*Id.* at 656. The court expanded its statement thus: "the *sine qua non* for the state's use of its parens patriae power as a justification for the forcible administration of mind-affecting drugs is a determination that the individual . . . lacks the capacity to decide for himself whether he should take the drugs." *Id.* at 657. Under the Massachusetts

violence in a particular situation outweighs the possibility of harm to the medicated individual."[270]

These legal decisions[271] should be considered in light of the following research findings.[272] First, antipsychotic medications benefit many patients and permit them to leave the hospital sooner.[273] In one major study, seventy-five percent of the patients treated with antipsychotic drugs showed marked to moderate improvement within six weeks of hospital admission, while only twenty-three percent of the placebo group achieved such improvement.[274]

Second, most patients need involuntary medication only infrequently, and patients usually come to accept the medication. Voluntary acceptance by an involuntary patient, of course, may be flawed. But it is striking that about two-thirds of involuntary patients continue voluntary treatment following a period of involuntary confinement, frequently with drug therapy.[275]

---

commitment law, no adjudication of incompetency was required for commitment. *Id.* at 658–59.

[270]*Id.* at 656. The Supreme Court later vacated and remanded the court of appeals' decision, without really challenging its reasoning. It remanded for reconsideration in light of an intervening state court decision that created the possibility that "Massachusetts [state law] recognizes liberty interests of persons adjudged incompetent that are broader than those protected directly by the Constitution of the United States." Mills v. Rogers, 102 S. Ct. 2442, 2450 (1982).

[271]*See also* Davis v. Hubbard, 506 F. Supp. 915 (N.D. Ohio 1980) (except in emergency, only a court may order forcible medication); Naughton v. Bevilacqua, 458 F. Supp. 610 (D.R.I. 1978) (use of phenothiazines simply to control patients is not justifiable); Gundy v. Pauley, 619 S.W.2d 730 (Ky. 1981) (upholding right to refuse ECT); *In re* Roe, 421 N.E.2d 40 (Mass. 1981) (if an incompetent person refuses antipsychotic drugs, judicial determination of substituted judgment must be sought); *In re* K.K.B., 609 P.2d 747 (Okla. 1980) (involuntary but competent patients have a right to refuse antipsychotic drugs).

[272]In this discussion we are particularly indebted to Roth and Appelbaum for their excellent article *What We Do and Do Not Know About Treatment Refusals in Mental Institutions*, *supra* note 199.

[273]*See, e.g.*, W. APPELTON & J. DAVIS, PRACTICAL CLINICAL PSYCHOPHARMACOLOGY (1980); Kessler & Weletzky, *Clinical Use of Antipsychotics*, 13 AM. J. PSYCHIATRY (1981); May, Tuma & Dixon, *Schizophrenia--A Follow-up Study of Results of Treatment*, 33 ARCHIVES GEN. PSYCHIATRY 474 (1976); Nat'l Inst. of Mental Health, Psychopharmacology Serv. Center Collaborative Study Group, *Phenothiazine Treatment in Acute Schizophrenia*, 10 ARCHIVES GEN. PSYCHIATRY 246 (1964) [hereinafter cited as *NIMH Study*]. Of course, some patients may benefit from treatment apart from drugs. *See* Carpenter, McGlashan & Strauss, *The Treatment of Acute Schizophrenia Without Drugs: An Investigation of Some Current Assumptions*, 134 AM. J. PSYCHIATRY 14 (1977).

[274]*NIMH Study*, *supra* note 273; *see also* P. MAY, TREATMENT OF SCHIZOPHRENIA: A COMPARATIVE SURVEY OF FIVE TREATMENT METHODS (1968); Gross, Hitchman, Reeves, Lawrence & Newell, *Discontinuation of Treatment with Ataractic Drugs*, in 3 RECENT ADVANCES IN BIOLOGICAL PSYCHIATRY 44 (J. Wortis ed. 1961).

[275]*See, e.g.*, Chodoff & Peele, *Involuntary Hospitalization and Treatability: Observations from the District of Columbia Experience*, 23 CATH. U.L. REV. 744 (1974); Sata & Goldenberg, *A Study of Involuntary Patients in Seattle*, 28 HOSP. & COMMUNITY

Third, many or most committed patients lack capacity to make reasoned treatment decisions, and therefore probably lack capacity to make reasoned refusals.[276] Moreover, many patients state reasons for refusing that are verbally incomprehensible, obvious manifestations of their disease (for example, "the CIA wants you to use this drug to kill me"), or manifestly ambivalent.[277]

Fourth, although there is considerable controversy over the frequency and seriousness of adverse side effects of antipsychotic medications, such as tardive dyskinesia, most research suggests that few cases develop from drug treatment lasting less than three months, and most cases occur only after use of antipsychotic drugs for more than two years.[278] The possibility of such side effects should not be minimized, and patients on such medications should be monitored carefully. Fortunately, medical means to avoid and to treat such effects are improving.

Fifth, the impact of treatment refusals must be considered. One study of the impact in Massachusetts following the *Rogers* decision found that more than half (89 of 159) of the patients who persistently refused treatment deteriorated.[279] These patients were

> subjected to the terrors and harshness of their psychotic inner lives for long periods of time. Many times these pa-

---

PSYCHIATRY 834 (1977); Spensley, *Involuntary Hospitalization: What for and How Long*, 131 AM. J. PSYCHIATRY 219 (1974); Zwerling, Karasu, Plutchik & Kellerman, *A Comparison of Voluntary and Involuntary Patients in a State Hospital*, 45 AM. J. ORTHOPSYCHIATRY 81 (1975).

[276]*See* Geller, *State Hospital Patients and Their Medication--Do They Know What They Take*, 139 AM. J. PSYCHIATRY 611 (1982) (55% of patients exhibited no understanding of the drugs they were taking and only 14% had a reasonably full understanding); Grossman & Summers, *supra* note 199; Soskis, *supra* note 58. Many patients subsequently are grateful for involuntary commitment and medication. *See* Kalman, *An Overview of Patient Satisfaction with Psychiatric Treatment*, 34 HOSP. & COMMUNITY PSYCHIATRY 48 (1983).

[277]In one study of 23 patients who refused treatment on 72 occasions, their stated reasons (sometimes more than one) were as follows: "no reason offered (nine patients); angry or seemingly irrelevant responses (seven patients); side effects (10 patients); overt delusions (nine patients); privacy (eight patients); legal rights (three patients)." Appelbaum & Gutheil, *Rotting with Their Rights On: Constitutional Theory and Clinical Reality in Drug Refusal by Psychiatric Patients*, 7 BULL. AM. ACAD. PSYCHIATRY & L. 306, 312–13 (1979) [hereinafter cited as Appelbaum & Gutheil, *Rotting with Their Rights On*]; *see also* Appelbaum & Gutheil, *Drug Refusal: A Study of Psychiatric Inpatients*, 137 AM. J. PSYCHIATRY 340, 342–44 (1980) [hereinafter cited as Appelbaum & Gutheil, *Drug Refusal*].

[278]*See* APA, TASK FORCE REP. NO. 18, TARDIVE DYSKINESIA (1979); Kline, *On the Rarity of "Irreversible" Oral Dyskinesias Following Phenothiazines*, 124 AM. J. PSYCHIATRY 48 (1968); Roth & Appelbaum, *supra* note 199, at 185–87.

[279]Address by M. Gill, "The Boston State Hospital Case: Its Impact on State Hospital Patients," APA Annual Meeting (May 6, 1980).

> tients deteriorated into states of physical emergency in which they resorted to physical assault on others or themselves or became self-destructive through refusal of food to such a degree that it endangered their physical health.[280]

Other studies confirm that the frequent exercise of the right to refuse treatment disrupts other patients, raises markedly the level of tension between staff and patients, and degrades the therapeutic environment.[281]

The current situation—in which an involuntarily committed patient may not be treated if he refuses to be—is unacceptable. The decisions in *Rogers* and *Rennie* also offered inadequate solutions. The *Rennie* court did not explain why referring the matter to a "patient advocate" and then to an "independent" physician improves on the judgments made by the treating physicians. This "solution" permits decisionmaking on pharmacology by a person who may not have medical training. Even an "independent physician" has had limited contact with the patient, and has no continuing responsibility for his care. Thus, the *Rennie* approach would use medication not to treat psychiatric crises and thereby prevent violent emergencies, but only to end them once they have occurred.

The *Rogers* "solution" also is unrealistic. It requires many complex steps before the agitated patient can receive the treatment he needs. As Stone has noted:

> Except in an emergency . . . the patient's right to refuse treatment must be respected and the patient must be presumed to be competent. The only way to overcome the patient's presumptive competence is for the hospital to go to the state court and institute a proceeding to have the patient judged incompetent and a guardian appointed. If the patient is not found incompetent, his refusal of treatment stands. If he is found incompetent and a guardian is appointed, then the decision to accept treatment and overrule the patient is left to the guardian. The judge gave no guidance to such guardians as to what should control their decisions. But the Supreme [Judicial] Court of Massachusetts has made it clear in a series of recent opinions about guardianship that the patient should be consulted about the choice of guardian and that the guardian should pay serious attention to the

---

[280] *Id.*

[281] *See* Appelbaum & Gutheil, *Drug Refusal*, *supra* note 277; Perr, *Effect of the* Rennie *Decision on Private Hospitalization in New Jersey: Two Case Reports*, 138 AM. J. PSYCHIATRY 774 (1981); Schultz, *The Boston State Hospital Case: A Conflict of Civil Liberties and True Liberalism*, 139 AM. J. PSYCHIATRY 183 (1982).

> patient's wishes. These provisions would very likely permit a patient to select a guardian who would uphold his right to refuse treatment and the hospital would have no recourse.[282]

Because the law on this subject is so murky, treatment staffs live in constant fear of lawsuits. It remains unclear when treatment refusals must be accepted and when they may be overriden. As a result, some patients will receive suboptimal treatment, extending their stay and raising the costs of care.[283] More patients will need to be secluded, restrained, or transferred to more secure facilities.[284] The quality of therapy that can be provided to other, nondisruptive patients also will decline. As a result, it will be harder to attract good staffs.[285] These and other factors reveal the right to refuse treatment as a "solution" which creates still larger problems.[286]

The Model Law rejects the reasoning of *Rogers* and *Rennie*. It rejects the notions that the treating staff and the patient's family usually act on motives adverse to the patient's interests, and that court-appointed decisionmakers will improve patient care.[287] A broad right to refuse treatment is in essence the right to refuse any parens patriae commitment. The Model Law does not endorse such a right.

Instead, the Model Law provides for a determination of incompetency prior to commitment, and grants treatment staff the power to treat involuntary patients without their consent. Under a Utah statute that is similar to the Model Law in this respect, a court concluded that "once the person has been afforded all the procedural safeguards . . . the judgment of the treating

---

[282]Stone, *supra* note 197.

[283]*See* Perr, *supra* note 281.

[284]*See* Schultz, *supra* note 281.

[285]*Id.*; *see also* Roth, *The Right to Treatment and the Right to Refuse Treatment*, in PSYCHIATRISTS AND THE LEGAL PROCESS (R. Bonnie ed. 1977).

[286]*See generally* Amarasingham, *Social and Cultural Perspectives on Medication Refusal*, 137 AM. J. PSYCHIATRY 353 (1980); Appelbaum & Gutheil, *Drug Refusal*, *supra* note 277; Appelbaum & Gutheil, *Rotting with Their Rights On*, *supra* note 277; Ford, *The Psychiatrist's Double Bind: The Right to Refuse Medication*, 137 AM. J. PSYCHIATRY 332 (1980); Gutheil, Shapiro & St. Clair, *Legal Guardianship in Drug Refusal: An Illusory Solution*, 137 AM. J. PSYCHIATRY 347 (1980); Reiser, *Refusing Treatment for Mental Illness: Historical and Ethical Dimensions*, 137 AM. J. PSYCHIATRY 329 (1980); Stone, *supra* note 22.

[287]Such an adversarial approach was flatly rejected by the Supreme Court as applied to juvenile proceedings in Parham v. J.R., 442 U.S. 584, 607 (1979) ("'Neither judges nor administrative hearing officers are better qualified than psychiatrists to render psychiatric judgments.'") (quoting *In re* Roger S., 19 Cal. 3d 921, 942, 569 P.2d 1286, 1299 (1977) (Clark, J., dissenting)).

professional physician in accordance with accepted medical standards is a factor of paramount importance."[288]

The Model Law does require the staff to engage in a therapeutic exchange with the patient and discuss the treatment alternatives with him.[289] Despite his inability to give informed consent, the patient may have useful observations or preferences that could be respected.

8.C. *Special Therapies.*

Notwithstanding subsections A and B above, a treatment facility shall not administer aversive therapy, experimental treatment, psychosurgery, or any other special therapy designated by the Department of Mental Health except as provided by law or in regulations promulgated by the Department of Mental Health.

COMMENTARY: It is widely acknowledged that the medical and ethical implications of some therapies are so great that they require special regulation. There is much dispute, however, over which therapies such rules should cover. The theoretical basis for subjecting particular therapies to such rules is also rather confused. Certain therapies are described as particularly "intrusive"—as if they could be rank-ordered on a unitary "intrusiveness" scale.[290] In our view, such analysis is not very meaningful. Is drug therapy more intrusive than group confrontation techniques because it places chemicals in the body, or less intrusive because it requires less time and emotional participation by the patient? Is a single, effective course of electroshock therapy more intrusive than a year of hospitalization? We are not sure. It is clear, however, that under the guise of calibrating the degree to which particular therapies are "intrusive," "unusual," "haz-

---

[288]A.E. and R.R. v. Mitchell, No. 78-466 (D. Utah 1980); *cf.* Winters v. Miller, 446 F.2d 65, 68–71 (2d Cir.), *cert. denied*, 404 U.S. 985 (1971); Gundy v. Pauley, 619 S.W.2d 730, 731–32 (Ky. 1981) (involuntary patient has right to refuse electroshock therapy "in the absence of a judicial declaration of incompetence or an emergency"); Price v. Sheppard, 307 Minn. 250, 259, 239 N.W.2d 905, 911 (1976) ("If [the] interest of the state [in protecting the well-being of its citizens] is sufficiently important to deprive an individual of his physical liberty, it would seem to follow that it would be sufficiently important for the state to assume the treatment decision.").

[289]*Cf. Rennie*, 462 F. Supp. at 1147 ("patients must be informed of and participate in the decision-making aspects of their treatment").

[290]*See* Price v. Sheppard, 307 Minn. 250, 239 N.W.2d 905 (1976); KY. REV. STAT. § 202A.180(7) (1982). A student author claimed to be able to rank various therapies on a scale of "coerciveness." Note, *Conditioning and Other Technologies Used to "Treat?" "Rehabilitate?" "Demolish?" Prisoners and Mental Patients*, 45 S. CAL. L. REV. 616, 619–23 (1972).

ardous," or "drastic,"[291] some ill advised, quasi-clinical judgments have been prematurely ensconced in the law.[292]

This subsection recognizes that while therapies cannot be so easily characterized for all individual patients, nevertheless a social consensus exists that certain therapies have such profound implications that they should be subject to special legal regulation. In many states this category of therapies includes some or all of the following: psychosurgery, electroconvulsive therapy, aversive therapy, and experimental treatment.[293]

Because of the ongoing medical and social debate on the use of such special therapies, we believe it is premature to freeze in the statute rigid criteria on their use.[294] Instead, this provision

---

[291]*See, e.g.*, CAL. WELF. & INST. CODE § 5325.1 (West Supp. 1982) ("hazardous procedures"); ILL. REV. STAT. ch. 91 1/2, § 2-110 (1981) (unusual, hazardous, or experimental services); KY. REV. STAT. § 202A.180(7) (1982) ("intrusive treatments"); OHIO REV. CODE ANN. § 5122.271(A)(5) (Page Supp. 1982) ("unusually hazardous treatment procedures"); OR. REV. STAT. § 426.385(2) (1981) ("potentially unusual or hazardous treatment"); VA. CODE § 37.1-84.1(5) (1976) ("hazardous treatment"); WIS. STAT. § 51.61(1)(k) (1979–1980) ("drastic treatment procedures"); *see also* Bell v. Wayne County Gen. Hosp., 384 F. Supp. 1085, 1100 (E.D. Mich. 1974) ("physically intrusive" therapy); Aden v. Younger, 57 Cal. App. 3d 662, 672, 129 Cal. Rptr. 535, 543 (1976) ("intrusive and hazardous"); Price v. Sheppard, 307 Minn. 250, 239 N.W.2d 905 (1976) ("intrusiveness").

[292]As one psychiatrist wrote to Mr. Stromberg on a draft of the Model Law, "[including] experimental, aversive, convulsive treatments, and psychosurgery under the rubric of intrusive therapies troubles me. Medication is often more troublesome, a longer treatment, and more risky than some of these treatments." (Personal communication on file with C. Stromberg); *see also In re* B., 156 N.J. Super. 231, 383 A.2d 760 (N.J. Super Ct. Law Div. 1977) (holding that psychotropic drugs are not "intrusive" under New Jersey statute); Price v. Sheppard, 307 Minn. 250, 239 N.W.2d 905 (1976) (involuntary patients may not refuse "customary" forms of treatment but may refuse more "intrusive" therapies). *Compare In re* Lundquist, 1 MENTAL DISABILITY L. REP. 190 (P. Ct., Ramsey County, Minn. 1976) (use of prolixin, a psychotropic drug, is intrusive psychiatric treatment), *with In re* Paul Fussa, 1 MENTAL DISABILITY L. REP. 332, 456 (Minn. 1976) (use of prolixin is a customary form of treatment).

[293]*See, e.g.*, ARK. STAT. ANN. §§ 59-1413, 59-1416(7)-(8) (1979); CAL. WELF. & INST. CODE §§ 5325–5326 (West Supp. 1982); CONN. GEN. STAT. § 17-206d(d) (1981); ILL. REV. STAT. ch. 91 1/2, § 2-110 (1981); N.J. STAT. ANN. § 30:4-24.2(d)(2) (West 1981); N.M. STAT. ANN. § 43-1-15(A) (1978); OHIO REV. CODE ANN. § 5122.271 (Page Supp. 1982); *see also* Aden v. Younger, 57 Cal. App. 3d 662, 674, 129 Cal. Rptr. 535, 543 (1976) (right to refuse psychosurgery); Moore v. Wangenstein, 6 CLEARINGHOUSE REV. 172 (D. Minn. 1972) (punishment of patients for refusing ECT challenged as cruel and unusual punishment; relief stipulated to).

[294]The City Council of Berkeley, California, recently voted to go beyond California's law restricting the use of ECT by banning it outright and making administration of ECT a crime. When the APA and others attacked this action as ill considered and illegal, a court enjoined enforcement of the statute. *See* Northern Cal. Psychiatric Soc'y v. City of Berkeley, No. 566778-3 (Cal. Super. Ct. Jan. 13, 1983).

ECT is now used to treat about 80,000 to 100,000 patients annually in the United States. *See generally* Note, *Mental Health: A Model Statute to Regulate the Administration of Therapy Within Mental Health Facilities*, 61 MINN. L. REV. 841 (1977); Note, *Regulation of Electro-Convulsive Therapy*, 75 MICH. L. REV. 362 (1976). An APA task force report summarized the evidence that ECT is effective, and concluded that it is

authorizes each state's Mental Health Department to establish, and revise as appropriate, regulations governing some or all of these special therapies. These regulations should address such topics as: informed consent, characteristics of eligible patients (age, diagnosis, and symptoms prior to treatment), accreditation of personnel and facilities, peer review of the administration of treatment, independent review by nonphysicians, appropriate court supervision (for example, in cases of substituted consent), and follow-up evaluations.[295]

8.D. ***Other Medical/Surgical Treatments.***

Consent for other medical/surgical treatments not intended primarily to treat a patient's mental disorder shall be obtained in accordance with applicable law.

COMMENTARY: No comment is required.

Section 9. PROVISION OF TREATMENT.

9.A. ***General Duty To Provide Treatment.***

Every patient shall be provided with prompt, competent and appropriate treatment that offers him a realistic prospect of improvement. Patients shall be afforded treatment by sufficient numbers of duly qualified personnel in facilities that meet applicable licensing and accreditation standards, that conform to applicable regulations of the Department of Mental Health, and that are able adequately to care for and treat the patients they serve.

COMMENTARY: This provision guarantees every patient a "right to treatment" as described in the Model Law. This right has been the subject of extensive litigation, administrative reform, and scholarly analysis for more than a decade. Lamentably, the confused history of several key cases and the difficulties of implementing both judicial orders and consent decrees have left in some doubt the rationale and exact contours of the right to

---

the most appropriate treatment for some disorders. APA, TASK FORCE REPORT ON ELECTROCONVULSIVE THERAPY (1978).

[295]Certain procedures, such as sterilization and psychosurgery, are so serious that absent specific statutory authority, it has been held that even a substitute decisionmaker such as a guardian should not be permitted to approve them for a patient without court review. *See* Sparkman v. McFarlin, 552 F.2d 172 (7th Cir. 1977), *rev'd and remanded sub nom.* Stump v. Sparkman, 435 U.S. 349 (1978); Holmes v. Powers, 439 S.W.2d 579 (Ky. 1968); Frazier v. Levi, 440 S.W.2d 393 (Tex. Civ. App. 1969); Culver, Ferrell & Green, *ECT and Special Problems of Informed Consent*, 137 AM. J. PSYCHIATRY 586 (1980); Comment, *Sterilization of Mental Defectives: Compulsion and Consent*, 27 BAYLOR L. REV. 174 (1975); Annot., 74 A.L.R.3D 1210 (1976); *see also* PSYCHOSURGERY REPORT, *supra* note 103, at 64.

treatment. In our view, whether or not the right to treatment is a constitutional requirement (and it may well be), it should be regarded as a basic social policy objective.[296]

Respect for the patient's right to treatment, however, need not entail legislating a precise set of staff-to-patient ratios, floor space allocations, medication procedures, or other detailed standards. Injunctions and consent decrees including such features were necessary at one time in order to jolt mental health systems and state legislatures out of their torpor of acquiescence regarding under-funded and disgraceful institutions. In the long run, however, neither judicial decrees nor rigid, overly detailed statutes are the ideal means of resolving treatment issues or running mental health systems.

The Model Law, therefore, seeks to define the right to treatment so as to protect patients' basic rights without unduly constraining the capacity of state mental health departments to respond to changing conditions. The patient is guaranteed "prompt, competent, and appropriate treatment" that offers him a "realistic prospect of improvement"; he is not guaranteed the best conceivable treatment. [297]

This provision also requires that treatment personnel be qualified[298] and that the facilities themselves meet both licensing and accreditation standards.[299] Treatment facilities also must be

---

[296]The authors previously have urged this view. See A. STONE (WITH C. STROMBERG), *supra* note 2, at 83–94; Stromberg, Wyatt v. Stickney *and the Right of Civilly Committed Mental Patients to Adequate Treatment*, 86 HARV. L. REV. 1282 (1973). The APA has supported the concept of the right to treatment. *See Position Statement on the Right to Adequate Care and Treatment for the Mentally Ill and Mentally Retarded*, 134 AM. J. PSYCHIATRY 354 (1977).

[297]*Cf.* Eckerhart v. Hensley, 475 F. Supp. 908, 915, 922 (W.D. Mo. 1979).

[298]Currently, because of poor compensation and very difficult working environments, many public mental hospitals are unable to attract adequately trained staff. A high proportion of staff are foreign-trained and, while many are highly motivated and competent, many others are poorly prepared and have difficulties communicating with patients. *See* Talbott, *The Problems and Potential Roles of the State Mental Hospital*, in STATE MENTAL HOSPITALS: PROBLEMS AND POTENTIALS 21, 23 (J. Talbott ed. 1980); Birnbach, *Back Ward Society, 1981: Implications for Residential Treatment and Staff Training*, 32 HOSP. & COMMUNITY PSYCHIATRY 550, 551 (1981); Knesper & Hirtel, *Strategies to Attract Psychiatrists to State Mental Hospital Work: Results from a Survey of Potential Employees*, 38 ARCHIVES GEN. PSYCHIATRY 1135 (1981); Torrey & Taylor, *Cheap Labor from Poor Nations*, 130 AM. J. PSYCHIATRY 428, 429 (1973); Bower, *Understaffing Severe in State Hospitals Nationwide*, Psychiatric News, Sept. 18, 1981, at 1, col. 1. One study found that foreign graduates constituted 58% of full-time staff in state mental hospitals and that only 43% were fully licensed in their state. J. JENKINS & M. WITKIN, FOREIGN MEDICAL GRADUATES EMPLOYED IN STATE AND COUNTY HOSPITALS (1976) (NIMH Mental Health Statistical Note No. 131); *see also* E.H. v. Matin, 284 S.E.2d 232, 236 (W. Va. 1981) (referring to harm to patients caused by "staff who are under-qualified," and the "amateurish care given to patients").

[299]*Cf.* Joint Comm'n on Accreditation of Hosps., *Standards on Patients' Rights*,

able to care adequately for the group of patients they serve. This may imply a higher standard than the legal minimum set by licensing standards. For example, an institution that serves only nonviolent adult psychiatric patients may have different treatment needs than an institution serving juveniles or violent adults. Rather than setting an absolute minimum, this provision would measure the adequacy of treatment against the standard of meeting the needs of the patients actually served.

9.B. ***Individual Treatment Plan***

1. A written individual treatment plan shall be prepared, with the participation of the patient to the extent he is able, during voluntary admission or emergency psychiatric evaluation or, if a person has been subject to neither, then within seven days of the beginning of a patient's thirty-day commitment. The individual treatment plan shall be approved by the responsible psychiatrist [or: "the responsible physician"],[300] and the course of treatment actually administered shall conform to the plan.

2. The patient's progress in attaining the objectives in the treatment plan shall be noted in his records, and revisions in the plan shall be made as appropriate. The patient and, if the patient desires, the next of kin or guardian, shall be afforded an opportunity to participate in considering any substantial change in the treatment plan.

3. The individual treatment plan shall be available upon request to the patient and to any other person designated by him, provided that the responsible psychiatrist [or: "the responsible physician"][301] may preclude disclosure of the individual treatment plan to the patient or others for a period not to exceed seven days from the request if he states in writing why disclosure would be harmful to the patient.

COMMENTARY: Other than in emergency situations, almost any effective treatment requires an individual treatment plan.[302] Under the Model Law, such a plan should be developed within seven days of voluntary admission or thirty-day commitment.[303] This subsection also encourages a patient's participation to the

---

¶ 14.3.f ("an adequate number of competent, qualified, and experienced professional clinical staff") *reprinted in* APA, RIGHTS OF THE MENTALLY DISABLED: STATEMENTS AND STANDARDS (1981) [hereinafter cited as *Standards on Patients' Rights*].

[300]Optional provision.

[301]Optional provision.

[302]*See* Eckerhart v. Hensley, 475 F. Supp. 908, 921 (W.D. Mo. 1979); Gary W. v. State, 437 F. Supp. 1209, 1226 (E.D. La. 1977); Davis v. Watkins, 384 F. Supp. 1196 (N.D. Ohio 1974); Wyatt v. Stickney, 344 F. Supp. 373 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974); *see also* 42 U.S.C. § 9501(B), (C) (Supp. V 1981); MD. HEALTH–GEN. CODE ANN. § 10-705 (1982); N.M. STAT. ANN. § 43-1-9 (1979); PA. STAT. ANN. tit. 50, § 7107 (Purdon Supp. 1982–1983).

[303]*Compare* ILL. REV. STAT. ch. 91 1/2, § 3-209 (1981) (within three days); OHIO REV. CODE ANN. § 5122.27 (Baldwin 1980) (within 20 days of admission); W. VA. CODE § 27-5-9(d) (1980) (within seven days).

extent of his ability, and grants him (or if he desires, his next of kin or guardian) the right to consider major changes in the treatment plan. It does not confer on the patient a veto or a right of prior approval for every minor alteration in treatment. Commitment necessarily involves placing a person under a physician's professional care, and thus authorizes professional judgments made in accordance with good medical practice.

This subsection encourages a discussion of therapy between the physician and patient. Paragraph 3 provides that while the treatment plan should be available to the patient or persons named by him, the physician may forbid disclosure for up to seven days if it might harm the patient. For example, a paranoid, panicky patient might be severely alarmed by remarks in the plan.[304] To protect against excessive withholding of information, the physician must state the reasons for nondisclosure in the patient's records.

9.C. ***Administration of Medications and Other Treatments.***

1. Medications and other treatments shall be prescribed, ordered, and administered only in conformity with accepted clinical practice. Medication shall be administered only in accordance with the written order of a physician or upon a verbal order noted in the patient's medical record and subsequently signed by the physician. Medication shall be administered only by a qualified physician, qualified nurse, or qualified other persons pursuant to procedures approved by the Department of Mental Health. The attending physician shall review regularly the drug regimen of each resident patient under his care and shall monitor any symptoms of harmful side effects. Prescriptions for psychotropic medications shall be written with a termination date not exceeding thirty days thereafter, but may be renewed.

2. Medications and other treatments shall be administered in accordance with all applicable law.

3. If a patient is given any psychotropic or other medication that has an effective duration of action including the day of a court hearing, the facts concerning its administration and effects and the patient's mental status and behavior in the absence of medication shall be brought to the attention of the court.

COMMENTARY: This subsection regulates the use of medications and other treatments. Although the Model Law generally does not contain such specifics, widespread public concern over the

---

[304] *Accord* ARIZ. REV. STAT. ANN. § 36-507(3) (Supp. 1982–1983); N.M. STAT. ANN. § 43-1-9 (1978). A broader exemption from disclosure was contained in the Mental Health Systems Act, 42 U.S.C. § 9501(1)(I) (Supp. V 1981), but it provided for review of the decision by a health professional chosen by the patient.

use, and alleged overuse and misuse, of psychotropic medications warranted particular rules on this subject.[305] Subsection 3 seeks to balance the patient's right to present himself at a court hearing as he feels appropriate,[306] the treatment facility's obligation to prevent the patient from deteriorating, and the court's need to learn about the patient's actual mental condition.

9.D. ***Other Medical/Surgical Care.***

All patients shall be provided with prompt, regular, and competent medical care for physical ailments under the supervision of a licensed physician. Every patient shall have a reasonably complete physical examination at appropriate intervals.

COMMENTARY: No comment is required.[307]

Section 10. RIGHTS OF PATIENTS.

COMMENTARY: This Section sets forth the rights of patients. Although the "right to treatment" cases generally hold that such rights apply only to involuntary patients (as the constitutional quid pro quo for their deprivation of liberty),[308] the Model Law guarantees these rights to voluntary patients as well. While some state laws grant the administrator of the treatment facility plenary powers to restrict a patient's rights in any manner he believes necessary,[309] the Model Law specifies which rights may be limited and for what reasons.

---

[305]*See generally* Rogers v. Okin, 478 F. Supp. 1342 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982); Rennie v. Klein, 462 F. Supp. 1131 (D.N.J. 1978), *supplemented*, 476 F. Supp. 1294 (D.N.J. 1979), *modified and remanded*, 653 F.2d 836 (3d Cir. 1981), *vacated and remanded*, 102 S. Ct. 3506 (1982); E.H. v. Matin, 284 S.E.2d 232, 236 (W. Va. 1981). For provisions barring the use of "unnecessary" or "excessive" medications, *see, e.g.*, CAL. WELF. & INST. CODE § 5325.1(c) (West Supp. 1982); ME. REV. STAT. ANN. tit. 34, § 2143.8 (1978); N.J. STAT. ANN. § 30:4-24.2(d)(1) (West 1981); N.M. STAT. ANN. § 43-1-6(I) (1978); N.C. GEN. STAT. § 122-55.6 (1981); N.D. CENT. CODE § 25-03.1-40 (1978); OHIO REV. CODE ANN. § 5122.27(F)(6) (Baldwin 1980); TEX. REV. CIV. STAT. art. 5547-300 (Vernon Supp. 1982–1983); WIS. STAT. § 51.61(1)(h) (1979 -1980).

[306]*See* Lessard v. Schmidt, 349 F. Supp. 1078, 1092 (E.D. Wis. 1972) (strong medications impaired patient's ability to present his case).

[307]*See* Rone v. Fireman, 473 F. Supp. 92, 104 (N.D. Ohio 1979); Davis v. Watkins, 384 F. Supp. 1196, 1207 (N.D. Ohio 1974); Wyatt v. Stickney, 344 F. Supp. 373, 380 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974); N.M. STAT. ANN. § 43-1-6(G) (1978); W. VA. CODE § 27-5-9(c) (1980).

[308]*See, e.g.*, Ellen S. v. Rhodes, 507 F. Supp. 734 (S.D. Ohio 1981); Rone v. Fireman, 473 F. Supp. 92 (N.D. Ohio 1979); Gary W. v. Louisiana, 437 F. Supp. 1209 (E.D. La. 1976); Wyatt v. Stickney, 344 F. Supp. 373 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974).

[309]*See, e.g.*, N.J. STAT. ANN. § 30:4-24.2(g)(1) (West 1981) (any right may be denied when "imperative" to do so).

10.A. ***Preservation of Rights.***

No right of any person (including but not limited to the right to register and vote at elections; rights to acquire, use, and dispose of property including contractual rights; rights to sue and be sued; rights relating to licenses, permits, privileges and benefits under law; and rights concerning domestic relations) shall be denied or reduced solely by reason of his having been evaluated, committed, or treated under this Act, except as otherwise specifically provided herein or in other applicable law. A finding of lack of capacity to make an informed decision concerning treatment under Section 6 shall not alone establish lack of competence for any other purpose. A treatment facility may for clinical reasons preclude a patient who is believed to lack competence from making substantial dispositions of his property until his competence to do so can be decided by a court.

COMMENTARY: This subsection provides that evaluation or commitment alone does not justify denying a patient his rights. Some rights, however, such as the right to bear arms, must necessarily be restricted in order to manage facilities safely. Clinical judgments about particular patients may warrant additional restrictions. States may modify the rights of persons with mental disorders by special provisions of law,[310] but such provisions should be developed very cautiously. Caution should be exercised because mental illness alone does not imply an inability to discharge all rights responsibly.[311] Persons who are or have been mentally ill should forfeit rights only when necessary to further some compelling state interest. Rights should be restricted only on the basis of legal provisions related to the person's actual competence to perform particular functions. (Persons found to be dangerous, of course, may be subject to special restrictions.) The final sentence of this provision is necessary to provide clear authority to prevent a possibly incompetent patient from disposing of items of great value before a court can evaluate his competency.

---

[310]*See* ARIZ. REV. STAT. ANN. § 36-506 (1981–1982 & Supp. 1982–1983); CAL. WELF. & INST. CODE § 5325.1 (West Supp. 1982); COLO. REV. STAT. §§ 27-10-104, 27-10-115 (1982); MASS. GEN. LAWS ANN. ch. 123, § 25 (West Supp. 1982–1983); N.J. STAT. ANN. § 30:4-24.2(5) (West 1981); N.M. STAT. ANN. § 43-1-5 (1978); N.Y. MENTAL HYG. LAW § 33.01 (McKinney 1978); OHIO REV. CODE ANN. § 5122.301 (Baldwin 1980); W. VA. CODE § 27-5-9(a) (1980 & Supp. 1982).

[311]*See generally* Davis v. Watkins, 384 F. Supp. 1197, 1206 (N.D. Ohio 1974); Wyatt v. Stickney, 344 F. Supp. 373 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974); Helvey v. Rednour, 86 Ill. App. 3d 154, 408 N.E.2d 17 (1980); Carroll v. Cobb, 139 N.J. Super. 439, 354 A.2d 355 (N.J. Super. Ct. App. Div. 1976); Note, *Mental Disability and the Right to Vote*, 88 YALE L.J. 1644 (1979).

10.B. ***Right to Treatment.***

Patients shall have a right to treatment to the extent provided in section 9 and subsections 10.C. and 10.D.

COMMENTARY: No comment is required.

10.C. ***Healthful and Humane Environment.***

Every patient shall have the right to a healthful and humane environment. Every treatment facility shall provide a clean, sanitary, safe and comfortable environment in a structure that complies with applicable licensing requirements governing physical facilities, nutrition, health and safety, and medical services, and for aspects of care for which there are not mandatory requirements, with generally accepted professional standards. In addition, every patient shall have a right to a humane psychological environment that protects him from harm or abuse, provides reasonable privacy, promotes personal dignity, and provides opportunity for improved functioning.

COMMENTARY: The often disgraceful conditions of public mental institutions have caused much of the legal and social furor over civil commitment in recent years. Thus, the wisdom of having a statutory provision clearly declaring that patients have the right to a decent physical and psychological environment would seem unarguable. The Model Law goes beyond the abstract constitutional right whose contours were gradually developed in the right to treatment cases. It attempts to graft at least some specifics onto the right to treatment concept. It must be noted, however, that the funds to make such this right a reality generally have not been provided.[312]

During the last decade, the rights encompassed in this provision have been defined in varying ways, including the right to "protection from harm,"[313] the right to decent physical surround-

[312]*Compare, e.g.*, New York State Ass'n for Retarded Children (NYSARC) v. Rockefeller, 357 F. Supp. 752 (E.D.N.Y. 1973) (declaring that there is a right to "protection from harm" but that the court could not impose national accreditation standards), *with* NYSARC v. Carey, 393 F. Supp. 715 (E.D.N.Y. 1975) (upholding the right to minimum quality of care and to "least restrictive alternative" placements), *and* NYSARC v. Carey, 596 F.2d 27 (2d Cir. 1979) (requiring the state to provide continued funding for advisory board and finding that adequate funds for alternative facilities had not been supplied). *See generally* Lottman, *Enforcement of Judicial Decrees: Now Comes the Hard Part*, 1 MENTAL DISABILITY L. REP. 69 (1976); Special Project, *The Remedial Process in Institutional Reform Litigation*, 78 COLUM. L. REV. 784 (1978); Note, *Implementation Problems in Institutional Reform Litigation*, 91 HARV. L. REV. 428 (1977); Note, *The* Wyatt *Case: Implementation of a Judicial Decree Ordering Institutional Change*, 84 YALE L.J. 1338 (1975).

[313]*See, e.g.*, NYSARC v. Rockefeller, 357 F. Supp. 752 (E.D.N.Y. 1973); 42 U.S.C. § 9501(1)(G) (Supp. V 1981); ARK. STAT. ANN. § 59-1416(23) (Supp. 1979); CAL. WELF. & INST. CODE § 5325.1(i) (West Supp. 1982) ("right to be free from hazardous proce-

ings,[314] the "right to treatment,"[315] and the right to "privacy, human dignity [and] a positive psychological environment which enhances the patient's self worth."[316] All of these elements are important and are embodied in the Model Law provision. However, none of these rights can be elevated to an absolute right without sacrificing other interests. In particular, tradeoffs must be made among the rights to privacy, treatment, and protection from harm. For example, to protect one patient from harm, it may be necessary to "invade" another patient's privacy by involuntarily administering medication to him to reduce his assaultive tendencies. To treat a patient effectively may mean permitting him to progress toward less and less restrictive settings, thereby increasing the chances that he will impair the privacy or dignity of others. Such tradeoffs ultimately must be guided by experience and by humane clinical judgment. The Model Law seeks to protect the basic interests of each patient, without imposing impossible obligations on mental health departments.[317] It also provides that if the Joint Commission on

---

dures"); *cf.* Youngberg v. Romeo, 102 S. Ct. 2452 (1982); Harper v. Cserr, 544 F.2d 1121 (1st Cir. 1976).

[314]*See, e.g.*, Eckerhart v. Hensley, 475 F. Supp. 908 (W.D. Mo. 1979); Wyatt v. Stickney, 344 F. Supp. 373 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974); N.M. STAT. ANN. § 43-1-6(D) (1978) (right to "clean, safe and comfortable environment").

[315]*See, e.g.*, Scott v. Plante, 532 F.2d 939, 947 (3d Cir. 1976); Welsch v. Likins, 373 F. Supp. 487, 491–97 (D. Minn. 1974), *aff'd*, 550 F.2d 1122, 1125 (8th Cir. 1977); Wyatt v. Stickney, 344 F. Supp. 373 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974); COLO. REV. STAT. § 27-10-116 (1982); FLA. STAT. ANN. § 394.459(2) (West Supp. 1983); N.M. STAT. ANN. § 43-1-7 (1978); N.Y. MENTAL HYG. LAW § 33.03 (McKinney 1978); UTAH CODE ANN. § 64-7-46 (1978 & Supp. 1981) (rights to "humane care and treatment" and "medical care and treatment in accordance with prevailing standards accepted in medical practice"); W. VA. CODE § 27-5-9(b) (1980); *cf.* Ohlinger v. Watson, 652 F.2d 775 (9th Cir. 1980) (two prison inmates granted the right to treatment).

[316]Rone v. Fireman, 473 F. Supp. 92, 121 (N.D. Ohio 1979); *see also* Goodman v. Parwatikar, 570 F.2d 801 (8th Cir. 1978); Harper v. Cserr, 544 F.2d 1121 (1st Cir. 1976); Gary W. v. Louisiana, 437 F. Supp. 1209 (E.D. La. 1976); Davis v. Watkins, 384 F. Supp. 1196 (N.D. Ohio 1974).

Mills reports that in the last decade, about 30 consent decrees have been entered in right to treatment cases, governing care in about 15 states. Mills, *The Right to Treatment: Little Law but Much Impact*, in PSYCHIATRY 1982, 361, 369 (L. Grinspoon ed. 1983). About 15 states protect the right to privacy and dignity by statute. *See, e.g.*, ARK. STAT. ANN. § 59-1416(9) (Supp. 1979) (right "to be respected as an individual with dignity and unique value"); CAL. WELF. & INST. CODE § 5325.1(b) (West Supp. 1982) ("dignity, privacy and humane care"); DEL. CODE ANN. tit. 16, § 5161(a)(1) (Supp. 1980) (dignity and personal integrity); N.J. STAT. ANN. § 30:4-24.2(e)(1) (West 1981) ("privacy and dignity"); N.M. STAT. ANN. § 43-1-6(D) (1978) ("humane psychological and physical environment"); OHIO REV. CODE ANN. §§ 5122.27(F)(3), 5122.29(B) (Baldwin 1980) ("humane psychological and physical environment," and "treated with . . . dignity"); *see also Standards on Patients' Rights*, *supra* note 299, ¶¶ 14.2.2, 14.2.4.

[317]*Cf.* Davis v. Hubbard, 506 F. Supp. 915 (N.D. Ohio 1980) (hospital's physical plant

Accreditation of Hospitals or a licensing authority has not established mandatory requirements for a particular aspect of care, treatment facilities nevertheless must meet generally accepted professional standards.

10.D. ***Least Restrictive Alternative and Leaves of Absence.***

1. Every patient shall have the right to treatment consistent with the least restrictive alternative principle.

2. Leaves of absence may be granted in appropriate cases at the discretion of the treatment facility. Police officers shall be authorized to and shall, at the request of a treatment facility, take into custody and return to the treatment facility any person who has been committed there and leaves without proper authorization or does not return at the end of an authorized leave of absence.

COMMENTARY: This provision ensures treatment "consistent with the least restrictive alternative principle," and encourages use of leaves of absence under appropriate safeguards.[318]

10.E. ***Institutional Labor.***

1. Patients shall have the right to perform labor as part of a therapeutic program.

2. Patients may not be required to perform labor, except that to the extent they are able, they may be required to perform (a) tasks necessary to care for their personal possessions, (b) routine, nondegrading housekeeping tasks necessary to maintain their living quarters, or (c) other tasks which the responsible psychiatrist [or: "the responsible physician"][319] approves and which are monitored as part of a therapeutic program for the patient. No patient shall be subjected to any loss of any right under this Act (as distinguished from a privilege that is conferred as part of a therapeutic program) because of his refusal to perform such tasks.

3. Any patient labor that confers an economic benefit upon the institution beyond merely supplementing employee performance of housekeeping tasks shall be compensated on a reasonable basis in accordance

---

would have to be totally reconstructed). For example, several consent decrees in right to treatment cases are expected to cost Massachusetts approximately $530 million in 1983. Despite this cost, there is concern that the emphasis is on "cosmetic changes in physical plant, changes that ought to be subordinate to substantive issues of treatment," and that the result of such decrees is accelerated "abandonment" of patients to haphazard outpatient care. Mills, *supra* note 316, at 368.

[318]*See also* 42 U.S.C. § 9501(1)(A) (Supp. V 1981) (right to treatment that "restrict[s] such liberty only to the extent necessary consistent with such person's treatment needs, applicable requirements of law, and applicable judicial orders"). Approximately half the states' laws provide for periodic review of a patient's condition and the advisability of discharge to a less restrictive facility. For an extreme example, *see* N.M. STAT. ANN. § 43-1-3(D) (1978). Concerning leaves of absence, *see generally* Interest of Henderson, 44 Colo. App. 102, 610 P.2d 1350 (1980); ARK. STAT. ANN. § 59-1418 (Supp. 1979).

[319]Optional provision.

with applicable law, and the proceeds of such labor shall be paid to the patient or his designee.

COMMENTARY: Institutional labor (sometimes denoted by the emotionally charged term "peonage") has been the subject of a hotly contested debate.[320] Although labor may be a useful part of the therapy program,[321] certainly a patient should not be exploited in any way.

The Model Law seeks to reconcile these sometimes conflicting considerations. Initially, it affirms a patient's right to perform work that is related to a written therapeutic program.[322] It then provides that a patient may be required to do certain tasks, some of which cannot seriously be characterized as labor (for example, making his bed or putting away his possessions), and some of which may be work that is inherent in a communal living arrangement (for example, vacuuming).[323] The Model Law, however, requires that such tasks be "nondegrading."

---

[320]*See, e.g.*, Bartlett, *Institutional Peonage: Our Exploitation of Mental Patients*, ATLANTIC MONTHLY, July 1964, at 116; Friedman, *The Mentally Handicapped Citizen and Institutional Labor*, 87 HARV. L. REV. 567 (1974); Lebar, *Worker-Patients: Receiving Therapy or Suffering Peonage*, 62 A.B.A. J. 219 (1976). A 1972 survey found that one in five residents of institutions for the mentally handicapped was engaged in a work program. *See* 4 PRESIDENT'S COMM'N ON MENTAL HEALTH, TASK PANEL REPORT ON LEGAL AND ETHICAL ISSUES 1385 (1978) [hereinafter cited as LEGAL AND ETHICAL ISSUES REPORT].

[321]*See, e.g.*, N.Y. MENTAL HYG. LAW § 33.09 (McKinney 1978); Schwartz, *Expanding a Sheltered Workshop To Replace Nonpaying Patient Jobs*, 27 HOSP. & COMMUNITY PSYCHIATRY 98 (1976); Wilder, *The Case for a Flexible Long Term Sheltered Workshop for Psychiatric Patients*, 27 HOSP. & COMMUNITY PSYCHIATRY 112 (1976).

[322]"Work therapy" was developed initially as a kind of antidote to the social decompensation suffered by those subject to long-term confinement in large institutions. The Model Law, in contrast, contemplates short-term commitment and a return to the community for almost all patients. Thus, the issue of patient labor must be considered in a somewhat different context than in the past. As to a patient's right to perform therapeutic work, *see* Wyatt v. Stickney, 344 F. Supp. 372 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974); ARIZ. STAT. ANN. § 36-510 (1974 & Supp. 1981–1982); MASS. GEN. LAWS ANN. ch. 123, § 29 (West Supp. 1982–1983). Compare *Standards on Patients' Rights*, *supra* note 299, at ¶ 14.6 (Patients may perform labor if the work "is part of the individual treatment plan" and is "performed voluntarily," if "the patient receives wages commensurate with the economic value of the work," and if the work project complies with all applicable laws.).

[323]*Cf.* Jobson v. Henne, 355 F.2d 129 (2d Cir. 1966); Davis v. Balson, 461 F. Supp. 842 (N.D. Ohio 1978); Wyatt v. Stickney, 344 F. Supp. 373 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974); Dale v. State, 44 A.D.2d 384, 355 N.Y.S.2d 485, *aff'd*, 36 N.Y.2d 833, 331 N.E.2d 686, 370 N.Y.S.2d 906 (1974); ILL. REV. STAT. ch. 91 1/2, § 2-106 (1981) (general housekeeping without compensation is allowed); *Standards on Patients' Rights*, *supra* note 299, at ¶ 14.6.1 ("A patient may be required to perform personal housekeeping tasks without compensation."). *But see* OHIO REV. CODE ANN. § 5122.28 (Baldwin 1980) (This statute is unusual in that it appears to provide that patients cannot be compelled to perform tasks related to the operation, support, or maintenance of the hospital, but can be compelled to perform other labor.).

Even unpleasant tasks such as cleaning the ward bathroom should not be deemed inherently degrading if responsibility rotates fairly among the patients in a living unit. On the other hand, if, as a disciplinary measure, one patient is repeatedly required to perform this duty, the task becomes degrading. Although a patient is not exempt from all tasks under the Model Law, this provision ensures that he will not be treated in an unfair way or in a way that impairs his self-esteem and rehabilitation. Other types of tasks that patients may be required to perform are those that are approved and monitored as part of the therapeutic program of the patient. It is a mere semantic point whether those tasks are considered therapy or labor.

While a patient may be required to perform therapeutic labor under paragraph 10.E.1., he may not be punished or deprived of any right because of his refusal to do so.[324] Compliance nevertheless may occur because patients, like other persons, often abide by social requirements for reasons other than the existence of sanctions. Also, while patients may not be deprived of any rights, privileges (such as extra desserts or minor tokens) may be granted to those who do perform required labor.[325] Encouraging therapeutic labor of this nature is proper. All state mental health systems operate on inadequate funds. These limited funds are better spent on treatment programs than on large numbers of housekeeping personnel.

The Model Law also permits voluntary labor that is an appropriate part of a treatment program. It does not permit nontherapeutic voluntary labor because it is difficult or impossible to eliminate the possibility of coercion in the institutional context. To be permissible, labor must have some appreciable therapeutic benefit. It need not be the only plausible therapy which could be provided.

Finally, the Model Law requires that the patient receive reasonable payment for any labor that confers an economic benefit on the institution beyond merely supplementing employee work.[326] Thus, if patients voluntarily work in a vegetable patch that merely supplies the hospital kitchen, and the work is ther-

---

[324]*Cf.* Parks v. Ciccone, 281 F. Supp. 805, 811 (W.D. Mo. 1968).

[325]*But see* Dale v. State, 44 A.D.2d 384, 389, 355 N.Y.S.2d 485, 490 (1974), *aff'd*, 36 N.Y.2d 833, 331 N.E.2d 686, 370 N.Y.S.2d 906 (1975).

[326]It generally is agreed that compensation increases the therapeutic value of labor. *See* Safier & Barnum, *Patient Rehabilitation Through Hospital Work Under Fair Labor Standards*, 26 HOSP. & COMMUNITY PSYCHIATRY 299, 302 (1975).

apeutic, it is permissible and need not be compensated. On the other hand, if the garden is worked so extensively that it produces cash crops which are sold at a profit, then the patient must be paid in accordance with law.[327] The payment must be made to the patient and cannot be applied directly to the costs of care.

10.F. ***Restraints and Seclusion.***

1. Restraints and seclusion may be of therapeutic benefit to some patients and therefore may be administered in conformity with good medical practice.

2. Every patient shall have the right to be free from unwarranted or inappropriate restraints or seclusion.

3. A patient may be physically restrained or placed in seclusion only at the written order of a physician or upon a verbal order noted in the patient's record and subsequently signed by the physician.

4. During any period in which a patient is restrained or secluded, he shall be checked periodically and cared for properly to assure his well-being.

COMMENTARY: The Model Law sets out only a general framework for detailed regulations to be developed in each state regarding restraints and seclusion.[328] Their use has become an issue of increasing importance as the incidence of violence in

---

[327]*Cf.* ILL. REV. STAT. ch. 91 1/2, § 2-106 (1981) (A patient "who performs labor which is of any consequential economic benefit to [the institution] shall receive wages which are commensurate with the value of the work performed in accordance with . . . laws."); N.Y. MENTAL HYG. LAW § 33.09 (McKinney 1978); Mass. Admin. Reg. tit. 104, § 14.03(7) (patient must be compensated for labor "to the extent of its economic value").

In Souder v. Brennan, 367 F. Supp. 808 (D.D.C. 1973), a case involving the most repulsive abuses of patients, the court declared that the minimum wage laws applied to work by mental patients. *See also* Wyatt v. Stickney, 344 F. Supp. 373, 381 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974). The decision aroused concern that it would end all patient labor activities, many of which were popular and beneficial. However, in National League of Cities v. Usery, 426 U.S. 833 (1976), the Supreme Court struck down similar requirements as an unconstitutional infringement on state governments' ability to control their budgets. Meanwhile, the Department of Labor had promulgated regulations permitting compensation to institutionalized mental patients at rates somewhat below the usual minimum wage, based on their actual productivity. 40 Fed. Reg. 5775, 5777 (1975). These regulations (29 C.F.R. § 529.4 (1982)) have never really been put into effect because of the uncertainty arising from the legal decisions. Thus, the required compensation for patient labor is an unsettled issue. *See* Weidenfeller v. Kidulis, 380 F. Supp. 445 (E.D. Wis. 1974); Note, *Institutionalized Patient Workers and Their Right to Compensation in the Aftermath of* National League of Cities v. Usery, 22 B.C.L. REV. 511 (1981). The Model Law does not resolve this issue. It simply requires compensation "in accordance with law."

[328]The APA is now developing detailed guidelines concerning restraints and seclusion. In our review of this subject we are indebted to Dr. Kenneth Tardiff, who shared with us the draft report on the use of seclusion and restraints of the APA task force that he chairs. Tardiff Progress Report, *supra* note 29.

psychiatric facilities has risen.[329] Whereas violent patients used to be treated primarily in public mental hospitals, because of deinstitutionalization, many violent patients now are found in emergency rooms, general hospital psychiatric wards, and other treatment facilities.

Restraints and seclusion, when properly used, are "highly respected form[s] of treatment, of great value to many severely disturbed patients and essential to the preservation of order and safety during psychiatric emergencies."[330] Both are useful in treating a variety of mental health problems, including cases involving "violent behavior [or] behavior disruptive of the therapeutic environment,"[331] "physical attack[s] on staff" or other patients,[332] and "agitated, uncontrolled behavior."[333]

Because of the relative paucity of quantitative studies, it is difficult to assess whether restraints and seclusion are used appropriately in most cases.[334] As a result, few generalizations can be made about the best use of these techniques. One premise that clearly is sound is that restraints and seclusion should be used only to prevent harm to the patient or to others.

Several states have statutory provisions governing restraints

---

[329]One study of full- and part-time faculty of a university teaching hospital found that 40% of them had been assaulted by patients. *See* Madden, Lion & Penna, *Assault on Psychiatrists by Patients*, 133 AM. J. OF PSYCHIATRY 422 (1976). Another study of suicidal patients in New York found that 7% of the female patients and 14% of the male patients committed assaults during their stay. Tardiff & Sweillam, *Factors Related to Increased Risk of Assaultive Behavior in Suicidal Patients*, 62 ACTA PSYCHIATRICA SCANDINAVICA 63 (1980). Unpublished data from the Veterans Administration for a five-year period reported more than 12,000 assaults. Tardiff Progress Report, *supra* note 29. And a study in a single public hospital over an eight-month period showed 379 cases of violence. Depp, *Violent Behavior Patterns in Psychiatric Wards*, in 2 AGGRESSIVE BEHAVIOR 295–306 (1976); *see also* Lanza, *The Reactions of Nursing Staff to Physical Assault by a Patient*, 34 HOSP. & COMMUNITY PSYCHIATRY 44 (1983); Petrie, Lawson & Hollender, *Violence in Geriatric Patients*, 248 J. A.M.A. 443 (1982).

[330]Amicus Brief of the Mass. Psychiatric Ass'n No. 79-1649, in Rogers v. Okin, 634 F.2d 650 (1st Cir. 1980); *see also* Gutheil, *Observations on the Theoretical Bases for Seclusion of the Psychiatric Inpatient*, 135 AM. J. PSYCHIATRY 325 (1978).

[331]Mattson & Sacks, *Seclusion: Use and Complications*, 135 AM. J. PSYCHIATRY 1210 (1978).

[332]Soloff & Turner, *Patterns of Seclusion: A Prospective Study*, 169 J. NERVOUS & MENTAL DISEASE 37, 38 (1981).

[333]Plutchik, Karasu, Conte, Siegel & Jerrett, *Toward a Rationale for the Seclusion Process*, 166 J. NERVOUS & MENTAL DISEASE 571, 573 (1978). *See generally* Rogers v. Okin, 478 F. Supp. 1342, 1376–81 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct 2442 (1982); Eckerhart v. Hensley, 475 F. Supp. 908, 926 (W.D. Mo. 1979).

[334]This assessment is particularly difficult because the definition, mode, duration, and clinical bases of "seclusion" and "restraint" vary widely in the existing studies. The techniques used under these rubrics also vary from "time out" from therapy, or separate rooms, to camisoles, primitive mechanical restraints, and "chemical restraints."

and seclusion. Some of these provisions are general,[335] while others are quite detailed.[336] More than twenty states have general regulations governing restraints and seclusion, and about twenty more states permit each institution to establish guidelines. A recent review of these rules suggests, however, the need for considerably more detail.[337]

Several courts—in cases that have involved some truly egregious practices—have declared that the excessive use of restraints or seclusion violates the Constitution.[338] In *Youngberg v. Romeo*,[339] the Supreme Court addressed the issue. Romeo, a profoundly retarded man, had been placed repeatedly in seclusion and restraints, including mechanical "shackles." While confined, he attacked, and was attacked by, other patients. The Supreme Court concluded that Romeo had "constitutionally protected liberty interests" in "safe conditions," "freedom from bodily restraint," and "minimally adequate or reasonable training to ensure safety and freedom from undue restraint."[340] The Court also recognized that the rights to "safety," and to "freedom from bodily restraint . . . are not absolute; indeed to some extent they are in conflict."[341] In some instances a patient must be restrained in order to ensure the safety of others. Accordingly, the Court declared that "[i]t is not appropriate for the courts to specify which of several professionally acceptable choices should have been made."[342] Such a professional decision is "presumptively valid," and liability can be imposed only if the professional's decision "is such a substantial departure from accepted professional judgment, practice, or standards as to

---

[335]*See, e.g.*, ARIZ. STAT. ANN. § 36-513 (1974 & Supp. 1982–1983); CAL. WELF. & INST. CODE § 5325.1(c) (West Supp. 1982); OHIO REV. CODE ANN. § 5122.27(f)(7) (Baldwin 1980); UTAH CODE ANN. § 64-7-47 (1978 & Supp. 1981); *cf.* 42 U.S.C. § 9501(1)(F) (Supp. V 1981).

[336]*See, e.g.*, ILL. REV. STAT. ch. 91 1/2, §§ 2-108–2-109, 2-201 (1981); N.J. STAT. ANN. § 30:4-24.2(d)(3) (West 1981); N.Y. MENTAL HYG. LAW § 33.04 (McKinney 1978 & Supp. 1982–1983).

[337]Tardiff Progress Report, *supra* note 29.

[338]*See, e.g.*, Halderman v. Pennhurst State School & Hosp., 612 F.2d 84 (3d Cir. 1979), *rev'd on other grounds*, 451 U.S. 1 (1981); Davis v. Hubbard, 506 F. Supp. 915 (N.D. Ohio 1980); Burks v. Teasdale, 492 F. Supp. 650 (W.D. Mo. 1980); Rogers v. Okin, 478 F. Supp. 1342, 1374 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982); Eckerhart v. Hensley, 475 F. Supp. 908, 925–28 (W.D. Mo. 1979); Wyatt v. Stickney, 344 F. Supp. 373, 380 (M.D. Ala. 1973), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974).

[339]102 S. Ct. 2452 (1982).

[340]*Id.* at 2454.

[341]*Id.* at 2460.

[342]*Id.* at 2461 (quoting Romeo v. Youngberg, 644 F.2d 147, 178 (3d Cir. 1980)).

demonstrate that the person responsible did not base the decision on such a judgment."[343]

While this decision set some parameters for judicial review, it did not establish clinical rules on the use of restraints and seclusion. Both the *Romeo* decision and the Model Law strongly suggest that these standards be set by professional bodies, and that conduct be reviewed if abuses transcend technical professional judgments.

10.G. ***Corporal Punishment.***

Every patient shall have the right to be free from corporal punishment.

COMMENTARY: No comment is required.[344]

10.H. ***Nutrition.***

Every patient shall have the right to a nutritionally sound and medically appropriate diet.

COMMENTARY: The right to an appropriate diet has been recognized in several states.[345] In defining this right, a pragmatic, flexible approach should be taken. In contrast, the New Mexico law confers a right to a "varied" diet, even though some patients may need restricted diets for medical reasons, and requires an "appetizing" diet, which although desirable, seems far too vague to constitute an enforceable right.[346] The Model Law attempts to accommodate the unique needs of each patient by requiring a "nutritionally sound and medically appropriate" diet, while placing requirements on the treatment facility which are realistic. In general, the food should be decent, wholesome, palatable, and varied to the extent medically appropriate. Further specifics were not thought appropriate for statutory definition.

10.I. ***Exercise and Recreation.***

Every patient shall have reasonable opportunities for physical and outdoor exercise and access to recreational areas and equipment. Reasonable limitations may be set by general rules or, for clinical reasons, in particular cases.

---

[343]*Id.* at 2462.

[344]*See* Davis v. Watkins, 384 F. Supp. 1196, 1206 (N.D. Ohio 1974) ("The institution shall prohibit corporal punishment, mistreatment, neglect or abuse in any form, of any patient.").

[345]*See, e.g., id.*; Wyatt v. Stickney, 344 F. Supp. 373 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974); N.M. STAT. ANN. § 43-1-6(F) (1979).

[346]N.M. STAT. ANN. § 43-1-6(F) (1979).

COMMENTARY: The Model Law, like several state statutes,[347] and cases,[348] recognizes the patient's right to reasonable exercise and recreation. The provision permits limitation of the right "by general rules or, for clinical reasons, in particular cases." This allows flexibility in light of the location, physical plant, and patient population of particular treatment facilities. Special limitations also might be imposed on an individual, such as a violent patient who insisted on hitting teammates with his baseball bat. Such limitations must be "reasonable," and if applied to particular individuals, must be justified by medical reasons.

10.J. ***Visitors.***

Every patient shall have the right to receive visitors of his choosing with reasonable privacy. Reasonable limitations on access of visitors may be set by general rules or, for clinical reasons, in particular cases.

COMMENTARY: The right to receive visitors also has been recognized in a number of state laws[349] and court decisions.[350] This right includes the right to be examined by a privately retained physician.[351] While institutional life necessarily limits privacy, the Model Law provides for reasonable privacy during visits.

10.K. ***Communications.***

1. Every patient shall have the right to send and receive mail. Reasonable rules governing inspection (but not reading) of incoming mail may be enforced, provided that they are necessary to substantial health care purposes and that they preserve the patient's rights of privacy to the extent compatible with his clinical status.

2. Every patient shall have the right to reasonably private access to telephones, including the right to make long-distance calls to the extent he can arrange for payment for such calls.

---

[347]*See, e.g.*, ARK. STAT. ANN. § 59-1416(26) (1971 & Supp. 1979); N.J. STAT. ANN. § 30:4.24.2(e)(8)–(9) (West 1981); N.M. STAT. ANN. § 43-1-6(E) (1979).

[348]Davis v. Watkins, 384 F. Supp. 1196, 1208 (N.D. Ohio 1974); Wyatt v. Stickney, 344 F. Supp. 373, 381 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974).

[349]*See, e.g.*, ARK. STAT. ANN. § 59-1416(4) (1971 & Supp. 1979) ("right to see visitors each day"); CAL. WELF. & INST. CODE § 5325(c) (West 1972 & Supp. 1981) (right to "see visitors each day"); N.M. STAT. ANN. § 43-1-6(A) (1979); N.Y. MENTAL HYG. LAW § 33.05 (McKinney 1978); OHIO REV. CODE ANN. § 5122.29(D)(1) (Page 1981). *But see Standards on Patients' Rights, supra* note 299, ¶ 142.4 (right to receive visitors "unless such visits are clinically contraindicated").

[350]*See, e.g.*, Schmidt v. Schubert, 422 F. Supp. 57 (E.D. Wis. 1976); Eckerhart v. Hensley, 475 F. Supp. 908 (W.D. Mo. 1975); Wyatt v. Stickney, 344 F. Supp. 373, 379 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974).

[351]*See, e.g.*, ARIZ. REV. STAT. ANN. § 36-514(1) (1974 & Supp. 1982); ARK. STAT. ANN. § 54-1416(17) (1971 & Supp. 1979).

3. A treatment facility shall provide reasonable assistance to patients in exercising their communication rights. Reasonable limitations on use of the mails and telephones may be set by general rules or, for clinical reasons, in particular cases.

COMMENTARY: The patient's right to send and receive mail and to make telephone calls is well established.[352] However, in some circumstances, these rights must defer to the needs of institutional management. Most states that provide for such rights also authorize some limitations.[353] In addition, development of the fluoroscope and related technologies has made it possible to examine incoming mail for contraband without reading it, and this is done in many prisons.[354] The Model Law ensures a patient's basic communication rights, while permitting limitations where necessary in light of "substantial health care purposes," or a patient's "clinical status." For example, a patient's mail might be inspected to ensure that he does not receive foods that might cause an adverse medical reaction. Similar, reasonable limitations may be imposed on outgoing mail or calls. For example, a patient should not be permitted repeatedly to make obscene or threatening telephone calls, or to send letters with such content.[355] A patient may make long-distance calls only to the extent to which he can arrange payment for such calls; this simply recognizes the limits of institutional resources.[356] Finally, this provision imposes an affirmative duty on institutions to help patients exercise their communication rights.

---

[352]*See, e.g.*, Gerrard v. Blackmun, 401 F. Supp. 1189 (N.D. Ill. 1975); Davis v. Watkins, 384 F. Supp. 1196, 1207 (N.D. Ohio 1974); Brown v. Schubert, 347 F. Supp. 1232 (E.D. Wis. 1974), *supplemented*, 389 F. Supp. 281 (E.D. Wis. 1975); Wyatt v. Stickney, 344 F. Supp. 373, 379 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974); CAL. WELF. & INST. CODE § 5325(d) (West 1972 & Supp. 1981); ILL. REV. STAT. ch. 91 1/2, § 2-103 (1981) (rights to "unimpeded, private, and uncensored" communications); N.M. STAT. ANN. § 43-1-6(B) (1978); OHIO REV. CODE ANN. § 5122.29(C)–(D) (Page 1981).

[353]*See, e.g.*, ARIZ. REV. STAT. ANN. § 36.514 (1974 & Supp. 1982); N.J. STAT. ANN. § 30:4-24.2(e)(6) (West 1981); N.Y. MENTAL HYG. LAW § 33.05 (McKinney 1978); *see also Standards on Patients' Rights, supra* note 299, ¶ 142.4.4–4.6 (right to communicate freely "unless clinically contraindicated").

[354]*See generally* Wolff v. McDonnell, 418 U.S. 539, 577 (1974) (procedures for opening inmates' prison mail); Eckerhart v. Hensley, 475 F. Supp. 908, 924 (W.D. Mo. 1979) (checking patients' mail for contraband without reading it); Davis v. Balson, 461 F. Supp. 842 (N.D. Ohio 1978) (presence of patient is necessary to ensure contents are not read).

[355]*See* ILL. REV. STAT. ch. 91 1/2, § 2-103(c) (1981); MASS. GEN. LAWS ANN. ch. 123, § 23 (West Supp. 1982); N.J. STAT. ANN. § 30:4-24.2(e)(6) (1981) ("reasonable access" to mail and phones); N.M. STAT. ANN. § 43-1-6(B) (1979) (physician may "for good cause" limit such rights).

[356]*Cf.* CAL. WELF. & INST. CODE § 5326 (West 1972 & Supp. 1981). Of course, special provisions should be made for the needs of indigent patients.

10.L. ***Practice of Religion.***

Every patient shall have the right to practice or refrain from practicing a religion, and pressure shall in no event be placed on those who do not wish to practice a religion. The treatment facility shall provide appropriate assistance so that patients wishing to practice a religion have a reasonable opportunity to do so.

COMMENTARY: The right to practice or refrain from practicing religion is a fundamental human and civil right. This subsection provides broad protection for patients' religous rights, as do several state laws.[357] While a treatment facility must provide a "reasonable opportunity" for patients to practice religion, and can justify restraining a patient only in the most compelling circumstances, the Model Law does not preclude all restraints on the exercise of religion. For example, religious belief might not justify smoking peyote or chanting all night on the ward floor. However, these would be rare and extreme cases.[358]

10.M. ***Personal Possessions.***

Every patient shall have the right to keep, use and store personal possessions and to maintain and use bank accounts or other sources of personal funds, unless precluded from doing so by order of a court. Reasonable limitations may be set by general rules or, for clinical reasons, in particular cases.

COMMENTARY: This provision ensures the patient's right to keep personal possessions, with reasonable limitations.[359] For example, the right to keep "personal possessions" in a treatment facility might not include a pet, a gun, or a collection of 6,000 records. In particular, limitations may be necessary on the use of personal funds. A patient should have the power to draw on his personal accounts for miscellaneous items without seeking the approval of anyone. However, a patient who suffers from a severe mental disorder and lacks the capacity to make reasoned decisions on various matters should not necessarily be allowed to convey his entire life savings to a visitor or to a charity on

---

[357]*See, e.g.*, ARIZ. REV. STAT. ANN. § 36-514(4) (1974 & Supp. 1982); N.J. STAT. ANN. § 30:4-24.2(e)(11) (West 1981); N.M. STAT. ANN. § 43-1-6(c) (1979); OHIO REV. CODE ANN. § 5122.29(H) (Page 1981).

[358]We do not address the unique and conflicting considerations in cases of "deprogramming" adherents of religious cults.

[359]*Cf.* ARIZ. REV. STAT. ANN. § 36-507(4)-(5) (1974 & Supp. 1982); CAL. WELF. & INST. CODE § 5325(a)-(b) (West 1972 & Supp. 1982); MASS. GEN. LAWS ANN. ch. 123, § 23 (West Supp. 1982); OHIO REV. CODE ANN. § 5122.29 (Page 1981). Several state laws recognize that reasonable restrictions are proper. *See, e.g.*, ILL. REV. STAT. ch. 91 1/2, § 2-104 (1981); N.Y. MENTAL HYG. LAW § 33.07 (McKinney 1978).

one sunny afternoon. The Model Law authorizes the treatment facility to intervene in such circumstances to prevent major dispositions of property until a court can determine the patient's competency to make such decisions and, if necessary, appoint a guardian or conservator for him.

10.N. ***Notice of Rights.***

As soon after admission as his medical condition permits, a patient shall be advised orally and given a written statement of his rights under this Act, and such a statement of rights shall be posted so that it is available to patients.

COMMENTARY: No comment is required.[360]

10.O. ***Non-Retaliation.***

No patient shall be retaliated against or subjected to any adverse change of conditions or treatment solely because of his having asserted his rights.

COMMENTARY: No patient should be punished for exercising a legal right. The Model Law defines punishment very broadly as "retaliat[ion] against or subject[ion] to any adverse change of conditions or treatment," so as to ensure that patients are protected. However, this provision must not be construed too broadly. If a patient's mode of expressing his rights causes serious harm to others, remedial steps may be taken. For example, if a patient has a right to exercise, but he attacks someone every time he engages in sports, he may be deprived of this right temporarily.

10.P. ***Access to Counsel.***

A patient may at any time have a telephone conversation with or be visited by his lawyer.

COMMENTARY: Because access to counsel is often critical to a patient's ability to be advised of and to enforce his rights, this subsection provides the right to consult with a lawyer virtually at any time. This subsection takes precedence over provisions in the Model Law that might otherwise limit this right, such as subsections 10.J. and 10.K.

---

[360]*See also* the commentary to subsection 4.E; *cf.* CONN. GEN. STAT. § 17-206i(c) (1981); DEL. CODE ANN. tit. 16, § 5161 (Supp. 1980); N.J. STAT. ANN. §§ 30:4-24.2(b), (d) (West 1981); N.Y. MENTAL HYG. LAW § 9.07 (McKinney 1979); OHIO REV. CODE ANN. § 5122.29 (Baldwin 1980).

Section 11. SUCCESSIVE PERIODS OF COMMITMENT.

11.A. ***Sixty-Day Recommitment.***

1. Any person who has been subject to a thirty-day commitment pursuant to Section 6 may be recommitted for up to sixty days upon a petition by the treatment facility or by the next of kin or guardian. The petition may be filed with the court at any time prior to the expiration of the thirty-day commitment. The petition shall include a statement as to why the person still meets the criteria for involuntary commitment, what treatment has been provided and what progress has been made, why a further period of commitment is warranted, and the identity of the person who has knowledge concerning the case. The petition shall be promptly served by the court on the patient, the next of kin or guardian, and the patient's attorney.

2. The patient shall be entitled to a hearing before the court on the petition on or before the first business day following the expiration of the thirty-day commitment, and shall have all other rights to which he was entitled at the hearing on thirty-day commitment.

3. The court shall order that the person be discharged unless it determines (a) by clear and convincing evidence that the person still satisfies the criteria for involuntary commitment, and (b) that there is a reasonable prospect that a substantial therapeutic purpose would be served by a further period of commitment.

COMMENTARY: If, following thirty-day commitment, it appears that a patient continues to meet all the criteria for commitment, this subsection authorizes the treatment facility or next of kin or a guardian to petition for a further period of sixty days of hospitalization. Ordinarily, the treatment facility will be in the best position to evaluate the patient's mental condition. However, in some cases, the family, because of its more intimate knowledge of the patient and of the home environment to which he would return, may have valid reasons for seeking further periods of treatment even if the treatment facility would not do so. This is not meant to suggest that a court should unquestioningly obey the family's desires; on the contrary, it should be aware of the possible conflicts of interest between the family and the patient. However, a wife should have an opportunity to explain to a court that she regularly visits her husband at the hospital and believes that he would assault her if he returned home.

At the hearing on sixty-day recommitment, a patient has the same rights as he had at the hearing on thirty-day commitment.[361] Sixty days is considered an intermediate term of com-

[361] *See* subsection 6.D. The state has the burden of proving by clear and convincing evidence the need for continued confinement. The respondent has no burden to prove

mitment during or by the end of which it should be possible to release most patients.[362]

11.B. ***One-Hundred-Eighty-Day Recommitments.***

1. Any person who has been subject to sixty-day recommitment pursuant to subsection 11.A. may be recommitted for up to one hundred eighty days upon a petition filed with the court by the treatment facility or by the next of kin or guardian. The petition shall include a statement as to why the person still meets the criteria for involuntary commitment, what treatment has been provided and what progress has been made, why a further period of commitment is warranted, and the identity of the person who has knowledge concerning the case. The petition shall be promptly served by the court on the patient, the next of kin or guardian, and the patient's attorney.

2. The patient shall be entitled to a hearing before the court on the petition on or before the first business day following expiration of the operative period of commitment and shall have all other rights to which he was entitled at the hearing on thirty-day commitment.

3. The court shall order that the person be discharged unless it determines (a) by clear and convincing evidence that the person still satisfies the criteria for involuntary commitment, and (b) that there is a reasonable prospect that a substantial therapeutic purpose would be served by a further period of commitment.

4. Additional recommitments for periods of up to one hundred eighty days each may be ordered in accordance with paragraphs 1–3 of subsection 11.B when warranted.

COMMENTARY: This subsection provides for successive periods of recommitment of up to 180 days each for persons who require treatment for longer than 60 days. The empirical data suggest that this provision will apply to a small number of persons who are demonstrably dangerous to others, and who, had they not been processed through the mental health system, could have

---

his sanity or his competence. *See* Matter of Rockman, 104 Misc. 218, 428 N.Y.S.2d 168 (Sup. Ct. 1980).

[362]The empirical evidence indicates that most committed persons can improve sufficiently to be released in 30 to 90 days. *See, e.g.*, Gove & Fain, *supra* note 23, at 673 (67% of involuntary patients discharged within 38 days); Tomellieri, Lakshminanayanan & Herjanic, *supra* note 23, at 291 (63% of committed patients discharged within 90 days). In Boston State Hospital, which was the subject of a major "right to refuse treatment" suit, the average length of stay in 1974 was 14 days. Rogers v. Okin, 478 F. Supp. 1342, 1356 (D. Mass. 1978), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982); *see also* Mattes, *The Optimal Length of Hospitalization for Psychiatric Patients: A Review of the Literature*, 33 HOSP. & COMMUNITY PSYCHIATRY 824 (1982).

As noted in the commentary to Section 11.B, the authors believe that parens patriae confinement should not continue beyond the end of such 60-day recommitment (i.e., a total of about 104 days).

been criminally indicted, possibly leading to a long-term sentence.[363]

The authors of this Article must state candidly that this provision contains two elements that we believe are unwise, but that a majority of the Assembly of the APA endorsed. First, we personally believe that pure parens patriae commitments, should end *prior* to 180-day commitment (that is, after emergency, thirty- and sixty-day commitment, or a total of about 104 days). Clinical research shows that such a limit would permit effective treatment for the vast majority of committed patients who can be treated.[364] We believe that after a certain period of treatment and involuntary confinement, society must give patients a chance on their own. We might agree that this should not apply to an imminently suicidal person if it could be predicted that some finite further period of treatment would likely end the person's suicidal impulses. Unfortunately, this is rarely possible. Nevertheless, the APA concluded that even as to parens patriae commitments, the duration of commitment should be open-ended, as long as the court determines that the person continues to meet the criteria for commitment. This position was motivated by concerns for patients who continue to need treatment. Although we disagree with the Model Law on this point, such concerns merit serious consideration.

The Model Law also provides that all commitment decisions (including successive 180-day commitments) should be made by the court. The authors personally believe that jury trials are proper.[365] The Supreme Court has not decided whether jury trials are constitutionally required in civil commitment hearings.[366] In other contexts, the Supreme Court has held that trial by jury is *not* an indispensable element of due process; nor is it required for accurate fact-finding.[367] The lower courts have held

[363]*See* Abramson, *The Criminalization of Mentally Disordered Behavior: Possible Side-Effects of a New Mental Health Law*, 23 HOSP. & COMMUNITY PSYCHIATRY 101 (1972)

[364]*See* Klerman, *supra* note 7, at 110.

[365]The reasons the APA rejected the requirement of a jury trial are captured in the following comments sent to Mr. Stromberg by various psychiatrists: (1) "To require a jury trial after the first 30 days of commitment is absurd [for it] would effectively prevent continuation of treatment beyond 30 days in almost all instances"; (2) "[t]he jury demand would simply be used to 'bludgeon' hospitals into releasing patients inappropriately"; (3) "[t]he requirements for jury trial invoke further unnecessary adversarial conflicts, and are unrealistic as an item of additional costs"; and (4) "[t]reatment staff would spend all their time in commitment hearings rather than in treating patients." (Personal communications on file with C. Stromberg.)

[366]*See* Humphrey v. Cady, 405 U.S. 504, 511 (1972).

[367]*See* Parham v. J.R., 442 U.S. 584 (1979) (upholding statute allowing juvenile commitment without jury trial); McKeiver v. Pennsylvania, 403 U.S. 528 (1971) (jury trial

in most cases that civil commitment does not require a jury trial,[368] although several courts have compelled jury trials for long-term commitments on the ground that the state law appeared to require it in analogous proceedings.[369] A number of states provide by statute for jury trials for long-term commitments.[370]

In our view, the constitutional question is not dispositive. The legal system and the public historically have regarded juries as a great bulwark protecting individuals against tyranny. In part, juries ensure that community values affect the decisionmaking process.[371] In recent years, the public has credulously accepted the view that psychiatrists coerce normal citizens into civil commitment, with the complicity of the courts. This cynical view undermines morale throughout the mental health system, and ultimately harms patients. This process needs to be ended. Jury trials, even if they lead to the same results as judges' decisions,[372] will be perceived as better protecting all citizens, which itself has value.

11.C. ***Waiver of Hearings.***

A patient may waive any hearing to which he is entitled under this Section 11 upon a written waiver that the court finds is knowingly and voluntarily executed by the patient.

---

not required in juvenile delinquency hearings); French v. Blackburn, 428 F. Supp. 1351 (D.N.C. 1977), *aff'd mem.*, 443 U.S. 901 (1979).

[368]*See, e.g.*, Suzuki v. Quisenberry, 411 F. Supp. 1113 (D. Hawaii 1976), *supplemented sub nom.* Suzuki v. Alba, 438 F. Supp. 1106 (D. Hawaii 1977), *aff'd in part, rev'd in part, and dismissed in part sub nom.* Suzuki v. Yuen, 617 F.2d 173 (9th Cir. 1980); Doremus v. Farrell, 407 F. Supp. 509 (D. Neb. 1975); County Att'y v. Kaplan, 124 Ariz. 510, 605 P.2d 912 (1980); *In re* Jones, 339 So. 2d 1117 (Fla. 1976), *cert. denied*, 430 U.S. 972 (1979); Markey v. Wachtel, 264 S.E.2d 437 (W. Va. 1979).

[369]*See, e.g.*, Lynch v. Baxley, 386 F. Supp. 378 (M.D. Ala. 1974); Bell v. Wayne County Hosp., 384 F. Supp. 1085 (E.D. Mich. 1974); Lessard v. Schmidt, 349 F. Supp. 1078, 1092 (E.D. Wis. 1972); Estate of Hofferber, 28 Cal. 3d 161, 616 P.2d 836, 167 Cal. Rptr. 854 (1980); *In re* Wagstaff, 93 Mich. App. 755, 287 N.W.2d 339 (1980). *See generally* Note, *The Confinement of Mabel Jones: Is There a Right to a Jury Trial in Civil Commitment Proceedings?*, 6 FLA. ST. U.L. REV. 103 (1978).

[370]*See, e.g.*, ARK. STAT. ANN. § 59-101 (1971); MICH. COMP. LAWS § 330.1458 (1979); N.M. STAT. ANN. § 43-1-12 (1979); N.Y. MENTAL HYG. CODE § 9.35 (McKinney 1978); TEX. REV. CIV. STAT. ANN. art. 5547-36(e) (Vernon Supp. 1982–1983).

[371]*See* Humphrey v. Cady, 405 U.S. 504, 509 (1972) (Under a Wisconsin statute, "the jury serves the critical function of introducing into the process a lay judgment reflecting values generally held in the community, concerning the kinds of potential harm that justify the State in confining a person for compulsory treatment."); Duncan v. Louisiana, 391 U.S. 145, 151, 155–56, 187 (1968) (preservation of right to jury trial for misdemeanors).

[372]Some research indicates that judges and juries are actually *more* likely to believe commitment is required than are psychiatrists. *See, e.g.*, Simon & Cockerham, *Civil Commitment, Burden of Proof, and Dangerous Acts: A Comparison of the Perspectives of Judges and Psychiatrists*, 5 J. PSYCHIATRY & L. 571 (1977).

COMMENTARY: No comment is required.[373]

Section 12. DISCHARGE.

12.A. ***Periodic Review.***

The responsible psychiatrist [or: "the responsible physician"][374] shall review periodically whether a patient still meets the criteria for lawful commitment, and if he concludes that the patient does not, he shall undertake discharge procedures as provided herein.

12.B. ***Patients Likely to Harm Themselves or to Suffer Substantial Deterioration.***

As to a patient committed because he was likely to cause harm to himself or to suffer substantial mental or physical deterioration, if the responsible psychiatrist [or: "the responsible physician"][375] concludes that the patient no longer meets the criteria for lawful commitment, he may discharge the patient directly.

12.C. ***Patients Likely to Harm Others.***

As to a patient committed solely because, or partly because, he was likely to cause harm to others, if the responsible psychiatrist [or: "the responsible physician"][376] concludes that the patient no longer meets the criteria for lawful commitment, or that the patient's treatment program has been completed or is unlikely to provide further benefits, he shall apply to the court for an order discharging or transferring the patient, as may be appropriate.[377] The application shall set forth the relevant facts. The court may conduct an informal hearing, subject to such procedures as the court sets. Nothing in this subsection shall reduce any rights to hearings which patients have pursuant to other provisions of this Act.

---

[373]*Cf.* OHIO REV. CODE ANN. § 5122.15(H) (Page 1981 & Supp. 1982) (such hearings "are mandatory and may not be waived"). Concerning the meaning of "knowingly and voluntarily," *see* Boyd v. Dutton, 405 U.S. 1, 3 (1972); Brady v. United States, 397 U.S. 742 (1970); Boykin v. Alabama, 395 U.S. 238 (1969); Miranda v. Arizona, 384 U.S. 436, 475 (1966); Carnley v. Cochran, 369 U.S. 506, 516 (1962); Johnson v. Zerbst, 304 U.S. 458, 464 (1938). Concerning the presumption against waiver, *see* Hodges v. Easton, 106 U.S. 408, 412 (1882); *see also* Lynch v. Baxley, 386 F. Supp. 378, 396 (M.D. Ala. 1974) (waiver of rights in commitment proceedings must be made by counsel with informed consent of the subject and court approval); Matter of Sherman, 98 Misc. 2d 431, 414 N.Y.S.2d 78 (Sup. Ct. 1979). *See generally* Pate v. Robinson, 383 U.S. 375 (1966) (court's failure to conduct a hearing on petitioner's defense of incompetency to stand trial deprived petitioner of a fair trial). In some states, such as Massachusetts, if a patient waives a hearing, the next hearing may not be waived.

[374]Optional provision.

[375]Optional provision.

[376]Optional provision.

[377]Optional provision. Many psychiatrists believe it is proper for them to make discharge decisions for all patients, whether committed as dangerous to others or on some grounds, unless a psychiatrist at his option seeks judicial consideration of a particular case. Some jurisdictions may prefer this approach.

12.D. ***Temporary Delays.***

Discharge of any patient may be delayed for a reasonable period of time in order to arrange transportation or lodging for the patient, or for other good cause.

12.E. ***Discharge and Recommitment.***

A person who has been discharged from emergency evaluation, thirty-day commitment or a subsequent period of commitment may be recommitted only pursuant to the same procedures provided in this Act and upon a showing of some new circumstances warranting such commitment which were not known at the time of discharge.

12.F. ***Release to Outpatient Treatment.***

The responsible psychiatrist [or: "the responsible physician"][378] may, as part of an individual treatment plan for a patient who is involuntarily committed, release such patient to outpatient treatment upon the condition that if the patient fails to follow through with or respond acceptably to such outpatient treatment, he may be returned to inpatient treatment for the remainder of the operative period of commitment.

12.G. ***Habeas Corpus.***

Nothing in this Act shall limit any other legal rights or remedies concerning discharge which a patient may have or acquire pursuant to law, regulation or policy, including the right to petition for a writ of *habeas corpus*.

COMMENTARY: The questions of how and when patients are discharged is second in importance only to the criteria for commitment. Section 12.A provides that treatment personnel must "review periodically" a patient's condition to determine whether the patient continues to need hospitalization and to satisfy the other criteria for commitment. If a patient no longer meets the commitment criteria, the responsible psychiatrist has an obligation to seek to discharge him.[379]

Different standards and procedures govern the discharge of patients who were committed under parens patriae and the police power. (Compare subsections 12.B and 12.C). In parens

---

[378]Optional provision.

[379]*See* O'Connor v. Donaldson, 422 U.S. 563 (1975). This provision does not deal with arrangements for care after discharge. An appropriate aftercare plan should be developed. *See* N.Y. MENTAL HYG. LAW § 29.15 (McKinney 1978 & Supp. 1982–1983); PA. STAT. ANN. tit. 50, § 7116 (Purdon Supp. 1982–1983); A Patient's Bill of Rights, *supra* note 148, ¶ 10; Bachrach, *Continuity of Care for Chronic Mental Patients: A Conceptual Analysis*, 138 AM. J. PSYCHIATRY 1449 (1981); Meyerson & Herman, *What's New in Aftercare? A Review of Recent Literature*, 34 HOSP. & COMMUNITY PSYCHIATRY 333 (1983).

patriae commitment, the patient's therapeutic interests are paramount. Since the psychiatrist can best judge whether release is warranted, he may release the patient directly.

As to patients committed under the police power, however, social protection must be weighed along with the therapeutic interests of the patient. If such a person no longer meets the commitment criteria—or if he meets the criteria but confinement in a hospital would serve no significant treatment purpose—he may be transferred to another secure facility, or discharged. This decision is made not by the psychiatrist, but by the court, which our society vests with the authority to weigh individual liberty against the need for social protection.[380]

In such cases, Section 12.C allows the discharge hearing to be conducted without formal evidentiary procedures. As in a parole hearing, basic due process rights must be protected, but trial-type procedures need not be followed.

Subsection 12.E ensures that the strict requirements and time periods for commitment cannot be avoided by the ruse of a brief release and new petition for commitment. If a patient is released, new behavior or other events must arise to warrant a new period of commitment.[381]

When appropriate, subsection 12.F permits release to an outpatient program, and ensures that if the patient does not participate conscientiously, he will be returned to inpatient care. This is similar to provisions in some statutes that allow release on "convalescent status."[382] Finally, subsection 12.G preserves the right to petition for the writ of habeas corpus.[383]

---

[380]This has been one of the most controversial features of the Model Law. In our view, it makes little sense to keep a chronic, violent, untreatable mental patient in a hospital ward devoted to treating acute, nonviolent patients. He will benefit little, and the nonviolent patients may suffer as a result. Moreover, to do so perpetuates the hypocrisy of custodial confinement under the guise of treatment. Some psychiatrists seek to maintain this ambiguity by suggesting that they can attempt to treat anyone. However, it is unclear whether there is enough good that can be offered the patient, or enough tangible societal protection, to justify the extraordinary act of preventive detention. On the other hand, processing such a person through a criminal trial followed by a prison term also is inadequate. This provision recognizes the need for secure facilities at which competent treatment is available, and provides some flexibility in permitting transfers to such facilities. We recognize, however, that at present there is a great unmet need for such institutional alternatives. We urge that the constitutional, ethical, and financial problems involved in this type of preventive detention be confronted openly, rather than concealed behind the facade of treatment in civil commitment.

[381]*Cf.* Application of True, 103 Idaho 151, 645 P.2d 891 (1982) (due process hearing required for rehospitalization after conditional release).

[382]*See, e.g.*, ALASKA STAT. § 47.30.200 (1979); N.M. STAT. ANN. § 43-1-21 (1979).

[383]*See, e.g.*, ARK. STAT. ANN. § 59-1424 (Supp. 1979); ILL. REV. STAT. ch. 91 1/2, § 3-905 (1981); NEB. REV. STAT. § 83-1066(9) (1981); N.J. STAT. ANN. § 30:4.24.2(h)

Section 13. CONFIDENTIALITY AND DISCLOSURE OF INFORMATION.

[Guidelines for this topic are contained in the American Psychiatric Association's "Model Law on Confidentiality of Health and Social Service Records."]

COMMENTARY: Because the APA previously approved a general "Model Law on Confidentiality of Health and Social Service Records," it did not include specific provisions governing confidentiality in this Model Law.[384]

Section 14. REPRESENTATION OF PATIENTS.

14.A. ***Right to Counsel at Hearings.***

Every patient shall have a right to counsel to represent him at court hearings under this Act, except that a patient need not be provided with counsel for the preliminary hearing on emergency evaluation provided in subsection 4.F.

COMMENTARY: The right to counsel provided in this subsection applies to all hearings except the "probable cause" hearing on emergency evaluation, because that hearing is intended to be a brief, informal inquiry without full adversarial presentations. Nothing precludes counsel at such hearings, however. The right to counsel includes the right to the assistance of counsel sufficiently in advance of hearings to prepare adequately. (See subsection 6.D.3.)

14.B. ***Resolution of Grievances in Treatment Facilities.***

Every treatment facility shall establish a fundamentally fair procedure for the assertion, resolution, and redress of patients' grievances, and

---

(West 1981); N.Y. MENTAL HYG. LAW § 33.15(a) (McKinney 1978 & Supp. 1982–1983); OHIO REV. CODE ANN. § 5122.30 (Page Supp. 1982). In California, until the decision in Doe v. Gallinot, 486 F. Supp. 983 (C.D. Cal. 1979), habeas corpus petitions were apparently the usual means of obtaining judicial review of confinement. *See* CAL. WELF. & INST. CODE § 5275 (West 1972 & Supp. 1982). For cases concerning the use of habeas corpus to challenge the conditions of confinement, see Williams v. Robinson, 432 F.2d 637 (D.C. Cir. 1970); Covington v. Harris, 419 F.2d 617 (D.C. Cir. 1969); Vaughan v. Bower, 313 F. Supp. 37 (D. Ariz. 1970), *aff'd without opinion*, 400 U.S. 884 (1970).

[384] *But see* subsection 6.D.6. Significant questions arise concerning the patient's access to his own records, disclosure to security firms or prospective employers, disclosure to third-party payors, and disclosure when a person may be assaulted by the patient. *See generally* Appelbaum, *Confidentiality in Psychiatric Treatment*, in PSYCHIATRY 1982, 325 (L. Grinspoon ed.). States may want to address these questions as part of their commitment laws. *Cf.* 42 U.S.C. §§ 9501(1)(H)–(I) (Supp. V 1981); ALASKA STAT. § 47.30.260 (1979); ARK. STAT. ANN. § 59-1416(14) (Supp. 1979); CAL. WELF. & INST. CODE § 5328 (West 1972 & Supp. 1982); N.M. STAT. ANN. § 43-1-19 (1979); N.Y. MENTAL HYG. LAW §§ 33.13–.14 (McKinney 1978 & Supp. 1982–1983); OHIO REV. CODE ANN. § 5122.31 (Page 1981); PA. STAT. ANN. tit. 50, § 7111 (Purdon Supp. 1982–1983); TEX. REV. CIV. STAT. ANN. art. 5547–87 (Vernon Supp. 1982–1983).

shall have a patients' representative or similar person who shall hear patients' grievances, attempt to resolve problems, and protect patients' interests.

COMMENTARY: The Model Law provides for advocacy of patients' rights, but does not choose among the various methods that have been proposed. Proponents of an "external" advocacy model stress the importance of independence, an adversarial system, and legally enforceable remedies, while proponents of an "internal" model argue that more can be accomplished by conciliation and administrative reform.[385] We believe that the full range of models, including legal "protection and advocacy," "treatment review boards," "ombudsmen," and "patients' representatives," can be effective.[386] It must be recognized that many patients' problems are only vaguely legal; that short lengths of stay may constrain traditional legal strategies; and that a mediating, fact-finding, and counselling posture is often preferable to litigation in resolving patients' problems other than discharge and recommitment.[387] Thus, while not rigidly prescribing the service model, this provision requires that the model chosen be "fundamentally fair" and provide for the assertion and resolution of patients' grievances.

14.C. ***Representation by Next of Kin or Guardian.***

Any right of patients provided in this Act may be exercised on behalf of a patient who is unable to exercise such right by a next of kin or guardian, in accordance with law.

COMMENTARY: This subsection ensures that a patient does not lose his rights because he cannot personally exercise them. If a treatment facility violates an incompetent patient's rights, the next of kin or guardian can sue on his behalf. Courts should

---

[385]*See generally* A. STONE (WITH C. STROMBERG), *supra* note 2, at 233–50; HHS REPORT ON THE CHRONICALLY MENTALLY ILL, *supra* note 7. In developing the Mental Health Systems Act, the Department of Health and Human Services and Congress also decided not to embrace a single model, but instead to study alternatives. *See* 42 U.S.C. § 9501(1)(L) (Supp. V 1981).

[386]*See* NYSARC v. Carey, 596 F.2d 27 (2d Cir. 1979); 42 U.S.C. §§ 9501(L)–(M) (Supp. V 1981) (right to some mechanism); LEGAL AND ETHICAL ISSUES REPORT, *supra* note 320, at 1366–70; Broderick, *One-Legged Ombudsman in a Mental Hospital: An Over the Shoulder Glance at an Experiment*, 22 CATH. U.L. REV. 517 (1973); Goldstein, *The Role of Defense Counsel in the Civil Commitment Process*, 1972 AM. CRIM. L. REV. 409; Gupta, *New York's Mental Health Information Service: An Experiment in Due Process*, 25 RUTGERS L. REV. 405 (1971); Regan, *When Nursing Home Patients Complain: The Ombudsman or Patient Advocate*, 65 GEO. L.J. 691 (1977).

[387]*See generally* A. STONE (WITH C. STROMBERG), *supra* note 2, at 233–50; Brakel, *Legal Aid in Mental Hospitals*, 21 AM. B. FOUND. RESEARCH J. 23 (1981) (an empirical review of various models of legal aid for mental patients).

ensure that such substituted assertion of rights occurs only if the patient is unable to exercise his rights, and the surrogate is acting in the patient's best interests. Other state laws (for example, governing bank withdrawals and exercise of stock rights) are incorporated by reference.

Section 15. TRANSPORTATION.

Whenever a patient is to be brought to or from a treatment facility, or is to be transferred to another facility or to a home, the court may direct the sheriff, state police, or other appropriate authorities to furnish suitable transportation.

COMMENTARY: In many states, the lack of statutory authorization for transporting mental patients has prevented court hearings and placements because the authorities either lacked the funds or feared liability for undertaking "unauthorized" acts. This provision specifically authorizes the court to direct police officers or other appropriate authorities (for example, a private ambulance service) to transport patients to or from treatment facilities, court hearings, or placements, including their homes.[388]

Section 16. NON-DEROGATION OF PATIENTS' RIGHTS.

Rights conferred upon patients by this Act shall be in addition to, and nothing in this Act shall revoke or reduce, any rights, privileges, or immunities that a patient may have or acquire by law, regulation, or policy.

COMMENTARY: This provision ensures that the Model Law will not be construed to repeal or to replace existing legal rights of patients, or to preclude future laws from conferring additional rights.[389] Prominent among the rights *not* repealed is the right to sue in court for violation of any rights, whether provided in the Model Law or elsewhere.

Section 17. COSTS OF CARE.

In accordance with law, indigent public patients shall receive care and treatment under this Act without charge to them. Patients committed under this Act who are able to pay may be required to pay some

---

[388]For similar provisions, see, e.g., ALASKA STAT. § 47.30.870 (Supp. 1981); N.M. STAT. ANN. § 43-1-22 (1978).

[389]*Cf.* 42 U.S.C. § 9501(2)(A) (Supp. V 1981); COLO. REV. STAT. § 27-10-104 (1982); MASS. GEN. LAWS ANN. ch. 123, § 25 (West Supp. 1982–1983); WASH. REV. CODE § 71.05.060 (1981). If a conflict arises in the definitions of such rights, established principles of statutory construction should apply. *See* SUTHERLAND, STATUTES AND STATUTORY CONSTRUCTION ch. 22 (Sands 4th ed. 1972).

reasonable costs of care and treatment, and to that end treatment facilities and the State shall be authorized to recover such costs from them or their estate, their family, custodians of their property, or third parties liable for the costs of their care or treatment, in conformity with law. The liability of patients, their families, and others for the long term care of patients committed as likely to cause harm to others shall be specially limited by regulations of the Department of Mental Health.

COMMENTARY: The enormous cost of institutional care for mental patients constitutes an increasing proportion of already overburdened state and local government budgets. Although the Model Law cannot solve this problem, it attempts to allocate the burdens in a fair manner. First, it provides that patients shall receive care and treatment even if they are unable to pay for it. Patients who are not indigent may be required to pay the "reasonable costs" of their care and treatment.[390] Courts have allowed the state to recover such costs from persons who are civilly committed. The power to promulgate specific regulations may be delegated to the Department of Mental Health.[391] In providing that the costs of care may be recovered from patients, families, banks, or insurers, the Model Law incorporates (and does not override) substantive state laws concerning the rights and obligations of those parties.[392]

A patient committed under the police power who has completed a course of treatment but is still considered dangerous by the court poses a special problem. His continued confinement is intended to advance not simply his own and his family's interests, but the safety interests of the state as well. In this respect, he is more like a prison inmate. It is not fair to impose upon him open-ended financial liability for the costs of such confinement. This provision requires that the Department of Mental Health fairly limit such liability by special rules.

Section 18. IMMUNITIES AND PENALTIES.

18.A. ***Immunities.***

1. In the absence of willful misconduct or gross negligence, no officer, director, staff member, or employee of a treatment facility shall be liable

[390]The requirement that costs be reasonable does not mean that they must be uniform. Many states charge patients on a sliding scale based on ability to pay and other factors. *See, e.g.*, ALASKA STAT. § 47.30.910 (Supp. 1981); COLO. REV. STAT. §§ 27-12-102, 27-12-104 (1982); N.M. STAT. ANN. § 43-1-25 (1978); N.Y. MENTAL HYG. LAW §§ 41.36, 43.03 (McKinney Supp. 1982–1983); UTAH CODE ANN. § 64-7-18 (1977).

[391]*See* State Dep't of Mental Hygiene v. Schneps, 95 Misc. 2d 828, 408 N.Y.S.2d 980 (App. Term 1978); Matter of Klisurich, 98 Wis. 2d 274, 296 N.W.2d 742 (1980); Matter of Nelson, 98 Wis. 2d 261, 296 N.W.2d 736 (1980).

[392]Some state laws exclude persons other than direct family members from financial liability. *See, e.g.*, COLO. REV. STAT. § 27-12-101 (1982).

for acts or omissions within the scope of his employment related to admission, evaluation, care, treatment, nonadmission, transfer, removal of restrictions upon, or discharge of a person, pursuant to this Act.

2. No other person who, acting in good faith and with a reasonable basis, participates in any of the processes provided in this Act shall be liable for such actions.

3. Notwithstanding any other provision of this Act, no police officer, no officer, director, staff member, or employee of a treatment facility, and no other person or entity performing actions pursuant to this Act, shall be liable for any action of a patient who is discharged from or is absent from a treatment facility pursuant to this Act.

4. Under no circumstances shall any person performing actions pursuant to this Act have a duty to, or be liable for failing to, notify, advise or warn anyone concerning the non-admission, transfer, removal of restrictions on, or discharge of any person.

COMMENTARY: Suing professionals and hospitals who have responsibility for the care of mentally ill patients has become increasingly frequent. Unless the problem is addressed in some way, many of the best professionals and facilities will stop serving civilly committed patients. The authors, like many others, however, are reluctant to carve out broad exceptions to the usual rules of legal accountability.

In considering the problem, it is important to focus on the various types of lawsuits which occur. For example, tort suits based on malpractice, other acts of negligence, or personal assaults by staff members, should not be deterred by special immunity rules.[393] They should be treated no differently than had the actions occurred in other contexts. If proven, these acts should be punished by criminal or civil liability and administrative sanctions. Another category of suits, broad constitutional "right to treatment" cases, usually challenge overall institutional conditions; such suits customarily name individual administrators or staff only in their official capacities.[394] The right to bring such actions also should not be limited by broad legal immunities.

Several other types of cases, however, now plague the mental health system. Psychiatric staff and hospitals have been sued

[393]*See, e.g.*, Sanz v. Puerto Rico, 535 F. Supp. 330 (D.P.R. 1982); Vigilant Ins. Co. v. Keiser, 391 So. 2d 706 (Fla. Dist. Ct. App. 1980); Florida v. Tsavaris, 382 So. 2d 56 (Fla. Dist. Ct. App. 1980), *aff'd*, 394 So. 2d 418 (Fla. 1981); Cucalon v. New York, 103 Misc. 2d 808, 427 N.Y.S.2d 149 (Ct. Cl. 1980).

[394]*See, e.g.*, Wyatt v. Stickney, 344 F. Supp. 373 (M.D. Ala. 1972), *modified and remanded sub nom.* Wyatt v. Aderholt, 503 F.2d 1305 (5th Cir. 1974); *cf.* O'Connor v. Donaldson, 422 U.S. 563, 577 (1975) (Mental health staff are not immune from suit if they "knew or should have known that their actions violated the plaintiff's rights," or if they "took the action with the malicious intention to cause a deprivation of constitutional rights or other injury.").

both for seeking to commit patients,[395] and for not recommending their commitment;[396] for failing to provide treatment,[397] and for treating patients against their will;[398] for medicating or restraining patients who may harm themselves or others,[399] and for failing to medicate or restrain such patients;[400] for releasing patients who might do harm,[401] and for retaining them.[402] They also have been sued for a "failure to warn" possible victims,[403] while they are subject to legal obligations of confidentiality. (Some courts have rejected the "duty to warn"—particularly if the potential victims are not "readily identifiable.")[404] Practitioners even have been sued by criminals who claimed their conduct was the doctor's fault for failing to cure them.[405]

What then is society telling therapists to do? It is easy to say that they must make "reasonable judgments." But because predictions of dangerous conduct are unavoidably unreliable, this standard invites judicial review of therapists' decisions in the blinding glare of hindsight. If one confers on patients the right to the least restrictive alternative, therapists should not also be

---

[395]*See, e.g.*, Hall v. Quillen, 631 F.2d 1154 (4th Cir. 1980); Turner v. Baptist Medical Center, 275 Ark. 424, 631 S.W.2d 275 (1982); McLean v. Sale, 54 N.C. App. 538, 284 S.E.2d 160 (1981).

[396]*See, e.g.*, Seibel v. Kimble, 63 Hawaii 516, 631 P.2d 173 (1981); Hendrix v. City of Topeka, 231 Kan. 113, 643 P.2d 129 (1982).

[397]*See* discussion of subsection 10.C.

[398]*See* discussion of subsection 8.B.

[399]*See, e.g.*, Moon v. United States, 512 F. Supp. 140 (D. Nev. 1981).

[400]*See, e.g.*, Skar v. City of Lincoln, 599 F.2d 253 (8th Cir. 1979); Harper v. Cserr, 544 F.2d 1121 (1st Cir. 1976); Koenigsmark v. New York, 80 A.D.2d 707, 437 N.Y.S.2d 745 (1981).

[401]*See, e.g.*, Bowers v. DeVito, 686 F.2d 616 (7th Cir. 1980); Buford v. California, 104 Cal. App. 3d 811, 164 Cal. Rptr. 264 (1980); Bellavance v. State, 390 So. 2d 422 (Fla. Dist. Ct. App. 1980); Bradley Center, Inc. v. Wessner, 161 Ga. App. 576, 287 S.E.2d 716 (1982); Fuhrmann v. Hattaway, 109 Mich. App. 429, 311 N.W.2d 379 (1981); Fiederlein v. City of New York, 79 A.D.2d 489, 437 N.Y.S.2d 321 (1981); Levenhelt v. State, 6 Ohio App. 2d, 399 N.E.2d 106 (1978); *cf.* Teasley v. United States, 662 F.2d 787 (D.C. Cir. 1980). A recent article reported 33 recent cases of this kind. Kroll & Mackenzie, *When Psychiatrists are Liable: Risk Management and Violent Patients*, 34 HOSP. & COMMUNITY PSYCHIATRY 29 (1983).

[402]*See, e.g.*, Whitree v. State, 56 Misc. 2d 693, 290 N.Y.S.2d 486 (Ct. Cl. 1968).

[403]*See, e.g.*, Williams v. United States, 450 F. Supp. 1040 (D.S.D. 1978); Tarasoff v. Regents of Univ. of Cal., 17 Cal. 3d 425, 551 P.2d 334, 131 Cal. Rptr. 14 (1976); Mavrouchis v. Superior Court, 102 Cal. App. 3d 594, 162 Cal. Rptr. 724 (1980); Bellah v. Benson, 73 Cal. App. 3d 911, 141 Cal. Rptr. 92 (1977); McIntosh v. Milano, 168 N.J. Super. 466, 403 A.2d 500 (1979). *See generally* Fleming & Maximov, *The Patient or His Victim: The Therapist's Dilemma*, 62 CAL. L. REV. 1025 (1974).

[404]*See* Hasenei v. United States, 541 F. Supp. 999 (D. Md. 1982); Leedy v. Hartnett, 510 F. Supp. 1125 (M.D. Pa. 1981), *aff'd without opinion*, 676 F.2d 686 (3d Cir. 1982); Thompson v. County of Alameda, 27 Cal. 3d 741, 614 P.2d 728, 167 Cal. Rptr. 70 (1980). *But see* Lipari v. Sears, Roebuck & Co., 497 F. Supp. 185 (D. Neb. 1980).

[405]*See, e.g.*, Veverka v. Cash, 318 N.W.2d 447 (Iowa 1982); Cole v. Taylor, 301 N.W.2d 766 (Iowa 1981).

sued for implementing that alternative.[406] The legal message to health professionals[407] is so threatening and unclear that many are refusing to treat patients who may be violent. Thus, the danger to society is increased, rather than decreased.[408]

Several judicial decisions have required therapists to warn possible victims of a patient's dangerous propensities. The courts have based such a duty on the "special relationship" between patient and therapist. Unfortunately, the decisions are unclear as to when such a relationship exists, how certain the therapist must be that harm will occur, what steps the therapist must take, and who he must warn. The "duty to warn" cases are based on the judgment that "[i]n this risk-infested society, we can hardly tolerate the further exposure to danger that would result from a concealed knowledge of the therapist that his patient was lethal."[409] They ignore the fact that therapists cannot predict dangerous conduct well enough to make the duty to warn better than a scattering of buckshot. As one judge noted, creation of such a duty would "take us from the world of reality into the wonderland of clairvoyance."[410] In addition, the legal duty to warn would undermine most therapeutic relationships.[411]

---

[406]*See, e.g.*, Moon v. United States, 512 F. Supp. 140 (D. Nev. 1981); Bradley Center v. Wessner, 161 Ga. App. 576, 287 S.E.2d 716 (1982). As the court said in Lipari v. Sears, Roebuck & Co.:

> "[D]espite the therapeutic benefits of this 'open door' approach, the practice admittedly entails a higher potential of danger both for the patient and [others]. In deciding the extent to which a patient should be released from restrictions, the treating physician must exercise his judgment and balance the various therapeutic considerations together with the possible dangers" . . . . [A] psychotherapist is not subject to liability . . . so long as he uses due care in assessing the risks of such a placement.

497 F. Supp. at 192 (quoting Johnson v. United States, 409 F. Supp. 1283, 1293 (M.D. Fla. 1976), *rev'd on other grounds*, 576 F.2d 606 (5th Cir. 1978)). Several courts have imposed liability on state psychiatrists on the ground that release of a particular patient does not impair the general policy of release to less restrictive facilities and therefore is not a "policymaking" decision protected by the sovereign immunity doctrine. *See, e.g.*, Bellavance v. Florida, 390 So. 2d 422 (Fla. Dist. Ct. App. 1980).

[407]The legal cases may apply not only to psychiatrists but also to nurses, social workers, psychologists, and guidance counsellors who perform various modes of therapy.

[408]*See* Stone, *supra* note 22.

[409]Tarasoff v. Regents of Univ. of Cal., 17 Cal. 3d 425, 442, 551 P.2d 334, 347–48, 131 Cal. Rptr. 14, 27 (1976).

[410]*Id.* at 452, 551 P.2d at 354, 131 Cal. Rptr. at 34 (Mosk, J., concurring and dissenting). The California Supreme Court tried to evade this problem by saying that psychiatrists' predictions must be only as good as "the standards of the profession." *Id.* at 431, 551 P.2d at 340; *see also* Thompson v. County of Alameda, 27 Cal. 3d 741, 752, 614 P.2d 728, 734, 167 Cal. Rptr. 70, 76 (1980). But where such standards are lacking, holding someone to a nonexistent standard is simply a legal tour de force.

[411]The authors believe there may be a reasonable middle position, in which in cases of manifest, imminent violence, the therapist would be obligated to warn the police or

In almost all cases, if a potential victim were warned, he would request that the police commit or confine the person in prison. Thus, the duty to warn is a slippery slope that quickly carries us into the bog of preventive detention based on unreliable predictions of dangerous conduct. In addition, the warned person might decide to harm the ostensibly dangerous patient—thus involving the therapist in hopeless ethical and legal tangles.

Moreover, the result will be a lower level of protection for society. It is well recognized by therapists that a duty to warn others—and the Miranda-type warnings to patients that it would require—would defeat the prospects for treating many dangerous patients.[412] Many psychiatrists will avoid treating patients who may be dangerous. Instead, they will refer these patients to public clinics or to other therapists. Patients may misunderstand the therapist's "message" and discontinue treatment. Until recently, there was no empirical evidence bearing on these assertions, although most psychotherapists believed them to be true. However, an empirical study in California following the *Tarasoff v. Regents of University of California*[413] decision revealed that about four-fifths of the psychotherapists (predominantly psychiatrists) believed that patients were inhibited by being told that their communications were not confidential, and fully one-quarter reported losing patients because of this warning.[414]

In addition, if a therapist continues to treat such patients, he will err in favor of predicting that a patient is dangerous. Because a therapist will be loath to have his judgment reviewed by a jury in a victim's tort action, the costs to him of such an overprediction (possible loss of a patient) will be far less than the costs of underprediction. As a result, many patients who are not violent will be needlessly subjected to confinement.[415]

---

to initiate civil commitment, but not directly to warn potential victims. Warning the police may not have the disruptive impact on therapeutic relationships that direct contact with objects of the patients' hostility may have. In addition, the police, not psychiatrists, are equipped with the communications and security equipment necessary to make the "duty to warn" practical and relatively safe. *See* Stone, *The* Tarasoff *Decision: Suing Psychotherapists to Safeguard Society*, 90 HARV. L. REV. 358, 370–71 (1976).

[412]*Id.*

[413]17 Cal. 3d 425, 551 P.2d 334, 131 Cal. Rptr. 14 (1976).

[414]Note, *Where the Public Peril Begins: A Survey of Psychotherapists to Determine the Effects of* Tarasoff, 31 STAN. L. REV. 165 (1978). There are some therapists, however, who believe that confronting the patient with the legal problem helps him to understand his situation and aids the treatment process.

[415]The empirical study cited above also found that many therapists reported "lowering" their threshold for considering a patient as possibly dangerous following *Tarasoff*.

In response to this problem, many states have enacted laws providing varying levels of legal protection or immunity for mental health professionals.[416]

Section 18 of the Model Law addresses these problems. Paragraph 1 of subsection 18.A ensures that professional participants in the commitment process will not be held liable for acts relating to admission or nonadmission unless they are guilty of "willful misconduct or gross negligence." Although the possible theoretical distinctions among ordinary and gross negligence, willfulness, recklessness, bad faith, and other negligence standards often are discussed, we believe that in most of the cases we are concerned with here, the challenged conduct will be either acceptable or unacceptable under all these standards. This is particularly true if courts infer willfulness or gross negligence from a course of conduct in which a physician did not know, but should have known, that his conduct violated a patient's rights. So interpreted, the standard in the Model Law is in accord with several state statutes and court decisions.[417]

Paragraph 2 provides a similar rule for ordinary citizens who participate in the commitment process, although it states their standard of care affirmatively. These persons must act "in good faith and with a reasonable basis." Good faith is a subjective standard that prohibits taking action with an invidious purpose in mind. To act "with a reasonable basis" means that the person has acted with ordinary due care from an objective viewpoint.[418]

---

[416]*See, e.g.*, ARK. STAT. ANN. § 59-1414(B) (1979) ("No officer, physician, or other person shall be held criminally or civilly liable for his actions pursuant to this Act . . . provided he acts in good faith and without malice."); CAL. WELF. & INST. CODE §§ 5113, 5154, 5173, 5257, 5267, 5278, 5306 (West 1972) (providing, e.g., that no staff person "shall be civilly or criminally liable for any action of a person released"); FLA. STAT. § 768.28(9) (West Supp. 1983) (no liability "unless such officer, employee, or agent acted in bad faith or with malicious purpose or in a manner exhibiting wanton and willful disregard of human rights, safety or property"); PA. STAT. ANN. tit. 50, § 7114 (Purdon Supp. 1982–1983) ("in the absence of willful misconduct or gross negligence," no staff member is liable for any treatment or release decision "or for any of its consequences").

[417]*See, e.g.*, O'Connor v. Donaldson, 422 U.S. 563 (1975); Downs v. Sawtelle, 574 F.2d 1 (1st Cir.), *cert. denied*, 439 U.S. 910 (1978); ARK. STAT. ANN. § 59-1414(B) (1979); CAL. WELF. & INST. CODE §§ 5113, 5154, 5173, 5257, 5267, 5278, 5306 (West 1972); FLA. STAT. § 768.28(9) (West Supp. 1983); PA. STAT. ANN. tit. 50, § 7114 (Purdon Supp. 1982–1983). Also, "[c]allous or wanton neglect may warrant a finding of malice . . . ." Rogers v. Okin, 478 F. Supp. 1342, 1382 (D. Mass. 1979), *aff'd in part and rev'd in part*, 634 F.2d 650 (1st Cir. 1980), *vacated and remanded sub nom.* Mills v. Rogers, 102 S. Ct. 2442 (1982); *accord* Harper v. Cserr, 544 F.2d 1121, 1125 (1st Cir. 1976); *cf.* Wood v. Strickland, 420 U.S. 308 (1975) (liability only if person "acted with such an impermissible motivation or with such disregard . . . of clearly established constitutional rights that [the actions] cannot be reasonably characterized as being in good faith").

[418]*Compare* ARIZ. REV. STAT. ANN. § 36-515 (1974 & Supp. 1982–1983) ("Any person acting in good faith upon either actual knowledge or reliable information who makes

Paragraph 3 exempts police officers and treatment personnel from any liability for dangerous acts by patients who are released. In the authors' view, this exemption is unacceptably broad. In the most extreme case, an intentional release of a patient whom a psychiatrist strongly believes will in fact commit a homicide should be actionable. The plaintiff should be required to prove that the psychiatrist was virtually certain of his prediction, and of course this would be a rare case. Under those circumstances, however, we think it would be appropriate for such a psychiatrist to be subjected to a civil fine and administrative sanctions in addition to monetary liability. In approving the Model Law, however, the APA chose a blanket immunity provision.

Paragraph 4 creates an immunity from suits based on a therapist's failure to warn possible victims of the prospective discharge of a mental patient. As noted above, the authors personally believe that in cases of manifest imminent violence, warning the police may be the proper approach.

While paragraphs 1 to 3 refer to immunity from liability, it should be noted that paragraph 4 negates any duty to warn under any circumstances, and would therefore result in dismissal at the outset of any suit based on such a theory. It would not be necessary to examine the other conditions for liability such as breach of a duty of due care or proximate cause.

18.B. ***Penalties.***

1. Any person who knowingly and willfully gives substantial, false information or takes other wrongful action for the purpose of distorting, corrupting or interfering with the processes provided in this Act shall be subject to a civil fine, and shall be liable for injunctive relief and money damages, in addition to any other liability under law.

2. Any person who takes into custody, admits for evaluation or commitment, detains for a further period of time, discharges, or administers medication or treatment to a patient, or takes other action affecting the substantial rights of a patient, doing so knowingly and willfully in substantial violation of this Act, shall be subject to a civil fine, and shall be liable for injunctive relief and money damages, in addition to any other liability under law. This paragraph shall not be invoked in cases of minor, merely technical, or otherwise justifiable breaches of the provisions of this Act.

---

application for evaluation or treatment of another person . . . is not subject to civil or criminal liability for such act."), *with* OHIO REV. CODE ANN. § 5122.34 (Page 1981) (no liability for person "acting in good faith, either upon actual knowledge or information thought by them to be reliable").

COMMENTARY: The Model Law seeks to protect both the respondents in a commitment proceeding, and the integrity of the commitment process itself, by providing for civil fines as well as private legal remedies against any person who gives false information or otherwise acts wrongfully to have a person committed.

Paragraph 2 imposes a civil fine upon any person who acts "knowingly and willfully" to violate the law and to deprive patients of substantial rights or otherwise corrupt the civil commitment process.[419] This provision would apply, for example, to a person who lied in order to have a person improperly committed, or to a police officer who took a person into custody knowing that he did not satisfy the criteria for commitment. A therapist who continued to apply medications or restraints knowingly in violation of the law also could be penalized under this provision.[420]

The penalty provision does not apply to minor or otherwise justifiable departures from the law. For example, a treatment facility would be immobilized if it could be held liable for keeping a voluntary patient a half-hour beyond the time limit for discharge upon request. Similarly, it would be absurd to hold a police officer liable for taking into custody a person whose suicide attempt occurred forty-nine hours rather than the permissible forty-eight hours earlier. This provision protects the integrity of civil commitment against egregious violations. It does not impose an unrealistic "strict liability" standard upon what is necessarily a flexible legal and therapeutic process.

Section 19. REGULATIONS.

The Commissioner of Mental Health is empowered to promulgate regulations to implement this Act that are consistent with its provisions.

COMMENTARY: Several provisions of the Model Law call for the issuance of more detailed regulations by each state's Department of Mental Health. This Section provides explicit statutory authority for promulgating such regulations, so long as they are consistent with the Model Law.[421]

---

[419]It should be noted that the "knowingly and willfully" standard is more difficult to meet than the standard of negligence. It approximates the standard of the criminal law.

[420]*Cf.* ARIZ. REV. STAT. ANN. § 36-515(B) (1956 & Supp. 1982–1983); CAL. WELF. & INST. CODE § 5203 (West 1972); CONN. GEN. STAT. § 17-184 (1981).

[421]*Cf.* ALASKA STAT. § 47.30.290 (1962).

Section 20. CONSTRUCTION.

20.A. ***Gender and Number.***

As used in this Act, pronouns shall refer to both male and female persons equally, and articles shall refer to singular and plural persons and things.

COMMENTARY: No comment is required.[422]

20.B ***Severability.***

If any provision of this Act or its application to any person or circumstance is held invalid, it is the legislative intent that such invalidity not affect other provisions or applications which can be given effect apart from that which is invalidated, and to this end the provisions of this Act shall be deemed severable.

COMMENTARY: In recent years, courts have invalidated provisions of several state commitment statutes.[423] This Section would allow the commitment process to continue, despite the invalidation of a particular provision of the Model Law. It encourages the courts to rule that the statute will continue to operate, if possible, apart from the invalidated provision.

20.C. ***Construction Against Implied Repeal.***

This Act is intended as a unified, general Act covering its subject matter, and accordingly none of its provisions shall be deemed to be repealed by implication by subsequent legislation if such a construction reasonably can be avoided.

COMMENTARY: This provision is intended to create a presumption against the interpretation that subsequent legislation (particularly in an unrelated area of law such as domestic relations), should be read to repeal by implication any of the specific provisions of the Model Law.

---

[422]The Model Law was not written in entirely gender-neutral language. However, this provision dictates that the use of gendered pronouns has no legal significance whatsoever. For an interesting effort to develop ways of achieving gender-neutral language, see AM. PSYCHOLOGICAL ASS'N, GUIDELINES FOR NONSEXIST LANGUAGE IN APA JOURNALS (1977).

[423]*See, e.g.*, Doe v. Gallinot, 486 F. Supp. 983 (C.D. Cal. 1979), *aff'd*, 657 F.2d 1017 (9th Cir. 1981); Colyar v. Third Judicial Dist. Ct., 469 F. Supp. 424 (D. Utah 1979); Suzuki v. Quisenberry, 411 F. Supp. 1113 (D. Hawaii 1976), *supplemented sub nom.* Suzuki v. Alba, 438 F. Supp. 1106 (D. Hawaii 1977), *aff'd in part, rev'd in part, and dismissed in part sub nom.* Suzuki v. Yuen, 617 F.2d 173 (9th Cir. 1980).

STATUTE

CONTRIBUTION AND CLAIM REDUCTION IN ANTITRUST LITIGATION: A LEGISLATIVE ANALYSIS

DONALD J. POLDEN*
E. THOMAS SULLIVAN**

The Supreme Court's recent decision in Texas Industries v. Radcliff Materials *affirmed the common law rule against contribution among antitrust defendants, and left congressional action as the only avenue for those seeking to allow contribution. In this Article, Professors Polden and Sullivan examine the relevant legislation proposed in the last two Congresses in light of the policies underlying the antitrust laws. They conclude that a contribution rule, if adopted by Congress, should apply only to offenses subject to a rule of reason analysis and that contribution shares should be determined by standards more flexible than those of market share analysis. The authors believe that the trial judge in antitrust cases should have discretion over whether or not to conduct a jury trial. Finally, they suggest that legislation in this area should not be applied retroactively.*

The right to contribution among antitrust defendants is a controversial issue in antitrust litigation. Because a plaintiff can choose to recover all his damages from a single defendant,[1] notwithstanding the possible existence of other, perhaps more culpable defendants, the absence of a contribution rule means that the sued defendant may pay damage judgments far in excess of his responsibility, while other responsible parties escape liability. The enactment of a contribution rule would remedy this unfairness by permitting the paying defendant to distribute a portion of his liability among his coconspirators. A claim reduction rule would extend this result to settlements in multidefendant cases. The outstanding damage claim would be reduced by a settling defendant's contribution share so that this settlement

---

*Professor of Law, Drake University. B.B.A., The George Washington University, 1970; J.D., Indiana University, 1975.

**Professor of Law, University of Missouri, Columbia. B.A., Drake University, 1970; J.D., Indiana University, 1973.

The authors express their appreciation to Dean Richard Calkins of the Drake University Law School and Professor John J. Quinn of the Osgoode Hall Law School for their thoughtful comments on earlier drafts and to Patricia Stegmaier of Drake University Law School and James Burt of the University of Missouri Law School for their capable research assistance. Errors, if any, are the authors' responsibility.

[1] *See, e.g.*, Walker Distrib. Co. v. Lucky Lager Brewing Co., 323 F.2d 1, 8 (9th Cir. 1963).

would not leave the remaining defendants liable for more than their shares.

These arguments convinced the United States Court of Appeals for the Eighth Circuit to adopt a contribution rule in 1979.[2] During the same year, Senator Birch Bayh (D-Ind.) introduced a pro-contribution provision as part of Senate Bill S. 390, entitled the "Antitrust Improvements Act of 1979."[3] The contribution provision was withdrawn from S. 390 and later introduced as a separate contribution bill, S. 1468.[4] While this proposal was being considered by the Senate, the Supreme Court held that no right of contribution existed under the antitrust laws.[5] Congressional action therefore became the only available mechanism for the creation of a contribution right.

In response to the Supreme Court decision, the Senate Committee on the Judiciary held hearings in 1981 and early 1982,[6] and a new pro-contribution bill, S. 995,[7] was reported out of committee by a twelve to six vote on March 31, 1982. The Committee on the Judiciary of the House of Representatives also considered several bills concerning contribution and claim reduction, including bills that were very similar to S. 995.[8] None of these bills were passed during the Ninety-seventh Congress, but S. 995 already has been reintroduced in the Ninety-eighth Congress as S. 380,[9] and H.R. 2244 recently was introduced in the House Judiciary Committee.[10]

---

[2]Professional Beauty Supply v. National Beauty Supply, 594 F.2d 1179 (8th Cir. 1979).

[3]S. 390, 96th Cong., 1st Sess. (1979).

[4]S. 1468, 96th Cong., 1st Sess. (1979).

[5]Texas Indus. v. Radcliff Materials, 451 U.S. 630 (1981).

For a discussion of the merits of a contribution rule, see generally Easterbrook, Landes & Posner, *Contribution Among Antitrust Defendants: A Legal and Economic Analysis*, 23 J.L. & ECON. 331 (1980); Floyd, *Contribution Among Antitrust Violators: A Question of Legal Process*, 1980 B.Y.U. L. REV. 183; Jacobson, *Contribution Among Antitrust Defendants: A Necessary Solution To A Recurring Problem*, 32 U. FLA. L. REV. 217 (1980); Polinsky & Shavell, *Contribution and Claim Reduction Among Antitrust Defendants: An Economic Analysis*, 33 STAN. L. REV. 447 (1981); Sullivan, *New Perspectives in Antitrust Litigation: Towards A Right of Comparative Contribution*, 1980 U. ILL. L.F. 389.

[6]*The Antitrust Equal Enforcement Act: Hearings on S. 995 Before the Senate Comm. on the Judiciary*, 97th Cong., 1st & 2d Sess. (1981–1982).

[7]S. 995, 97th Cong., 2d Sess. (1982).

[8]For example, three bills were introduced in the House in the first session of the Ninety-seventh Congress: H.R. 1242, H.R. 4072, and H.R. 5794. As discussed in Section II, some of these bills represent ongoing efforts to gain passage of legislation like S. 1468 and S. 995, while other bills represent independent efforts to resolve legislatively critical issues concerning contribution and claim reduction in antitrust cases.

[9]S. 380, 98th Cong., 1st Sess. (1983).

[10]H.R. 2244, 98th Cong., 1st Sess. § 3 (1983) ("Antitrust Fairness Amendments of 1983") (adding a new section 4J to the Clayton Act, 15 U.S.C. §§ 12–27 (1976)), *reprinted in* 44 ANTITRUST & TRADE REG. REP. (BNA) No. 1107, at 673 (Mar. 24, 1983).

The premise of this Article is that the thoughtful consideration of the values underlying antitrust law and practice will dictate proper legislative choices in the composition of contribution and claim reduction legislation. Section I will summarize the courts' treatment of contribution rights in antitrust cases. The traditional rule holding that antitrust defendants had no right to contribution was challenged recently in the Eighth Circuit, but was reaffirmed by the Supreme Court. In Section II, the legislative proposals offered in the current and prior Congresses will be examined. Section III will develop the concerns underlying antitrust law—deterrence, compensation, complexity, and fairness—in the contribution context. Section IV will analyze the legislative proposals discussed in Section II. Specifically, the authors will discuss the provisions of the bills that concern the coverage of the contribution right, the computation of contribution shares, claim reduction, the jury trial right, and the retroactive applicability of the statute. The conclusion will present the authors' resolution of these issues and a model statute.

I. Case Law Developments

The first federal court to consider contribution between antitrust tortfeasors was the United States Court of Appeals for the Third Circuit in *Goldlawr, Inc. v. Shubert*.[11] Shubert filed a third-party complaint which alleged that the third-party defendant was responsible for the damages suffered by Goldlawr. The court ruled that the two conspiracies were separate and distinct and concluded that Shubert and the impleaded third-party defendant were not joint tortfeasors. In reaching this decision, the court opined in dictum that the federal common law did not provide a right of contribution between antitrust tortfeasors.[12]

The first United States court of appeals to hold that a right to contribution *did* exist under the antitrust laws was the Eighth Circuit in *Professional Beauty Supply v. National Beauty Supply*.[13] Several district courts previously had decided the issue

---

[11]276 F.2d 614 (3d Cir. 1960).

[12]*Id.* at 616. But compare *Goldlawr, Inc. v. Shubert* with *Gomes v. Brodhurst*, 394 F.2d 465 (3d Cir. 1968), a personal injury action predicated on a negligence theory, where the Third Circuit adopted a comparative fault contribution rule. The *Goldlawr* court observed that there was "no longer a legitimate place in our system, if indeed, there ever was, for a rule of law which places the full burden of restitution upon one who is only in part responsible for a plaintiff's loss." *Id.* at 467.

[13]594 F.2d 1179 (8th Cir. 1979).

directly to the contrary.[14] The defendants in *Professional Beauty* filed a third-party complaint seeking contribution. The district court dismissed the complaint for failure to state a claim upon which relief could be granted. On appeal, the Eighth Circuit reversed and required contribution by the third-party antitrust defendant. The court, in a two to one decision, reasoned that both deterrence and fairness dictated the adoption of a contribution rule.[15]

The *Professional Beauty* court was persuaded that a no-contribution rule could reduce the deterrent effect of the antitrust laws by allowing violators to escape liability. The possibility of escaping liability was perceived by the court to be greatest in situations where a plaintiff is so dependent on a violator that the plaintiff chooses not to sue that potential defendant for fear of retaliation by the violator, and instead sues other, relatively less powerful defendants.[16] This is likely to arise when the parties are in a vertical relationship. Apart from its concern about the deterrent effect of the existing law, the court objected to the misallocation of damages caused by the absence of a contribution rule.[17]

While noting the significance of the deterrence and fairness concerns in the case before it, the court conceded that such equitable considerations would not exist in every case. Accordingly, it left the decision on whether to use the contribution rule to the discretion of the trier of fact.[18] This discretion was to be exercised in light of the relative bargaining power of the wrongdoers and the extent of their participation in the illegal conduct.[19] Each defendant's contribution portion was to be calculated on a pro rata basis, with each violator bearing an equal share of the damage liability.[20]

---

[14]Olson Farms, Inc. v. Safeway Stores, 1977-2 Trade Cas. (CCH) ¶ 61,698 (D. Utah 1977), *aff'd*, 1979-2 Trade Cas. (CCH) ¶ 62,995 (10th Cir. 1979); El Camino Glass v. Sunglo Glass Co., 1977-1 Trade Cas. (CCH) ¶ 61,533 (N.D. Cal. 1976); Sabre Shipping Corp. v. American President Lines, 298 F. Supp. 1339 (S.D.N.Y. 1969); Washington v. American Pipe & Constr. Co., 280 F. Supp. 802, 804–05 (S.D. Cal.), *mandamus denied*, 393 F.2d 568 (9th Cir.), *cert. denied*, 393 U.S. 842 (1968).

[15]*Professional Beauty*, 594 F.2d at 1185–86.

[16]During depositions in *Professional Beauty*, the principal shareholders of plaintiff "stated that they were persuaded not to name [an additional defendant] because of [its] decision to renew Professional's franchise." *Id.* at 1185.

[17]*Id.* at 1185–86.

[18]*Id.* at 1186. The court did not specify whether the determination was to be made by the judge or by the jury.

[19]*Id.* at 1186.

[20]*Id.* at 1182 & n.4.

The discretionary contribution approach of *Professional Beauty* was rejected by every other court that considered the issue.[21] The principal argument against a contribution approach was articulated by the Fifth Circuit in *Wilson P. Abraham Construction Corp. v. Texas Industries*.[22] That court concluded that a contribution rule would dilute the deterrent effect of the antitrust law. Support for this conclusion was drawn from the economic theory of risk aversion. This theory suggests that business decisionmakers "are deterred more by the slight prospect of a large loss than by the prospect of a small loss."[23] The court, therefore, reasoned that a no-contribution approach, because it increases the size of the potential liability, would produce greater deterrence. The fairness issue arose in the context of a claim that a no-contribution rule violated the Equal Protection Clause by making the sued defendant responsible for the damages caused by others. The court rejected this argument, finding the no-contribution rule rationally related to the deterrent purposes of the antitrust law.[24]

The question concerning the existence of a right of contribution was resolved by the Supreme Court in 1981. In an appeal from *Abraham Construction*, the Court unanimously held in *Texas Industries v. Radcliff Materials*[25] that Congress, in adopting the Sherman and Clayton Acts, did not intend to create contribution rights, and that the federal courts were without power to fashion common law rules sanctioning contribution in antitrust litigation. The Court supported its first conclusion by examining the legislative history of the antitrust laws and concluded that Congress did not provide for contribution either explicitly or implicitly.[26] Reasoning that Congress enacted the antitrust treble-damage provision as a means to "punish past,

---

[21]*See, e.g.*, Olson Farms v. Safeway Stores, 1979-2 Trade Cas. (CCH) ¶ 62,995 (10th Cir. 1979); *In re* Beef Indus. Antitrust Litig., 607 F.2d 167 (5th Cir. 1979); Wilson P. Abraham Constr. Corp. v. Texas Indus., 604 F.2d 897 (5th Cir. 1979), *aff'd. sub nom.* Texas Indus. v. Radcliff Materials, 451 U.S. 630 (1981); *In re* Corrugated Container Antitrust Litig., 1979-1 Trade Cas. (CCH) ¶ 62,689 (S.D. Tex.), *aff'd mem.*, 606 F.2d 319 (5th Cir. 1979).

[22]604 F.2d 897 (5th Cir. 1979), *aff'd sub nom.* Texas Indus. v. Radcliff Materials, 451 U.S. 630 (1981). In *Abraham Construction* defendant was accused of fixing prices in violation of section 1 of the Sherman Act. The defendant, Texas Industries, filed a third-party complaint against other coconspirators. The district court dismissed defendant's third-party complaint and the Fifth Circuit affirmed on appeal.

[23]*Id.* at 901 n.8.

[24]*Id.* at 904–05.

[25]451 U.S. 630 (1981).

[26]*Id.* at 640.

and to deter future, unlawful conduct,"[27] the Court found no evidence that Congress intended to "ameliorate the liability of wrongdoers."[28]

In evaluating the federal common law as a potential source of the right, the Court confirmed that the formulation of a federal common law is limited to protecting "uniquely federal interests" and "those areas in which Congress has given the courts the power to develop substantive law."[29] Although the Court recognized that federal interests are at stake under the congressionally enacted antitrust laws, it stated:

> Contribution among antitrust wrongdoers does not involve the duties of the federal Government, the distribution of powers in our federal system, or matters necessarily subject to federal control even in the absence of statutory authority . . . . In short, contribution does not implicate "uniquely federal interests" of the kind that obligate courts to formulate federal common law.[30]

The Court similarly distinguished between its authority to develop substantive law principles within the broad mandate of the statutes and its power to formulate remedies for violations of the statutes.[31] Because Congress explicitly provided for specific remedies, in contrast to the open-endedness of the substantive provisions, the Court held that it was without authority to expand the remedial provisions to include contribution among defendants.[32] The Court invited Congress to consider the competing policies, values, and interests at stake, but it did not signal its own views on how those interests should be weighed.[33]

The *Texas Industries* decision was extended to claim reduction by the Fourth Circuit in *Burlington Industries v. Milliken & Co.*[34] The plaintiff (Burlington) had settled with one defendant before bringing its suit. The trial judge found for the plaintiff, but reduced its damage judgment by treble the settlement fig-

---

[27]*Id.* at 639.
[28]*Id.*
[29]*Id.* at 640.
[30]*Id.* at 642.
[31]*Id.* at 643 (emphasis added).
[32]*Id.* at 646 (citing Northwest Airlines v. Transport Workers Union, 451 U.S 77, 97 (1981)).
[33]*Id.* at 646–47. In light of the Court's statement that "treble damages reveal an intent to punish past, and to deter future, unlawful conduct, not to ameliorate the liability of wrongdoers", Congress will be called upon to balance this punishment/deterrence rationale with the difficult equity issues created by the present status of the law. *See generallv* Sullivan, *supra* note 5.
[34]690 F.2d 380 (4th Cir. 1982).

ure.[35] The Fourth Circuit overturned the trial judge's damage determination when it reaffirmed the prevailing rule that any amount received in settlement should be deducted from plaintiffs' damages *after* trebling, and not before.[36] The court reasoned that claim reduction was analogous to contribution and that, in the absence of a clear legislative indication otherwise, a district court does not have the inherent discretionary power to order claim reduction.[37] Finally, the court stated that claim reduction would retard the beneficial effects of partial settlements in antitrust cases.[38]

Following the Supreme Court's decision in *Texas Industries*, interested congressional representatives proposed legislation that would create a federal right of action for contribution. This legislation is the subject of the next section.

II. LEGISLATIVE PROPOSALS

This Section briefly describes the various legislative proposals considered by Congress during the Ninety-seventh and Ninety-eighth sessions.[39] Although there is considerable variation in the scope of the proposals, they seem to fall into three categories.

---

[35]Burlington Indus. v. Deering Milliken, Inc., 42 ANTITRUST & TRADE REG. REP. (BNA) No. 1052, at 413, 414 (D.S.C. 1982).

[36]Burlington Indus. v. Milliken & Co., 690 F.2d at 391.

[37]*Id.* at 392–95.

[38]*Id.* at 394.

[39]Of the bills considered by Congress during the Ninety-seventh session, only S. 995, 97th Cong., 2d Sess. (1982), had been reintroduced in the Ninety-eighth Congress as of the time that this Article was written. S. 380, 98th Cong., 1st Sess. (1983). This bill will henceforth be referred to as S. 380. The House Committee on the Judiciary also considered a number of bills during the Ninety-seventh session, but a substantially different bill recently has been introduced in committee. H.R. 2244, 98th Cong., 1st Sess. § 3 (1983), *reprinted in* 44 ANTITRUST & TRADE REG. REP. (BNA) No. 1107, at 673 (Mar. 24, 1983).

The authors' opinion is that because the final version of the contribution legislation probably will incorporate provisions from a number of different bills, a discussion of the relevant bills introduced in the Ninety-Seventh Congress will be helpful to our readers and also will provide a comparative perspective.

The two Senate bills provide as follows:

> Clayton Act (15 U.S.C. §§ 12–27 (1976)) is amended by inserting after section 4H the following new section:
>
> Sec. 4I. (a) Two or more persons who are subject to liability for damages attributable to an agreement to fix, maintain, or stabilize prices under section 4, 4A, or 4C of this Act may claim contribution among them according to the damages attributable to each such person's sales or purchases of goods or services. A claim for contribution by such person or persons against whom an action has been commenced may be asserted by cross-claim, counterclaim, third-party claim, or in a separate action, whether or not an action has been

A. *The Price-Fixing Bills*

Three legislative provisions share close similarities in substance and procedure and differ only in their application to pending antitrust cases. H.R. 4072[40] and H.R. 1242[41] were considered by the House Judiciary Committee last session, and S. 380[42] is presently before the Senate.

The House bills were expressly applicable only to horizontal price-fixing agreements, while S. 380 is cast in terms of price-fixing, and thus possibly covers resale price maintenance. All three bills provide a right of contribution according to the dam-

---

brought or a judgment has been rendered against the persons from whom contribution is sought.

(b) A release or a covenant not to sue or not to enforce a judgment received in settlement by one of two or more persons subject to contribution under this section shall not discharge any other persons from liability unless its terms expressly so provide. The court shall reduce the claim of the person giving the release or covenant against other persons subject to liability by the greatest of: (1) any amount stipulated by the release or covenant, (2) the amount of consideration paid for it, or (3) treble the actual damages attributable to the settling person's sales or purchases of goods or services. Under item (3) above, actual damages shall not be trebled in proceedings under section 4A of this Act.

(c) A release or covenant, or an agreement which provides for a release or covenant, entered into in good faith, relieves the recipient from liability to any other person for contribution, with respect to the claim of the person giving the release or covenant, or agreement, unless the settlement provided for in any such release, covenant, or agreement is not consummated.

(d) Nothing in this section shall affect the joint and several liability of any person who enters into an agreement to fix, maintain, or stabilize prices.

(e) This section shall apply to all actions under section 4, 4A, or 4C of this Act commenced after the date of enactment of this section.

(f)(1) The claim reduction principle of subsection (b) of this section shall also apply to actions alleging an agreement to fix, maintain, or stabilize prices under section 4, 4A, or 4C of this Act, which are pending on the date of enactment of this section, if upon proof by any party subject to liability for damages in such an action, the court determines that it would be inequitable, in light of all the circumstances and notwithstanding subsection (f)(2), not to apply the principle in that action. In ruling on a request to apply claim reduction, the court shall find the facts specially.

(f)(2) No agreement to settle, compromise, or release a claim under section 4, 4A, or 4C of this Act which has been signed by the parties prior to the date of enactment of this section may be rescinded, disapproved, reformed, or modified by the parties or by the court because of the application of the claim reduction principle, except upon the written consent of all the parties thereto.

(g) Each subsection of this section is severable from all other subsections, and the invalidity of any subsection for any reason shall not affect the validity of the remaining subsections: Provided, that subsections (f)(1) and (f)(2) are not severable from each other, and the invalidity of any provision of those subsections as applied in an action shall render the remainder of those subsections inapplicable in that action."

[40]97th Cong., 1st Sess. (1981).

[41]97th Cong., 1st Sess. (1981).

[42]98th Cong., 1st Sess. (1983).

ages attributable to each firm's sales or purchases of goods and services. The bills also permit the assertion of contribution claims via a number of procedures.[43]

Where the plaintiff settles with one or more defendants, all three bills would require the reduction of a plaintiff's claim for damages by the greatest of: (1) the amount stipulated in the settlement agreement, (2) the actual settlement figure paid by the settling party or parties, or (3) treble the actual damages attributable to the settling party's sales or purchases. Obviously, it seems probable that the third provision will be invoked most frequently.[44]

The bills all provide that any defendant who settles in good faith would not be liable for contribution to any other person concerning the claim that was settled. In the Senate bill this provision is intended to encourage settlements by ensuring that any settling defendants would not be liable in contribution claims asserted by nonsettling defendants.[45]

Finally, the bills differ somewhat with respect to their retroactive effect. H.R. 1242 would apply only to actions filed after the bill was enacted, while S. 380 and H.R. 4072 are expressly applicable to pending cases. The claim reduction features of H.R. 4072 would be applicable to pending cases except where they "would result in manifest injustice."[46] S. 380 has a more elaborate retroactivity provision that would permit application of the claim reduction provision to pending cases if "the court determines that it would be inequitable, in light of all the circumstances . . . , not to apply the principle in that action."[47] A closely related provision of S. 380 conditions the validity of the retroactivity provision on a paragraph of the bill that protects settlement agreements signed before the effective date of the act.[48]

---

[43]These proposals, like most of the contribution and claim reduction legislation considered by the Congress, provide that a party seeking contribution may assert a claim by counterclaim, cross-claim, third-party action, or a separate lawsuit. Although the proposals do not articulate the circumstances in which severance or joinder of claims is appropriate, it has been argued that the federal district courts have ample authority to dictate the appropriate procedure. *See* Jacobson, *supra* note 5, at 235, 239–40.

[44]S. REP. NO. 359, 97th Cong., 2d Sess. 47 (1982) (supplemental views of Sens. Kennedy and Metzenbaum).

[45]*Id.* at 29–30.

[46]H.R. 4072, 97th Cong., 1st Sess. § 4I(e) (1981).

[47]S. 380, 98th Cong., 1st Sess. § 4I(f)(1) (1983).

[48]According to the committee report on S. 995, subsection 4I(g) of S. 380 makes the dual retroactivity provisions of subsection 4I(f) unseverable from one another in the event either section is struck down. S. REP. No. 359, 97th Cong., 2d Sess. 15, 30–31

B. *Vertical and Horizontal Conspiracy Bills*

The Justice Department submitted one bill, H.R. 5794[49], to the House, and the House Committee on the Judiciary prepared two additional bills that modified the Justice Department bill in significant ways. A fourth bill, H.R. 2244, recently has been introduced in the House Judiciary Committee. These four bills would be applicable to both vertical and horizontal antitrust activity.

1. H.R. 5794. This bill is significantly different from the three preceding bills, both in substance and procedure. First, this bill expressly reaches any antitrust activity, not just horizontal price-fixing.[50] Second, the bill sets forth a definite statute of limitations of six months following final judgment in the litigation for which contribution is sought.[51] Third, the bill provides a bar against contribution claims by or against a settling defendant.[52]

With respect to claim reduction, H.R. 5794 requires a reduction of the plaintiff's claim by the greatest of: (1) the stipulated amount, (2) the actual consideration paid, or (3) the contribution share of the settling party.[53] The bill further provides that contribution and claim reduction must, "to the extent consistent with the fair and expeditious conduct of litigation," be determined by the court, and not a jury, after trial of the antitrust case.[54]

In establishing contribution and claim reduction shares and rights, the proposed bill distinguishes horizontal price-fixing from other antitrust violations. Contribution shares in price-fixing cases are determined "on the basis of the relative magnitude in the affected market of each such competitor's sales or purchases."[55] Contribution shares in other claims shall be determined by the court according to the "relative responsibility of each party for the origination or perpetration of the

---

(1982). This unseverability feature apparently serves the purpose of ensuring that if the retroactive claim reduction provision of subsection 4I(f)(1) is declared unconstitutional, then subsection 4I(f)(2), which protects settlements previously reached in pending cases, will also be severed. *Id.* at 30–31; *see also* G. BELL, ANTITRUST CONTRIBUTION AND CLAIM REDUCTION: AN OBJECTIVE ASSESSMENT 19 (1982).

[49]H.R. 5794, 97th Cong., 2d Sess. (1982).

[50]*Id.* § 4I(a).

[51]*Id.* § 4I(c).

[52]*Id.* § 4I(d).

[53]*Id.* § 4I(e).

[54]*Id.* § 4I(g).

[55]*Id.* § 4I(f)(2).

violation"[56] for which the plaintiff has been harmed or the defendants unjustly enriched.

2. Discussion Draft No. 1.[57] This bill is a modified version of H.R. 5794,[58] and materially differs from H.R. 5794 only with respect to the liability of a settling defendant to contribution claims. It provides that a settling defendant may not claim contribution, and it gives a settling plaintiff and defendant a choice. First, the parties may, as part of the settlement, agree that the plaintiff waives any claim against other defendants for the settling defendant's contribution share. The settling defendant would not be subject to any contribution claims by other defendants. Alternatively, the plaintiff could withhold the waiver and the settling defendant could be subject to contribution claims by other defendants. However, the bill further provides that the total damages collectible from nonsettling defendants shall be reduced by the greater of the stipulated settlement amount or the consideration paid for release.[59] In all other respects, the bill is identical to H.R. 5794.

3. Discussion Draft No. 2.[60] This bill is also a modification of H.R. 5794,[61] and differs from H.R. 5794 by its specification of contribution shares in vertical agreement cases. Significantly, the proposed language states that in vertical agreement cases contribution shares shall be allocated by "dividing the total damages equally between the functional levels and allocating shares within each level in accordance with the relative magnitude in the affected market of each person's sales or purchases . . . ."[62] This draft also permits the court to allocate liability among different levels in any just proportion, when necessary to avoid an unjust result. The bill apportions contribution claims in horizontal agreement cases on the basis of each defendant's relative magnitude and defines the term "relative magnitude" as

---

[56]*Id.* § 4I(f)(3).

[57]The text of this proposed bill is set forth in a letter by Congressman Peter Rodino, Chairman of the House Comm. on the Judiciary, to Att'y Gen. William French Smith, 42 ANTITRUST & TRADE REG. REP. (BNA) No. 1056, at 601–02 (1982) [hereinafter cited as Discussion Draft No. 1].

[58]*Id.* at 601.

[59]Discussion Draft No. 1, *supra* note 57, § 4I(f).

[60]The text of this proposed bill is set forth in a letter by Congressman Peter Rodino, Chairman of the House Comm. on the Judiciary, to Att'y Gen. William French Smith, 42 ANTITRUST & TRADE REG. REP. (BNA) No. 1056, at 603–04 (1982) [hereinafter cited as Discussion Draft No. 2].

[61]*Id.* at 601.

[62]Discussion Draft No. 2, *supra* note 60, § 4I(f)(1).

relative market shares in the relevant market.[63] In all other respects, this bill is identical to H.R. 5794.

4. H.R. 2244.[64] This bill, which recently was introduced in the House Judiciary Committee, applies to any antitrust action, but is significantly different from the other bills in several respects. This bill permits claim reduction or contribution of damages among defendants only if the district court, after determining the amount of damages, concludes that failure to reallocate or reduce damages "would be substantially unjust with respect to a defendant."[65] However, the bill also provides that claim reduction or contribution is permissible only where the plaintiff's conduct in the lawsuit "significantly contributed to the substantial injustice."[66] Finally, the bill establishes a floor on the amount of defendant's damages equal to the minimum of its relative fault or to its sales or purchases, which serves to limit the reduction or reallocation of damages.[67]

III. An Analytic Approach to Contribution Legislation

A contribution rule, as a matter of antitrust procedure, is relevant to the process of allocating damages, and, as a matter of policy, is relevant to the fundamental values underlying private antitrust actions. These procedural concerns and policy norms have been repeatedly considered and articulated by legal commentators and the United States Supreme Court, and they include the deterrent effect of private actions, judicial economy and litigation complexity, compensation for plaintiffs, and fairness to litigants. Because these considerations and values may be implicated by legislative proposals for contribution and claim reduction, it is necessary to consider the nature and effect of the proposed rules on private antitrust litigation.

A. *Deterrence*

The first factor concerns the deterrent effect of private antitrust actions. The Supreme Court has repeatedly stated that

[63] *Id.* § 4I(f)(2).
[64] H.R. 2244, 98th Cong., 1st Sess. § 3 (1983), *reprinted in* 44 Antitrust & Trade Reg. Rep. (BNA) No. 1107, at 673 (Mar. 24, 1983).
[65] *Id.* § 4J(a)(1).
[66] *Id.* § 4J(a)(2)(A).
[67] *Id.* § 4J(a)(2)(B).

private antitrust actions serve as a powerful deterrent to anticompetitive conduct. In the Court's recent antitrust decision, *American Society of Mechanical Engineers v. Hydrolevel Corp.*, Justice Blackmun stated that "[a] principal purpose of the antitrust private cause of action . . . is, of course, to deter anticompetitive practices."[68] In that case, the Court extended the scope of antitrust liability to nonprofit corporations when their voluntary agents engage in anticompetitive conduct. The Court made it clear that its decision to impose liability on such organizations was based largely on "the congressional intent that the private right of action deter antitrust violations."[69]

Supporters of the contribution rule argue that it would ensure that all violators would be subject to liability, and that this added certainty would enhance the deterrent effect of the antitrust laws.[70] The contrary argument is that allowing a defendant to redistribute liability reduces the consequences of a violation, and that this reduction in liability lessens deterrence.[71] The resolution of this question depends upon the attitude of the potential violator toward the risks involved. If a corporation prefers risking a low probability of a large penalty to a high probability of a small penalty, then a contribution rule should be favored, and vice versa. The absence of empirical data prevents a more definitive resolution of this issue.[72]

---

[68]102 S. Ct. 1935, 1945 (1982).

[69] *Id.* at 1946; *see also* Burlington Industries v. Milliken & Co., 690 F.2d 380, 392–93 (4th Cir. 1982) (discussing *Hydrolevel Corp.* decision in claim reduction context).

[70]There is a general body of literature which indicates that certainty of detection and punishment is a greater deterrent than severity of the penalty. *See, e.g.*, Andenaes, *The General Preventive Effects of Punishment*, 114 U. PA. L. REV. 949, 964 (1966); *see also* R. CYERT & J. MARCH, A BEHAVIORAL THEORY OF THE FIRM 119 (1963); O. WILLIAMSON, THE ECONOMICS OF DISCRETIONARY BEHAVIOR 48 (1978); Becker, *Crime and Punishment: An Economic Approach*, in ESSAYS IN THE ECONOMICS OF CRIME AND PUNISHMENT 9–11 (1974); Coffee, *Beyond the Shut-Eyed Sentry: Toward a Theoretical View of Corporate Misconduct and an Effective Legal Response*, 63 VA. L. REV. 1099 (1977); Ehrlich, *Participation in Illegitimate Activities: An Economic Analysis*, in ESSAYS IN THE ECONOMICS OF CRIME AND PUNISHMENT 111 (1974); Polinsky & Shavell, *supra* note 5, at 447, 449 (1981); Polinsky & Shavell, *The Optimal Tradeoff between the Probability and Magnitude of Fines*, 69 AM. ECON. REV. 880 (1979). *See generally* Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 168 (1968); Becker & Stigler, *Law Enforcement, Malfeasance, and Compensation of Enforcers*, 3 J. LEG. STUD. 1 (1974); Wheeler, *Antitrust Treble-Damage Actions: Do They Work?*, 61 CALIF. L. REV. 1319 (1973).

[71]K. ELZINGA & W. BREIT, THE ANTITRUST PENALTIES 128–29 (1976); Breit & Elzinga, *Antitrust Penalties and Attitudes Toward Risk: An Economic Analysis*, 86 HARV. L. REV. 693, 705–06 (1973).

[72]Whether a potential antitrust wrongdoer is deterred more by a small prospect of an increased damage award or by a greater certainty of a lower penalty is unclear. The recent economic analysis which has been done using the "expected utility" model of predicting behavior is equivocal. *See, e.g.*, Easterbrook, Landes, & Posner, *supra* note 5, at 344–53, 364–65. The question remains whether the *certainty* of detection and

B. *Complexity*

The second factor concerns complexity. The enactment of a contribution rule would add complications to court proceedings that already are intricate and expensive. The existing complexity in private antitrust actions may be attributable to the number of parties in the lawsuit, their difficulties in explaining and supporting their claims, and the courts' difficulties in deciding on appropriate relief. The pressures on the courts resulting from these and other issues have caused the Supreme Court to express concern over the "feasibility and consequences" of various damage theories[73] and to redefine traditional standards of statutory standing, remoteness, and causation in private antitrust action.[74] These cases, which strongly reflect growing judicial concern over litigation complexity, speculative proof of damages, and the unfairness of duplicative recoveries, counsel caution in adding new rights of action and novel approaches to the computation of damages to a litigation system already known for its complexity and delay.

The effects on the judicial system of the creation of contribution rights can perhaps best be understood within the analytical framework of transaction costs.[75] If contribution legislation

---

punishment is a greater deterrent than *severity* of the penalty. *See, e.g.*, STAFF OF THE SUBCOMM. ON MONOPOLIES AND COMMERCIAL LAW, HOUSE COMM. ON THE JUDICIARY, 98TH CONG., 1ST SESS., PROPOSED LEGISLATION TO ALLOCATE DAMAGES AMONG DEFENDANTS IN PRIVATE ANTITRUST LITIGATION, *reprinted in* 44 ANTITRUST & TRADE REG. REP. (BNA) No. 1101, at 280, 288 (Feb. 10, 1983) [hereinafter cited as STAFF REPORT]. Notwithstanding the equivocal nature of the theoretical work and the lack of empirical work, certain tentative observations can be drawn concerning the effect of contribution and claim reduction legislation on deterrence objectives.

[73]Blue Shield v. McCready, 102 S. Ct. 2540, 2546 n.11 (1982).

[74]*Id.*; Illinois Brick Co. v. Illinois, 431 U.S. 720 (1977); Brunswick Corp. v. Pueblo Bowl-O-Mat, 429 U.S. 477 (1977). *See generally* Berger & Bernstein, *An Analytical Framework for Antitrust Standing*, 86 YALE L.J. 809 (1977); Page, *Antitrust Damages and Economic Efficiency: An Approach to Antitrust Injury*, 47 U. CHI. L. REV. 467 (1980).

[75]The term "transaction cost" is capable of many meanings. *See* Polinsky, *Economic Analysis as a Potentially Defective Product: A Buyer's Guide to Posner's Economic Analysis of Law*, 87 HARV. L. REV. 1655, 1667–68, 1671–74 (1974). This is especially true in the contribution and claim reduction context. As a general matter, transaction costs are incurred whenever a group of people get together to bargain or settle. Types of transaction costs include process costs that are associated with the incremental expenses of administering a rule, such as additional judge or jury time; coordination costs associated with additional efforts at generating consensus among many parties to a transaction; insurance costs associated, for example, with the need to prevent cheating on a negotiated agreement; and error costs associated with the greater prospect of incorrect outcomes as complexity of coordination and the number of parties increase. *See* R. POSNER, ECONOMIC ANALYSIS OF LAW 45–47, 423, 458 (2d ed. 1977); Calabresi, *Transaction Costs, Resource Allocation and Liability Rules--A Comment*, 11 J.L. &

is enacted, a new set of "rights" will be created, and the working out of these "rights" at different levels of the system will entail additional costs. To the extent that many of the issues associated with contribution and claim reduction are already raised in antitrust litigation,[76] however, the consequences to the judicial system are reduced. Nevertheless, the addition of this right of action could have significant repercussions on both in-court and out-of-court costs.

A contribution rule will increase the complexity of antitrust lawsuits by bringing additional parties into the litigation. Under the traditional rule, the plaintiff could sue one coconspirator, litigate his case, and either win or lose. After a contribution rule is enacted, however, a defendant will be able to implead other coconspirators as third-party defendants.[77] Merely having additional parties in the litigation will increase costs to some degree. Costs and complexity will increase further to the extent that impleaded parties raise defenses or claims not raised by the original defendant. However, joinder may not significantly increase discovery costs because the antitrust discovery net generally is broadly cast to encompass nondefendants that have knowledge leading to relevant information regarding the alleged violation.[78]

Additional complexities may arise in the course of allocating damages among defendants. If some or all of the impleaded third-party defendants were to be found liable, a hearing would be required to determine each defendant's contribution share. If a method other than the easily applied pro rata formula were to be chosen for the allocation of damages, this proceeding would become quite complex. Moreover, if contribution were sought through the filing of an independent action, rather than through joinder of parties in the main action, duplication of effort would result.

A final consideration with regard to complexity is the effect of a proposed contribution rule on settlements. The existence

---

ECON. 67 (1968); Dahlman, *The Problem of Externality*, 22 J.L. & ECON. 141, 142–50 (1979); Heymann, *The Problem of Coordination: Bargaining and Rule*, 86 HARV. L. REV. 797, 827–43 (1973).

[76]For example, if the plaintiff has decided to use market share analysis to prove his damages, the addition of a contribution rule that uses market share analysis to determine contribution shares would not significantly increase litigation costs. *See infra* note 106.

[77]Of course, if the plaintiff named all antitrust violators in the original complaint, a contribution rule would result in no additional costs. *See* Note, *Contribution in Private Antitrust Actions*, 93 HARV. L. REV. 1540, 1549 (1980).

[78]Jacobson, *supra* note 5, at 235.

of additional defendants means that more settlements will have to be reached, and the relationship between the defendants may create uncertainties in the minds of both plaintiff and defendant that will make settlements harder to reach.[79] A contribution rule also will affect the parties' incentives to settle. It is frequently stated that private antitrust actions and the prospect of treble damage recoveries provide a substantial incentive to settle litigation.[80] If a defendant is permitted to shift a portion of liability to another, then there is less incentive to settle before liability is established, and to the degree one is permitted to decrease damages owed, the incentive to settle is decreased correspondingly.[81] Any decrease in the settlement rate could have a severe impact on the efficiency of the courts, given the large number of antitrust filings and the scarcity of judicial resources.[82]

It should be noted that despite the fact that a contribution rule may increase the complexity of the litigation, federal trial courts have had substantial experience with contribution rules. Since 1933, the courts have successfully applied contribution

---

[79]*See supra* note 75.

[80]*See, e.g.*, Easterbrook, Landes & Posner, *supra* note 5, at 365.

[81]*Id.* at 365–67; Note, *A Case Against Contribution in Antitrust*, 58 TEX. L. REV. 961, 980–82 (1980).

Because a plaintiff can choose to settle with each defendant sequentially or with them all together, the analysis of settlement incentives becomes complicated. An important issue is whether a contribution rule will result in a higher rate of terminated litigation than a no-contribution rule, and this requires a consideration of incentives and disincentives in partial, global, and sequential settlements strategies.

Some tentative observations are possible. A policy favoring partial settlements might result in more settlements than a policy favoring global settlements. The first reason is that coordination costs are quite high in a global settlement. Second, there is a greater inducement to settle when one alleged coconspirator has settled and the remaining defendants may have to ante up an amount greater than their aliquot portion of the gains from illegal conduct. Moreover, it would seem that the effect of the sequential settlement strategy has been quite positive. *See* Withrow and Larm, *The "Big" Antitrust Case: 25 Years of Sisyphean Labor*, 62 CORNELL L. REV. 1 (1976). Former Attorney General Griffin Bell has contended that contribution and claim reduction favor "global" settlements because they remove the plaintiff's incentive to enter into partial settlements by eliminating market share "uncertainty" in plaintiff's negotiation and settlement. *See* G. BELL, *supra* note 48, at 14–15; *see also* Polinsky & Shavell, *supra* note 5, at 457–62. However, at this point there is no clear indication of the effect of contribution and claim reduction legislation on attaining global settlements, and there is no demonstrable indication that a rule favoring global settlements would be more effective in increasing pretrial settlements than the current sequential settlement strategy.

Finally, the Fourth Circuit expressly rejected Bell's arguments, concluding that antitrust law and policy favored partial settlements. Burlington Indus. v. Milliken & Co., 690 F.2d 380 (4th Cir. 1982).

[82]*See* Conrac Corp. v. AT&T, 546 F. Supp. 429, 432 (S.D.N.Y. 1982); *see also* Withrow & Larm, *supra* note 81, at 6, where the authors detail the tremendous number of private antitrust cases filed, pending, and resolved in the federal courts. Of particular importance is the authors' statement that approximately 40% of filed private actions are settled before litigation.

rules under the federal securities laws. Securities litigation is not significantly more complex than antitrust litigation.[83] Both types of cases generally raise complicated economic, factual, and theoretical issues in a multiparty setting. In addition, courts, while exercising diversity jurisdiction, are experienced in applying contribution rules under state law, and various procedural rules[84] have assisted trial courts in the management of complex cases.

C. *Compensation for Plaintiffs*

A primary concern of antitrust plaintiffs is undoubtedly the recovery of damages. The Supreme Court has consistently found that a congressional purpose to compensate victims is manifest within the antitrust laws.[85] Though this objective recently has come under attack,[86] it is clear that any congressional action that threatens plaintiffs' ability to recover their damages will be intensely scrutinized by the Court.

D. *Fairness to the Litigants*

The final factor to be considered is perhaps the most difficult one: fairness to the parties. Fairness issues are encountered in the analysis of a contribution and claim reduction statute when there is either a transfer of part of the plaintiffs' control over the litigation to the defendants, or a shift of monetary liability among defendants, or a change in relative strengths in the negotiation of settlements.[87] For example, and as discussed more fully below, permitting defendants to allocate monetary liability

---

[83]*See* Sullivan, *supra* note 5, at 398–401.

[84]FED. R. CIV. P. 42(b) specifically permits the courts to provide for separate trials on "any claim, crossclaim, counterclaim or third-party claims or issues" to further convenience or avoid prejudice. FED. R. CIV. P. 21 establishes trial court discretion to drop parties or to sever claims and proceed separately. In addition, the rules permit the submission of special verdict forms or interrogatories to the jury in complex cases. *See* FED. R. CIV. P. 52.

[85]*See, e.g.*, American Soc'y of Mechanical Eng'rs v. Hydrolevel Corp., 102 S. Ct. 1935 (1982); Pfizer, Inc. v. Government of India, 434 U.S. 308, 314 (1978).

[86]*See, e.g.*, Schwartz, *An Overview of the Economics of Antitrust Enforcement*, 68 GEO. L.J. 1075, 1091–96 (1980) (Compensation in antitrust cases should be viewed as an incentive to vindicate anticompetitive conduct rather than as compensation for injuries suffered.).

[87]*See* Easterbrook, Landes & Posner, *supra* note 5, at 339–44; STAFF REPORT, *supra* note 72, at 283–87.

on the basis of relative fault may be fair to the defendants inter sese; but reducing a victim's damage entitlement by some multiple of settlement may not be just as between the victim and the defendants.

Supreme Court decisions, although not entirely consistent, have articulated a need to consider issues of fairness presented by legislative enactments. For example, in *United States v. Reliable Transfer Co.*,[88] the Court held that the mere fact that contribution may complicate litigation, because of difficulty in determining degrees of fault, does not justify the unjust results caused by an equal division of damages. The Court observed that "congestion in the courts cannot justify a legal rule that produces unjust results in litigation simply to encourage speedy out-of-court accommodation."[89] Clearly, then, contribution legislation will require a balancing of economic realities and judicial constraints against "unjust results" in particular cases.

The adoption of contribution legislation would resolve the perceived unfairness resulting from the "joint and several liability" rule of antitrust common law. A single defendant, under such legislation, no longer would be held liable for the plaintiff's entire damage claim while coconspirators went free.[90] A number of other fairness questions would remain to be considered in the legislation, however. First, in designing a rule that affects both antitrust plaintiffs and defendants, Congress should be concerned with unduly restrictive procedures that limit the parties' ability to advance the action. Second, in large part antitrust liability follows intentional and avoidable conduct, and notions of fairness are stronger for innocent victims than for intentional wrongdoers.[91] Third, antitrust violations and antitrust litigation impose costs on society and Congress should consider carefully the fairness of any rule that increases those costs, even though the rule seems fair to the litigants.[92] Fourth, notions of fairness in the allocation of liability among defendants are stronger where a reallocation is necessary to achieve some broader public objective, such as preservation of small busi-

---

[88]421 U.S. 397, 407–08 (1975). *But cf.* Texas Indus. v. Radcliff Materials, 451 U.S. 630 (1981).

[89]*Reliable Transfer*, 421 U.S. at 408.

[90]*See* Schwartz, *supra* note 86, at 1077 (the possibility of an error is a "process cost"). *See generally* R. POSNER, *supra* note 57, at § 21.2.

[91]Easterbrook, Landes & Posner, *supra* note 5, at 339–44.

[92]*See* Pfizer, Inc. v. Government of India, 434 U.S. 308, 314–15 (1978).

nesses.[93] Such public objectives, however, should be clearly articulated and closely related to the statutory alteration.

Ultimately, serious consideration of contribution legislation requires a sensitive balancing of procedural difficulties and policy norms. This effort, in an imperfect world, should attempt to achieve equitable ends without disproportionately increasing transaction costs, and should involve a search for the best procedural mechanism to implement antitrust values in an efficient and fair manner.

IV. Contribution Legislation: An Analysis

The legislative proposals surveyed in Section II will have various effects on antitrust policy. As the Supreme Court's decisions implicitly recognize, there is no commonly recognized hierarchy of values—only an array of factors, differing in weights and importance, from which decisions may be made.[94] It is therefore difficult to assign specific values or weights to factors that are implicated by a legislative proposal for contribution rights. However, there may be value to the decisionmaker in the identification of the complexities and incongruities created by a statute, as well as an assessment of a statute's probable likelihood of achieving its intended results. This Section analyzes the contribution proposals in light of certain characteristics that bring into question the ability of the statutes to fairly and efficiently advance the goals of antitrust law.

A. *Coverage of the Right to Contribution*

Several proposed bills extend the right to obtain contribution only to defendants in horizontal price-fixing cases.[95] This narrow coverage seems counterintuitive, for price-fixers receive little sympathy from the courts.[96] For example, in *Arizona v. Maricopa County Medical Society*,[97] a case involving allegations of

[93]*See* Reiter v. Sonotone Corp., 442 U.S. 330, 344–45 (1979).

[94]Blue Shield v. McCready, 102 S. Ct. 2540 (1982); *see* American Soc'y of Mechanical Eng'rs v. Hydrolevel Corp., 102 S. Ct. 1935 (1982).

[95]*See, e.g.*, S. 380, 98th Cong., 1st Sess. § 4I(a) (1983); H.R. 1242, 97th Cong, 1st Sess. § 4I(a) (1981).

[96]Catalano, Inc. v. Target Sales, 446 U.S. 643 (1980); Kiefer-Stewart Co. v. Joseph E. Seagram & Sons, 340 U.S. 211 (1951); United States v. Socony-Vacuum Oil Co., 310 U.S. 150 (1940).

[97]102 S. Ct. 2466 (1982).

horizontal price-fixing among physicians, the Court reiterated its longstanding position that horizontal price-fixing lacks any redeeming value and should receive the most stringent antitrust scrutiny. Price-fixing defendants therefore may be unworthy of any special consideration with regard to the distribution of damages, and to the extent that a contribution rule reduces deterrence, it is most unjustifiable here.[98]

The limitation of a right to contribution to price-fixers excludes several classes of defendants that deserve more favorable treatment. For example, in dealer termination cases,[99] where a vertically related party (e.g., manufacturer or supplier) may be the most culpable actor, the defendant may not be permitted to assert contribution rights against the unnamed manufacturer or supplier.[100] Moreover, vertical and horizontal arrangements should be distinguished, as they are in large part in standards of antitrust liability,[101] because beneficial effects often accompany vertical nonprice restrictions, and the harshness of the no-contribution rule may unnecessarily inhibit otherwise procompetitive conduct.[102]

---

[98] Note, *supra* note 81, at 971–77 (1980).

Former Attorney General Griffin Bell, a principal proponent of legislation like S. 380, has failed to address the rather perverse set of priorities seemingly established by the legislation. *See* G. BELL, *supra* note 48, at 11–12. He first makes the theoretical argument that a contribution and claim reduction regime may reduce price-fixing because the current no-contribution scheme may encourage price-fixing by letting some coconspirators enter into cheap settlement agreements, with the implicit encouragement of price-fixing activity. *Id.* at 11. Bell then dismisses "economic theorists" who conclude to the contrary, as not providing "any useful guidance on the question of deterrence." *Id.* at 12. Bell then suggests that deterrence may not be an appropriate objective of private actions because a great deal of price-fixing is perpetrated by "low-level employees" who, by a single conversation about prices with a competitor, can make an employer liable for the total damages caused by a large conspiracy. *Id.*

Bell's analysis is reminiscent of the "tail-wagging-the-dog" metaphor because the purpose of a deterrence function to an enactment is to compel an employer to take a more active and meaningful supervisory role over employees to prevent violation of public laws. *See* American Soc'y of Mechanical Eng'rs v. Hydrolevel Corp., 102 S. Ct. 1935 (1982). Moreover, it suggests that some proponents of contribution and claim reduction legislation are more interested in changing standards of liability, particularly the per se rule, than in serious consideration of the antitrust implications of the legislation.

[99]This term refers to a a situation where a manufacturer has refused to sell to a dealer because the dealer did not act in accordance with the manufacturer's wishes with respect to the product.

[100]The factual situation in Professional Beauty Supply v. National Beauty Supply, 594 F.2d 1179 (8th Cir. 1979), presents a classic illustration of an unfair or unjust result where contribution or claim reduction is not permitted in a vertical restriction case. *See* S. REP. No. 359, 97th Cong., 2d Sess. 42 (1982) (supplemental views of Sens. Metzenbaum and Kennedy); Sullivan, *supra* note 5, at 418–19.

[101]Sullivan, *supra* note 5, at 418–19.

[102]Easterbrook, Landes & Posner, *supra* note 5, at 367–68; Sullivan, *supra* note 5, at 418–19.

Finally, providing contribution rights to price-fixing cartel members may provide a strong incentive for refining the cartel agreement for prospective apportionment of damages in the event of detection.[103] This form of transaction cost associated with maintenance of a cartel agreement is an additional cost borne by consumers and other cartel victims.

B. *Computation of Contribution Shares*

One critical aspect of several contribution and claim reduction bills is the reliance on market share analysis to compute contribution shares and reduced claims. For example, H.R. 5794 requires contribution shares to be computed "on the basis of the relative magnitude in the affected market of each competitor's sales or purchases of goods or services,"[104] while H.R. 1242 requires computation "according to the damages attributable to each [price-fixer's] sales or purchases of goods or services."[105] The use of market share analysis as an integral part of a right to contribution or a defense of claim reduction is inappropriate for at least two reasons. First, a plaintiff may choose not to prove damages by market share analysis, but rather may select

---

The teaching of the Supreme Court's decision in Continental T.V. v. GTE Sylvania, 433 U.S. 36 (1977), is that strong policy considerations exist for treating vertical nonprice restrictions different than horizontal restrictions. These considerations, according to the Court, include economic justifications for vertical restrictions and a greater likelihood that such restrictions will achieve procompetitive results. *See* Pitofsky, *The* Sylvania *Case: Antitrust Analysis of Non-Price Vertical Restrictions*, 78 COLUM. L. REV. 1 (1978); Posner, *The Rule of Reason and the Economic Approach: Reflections on the* Sylvania *Decision*, 45 U. CHI. L. REV. 1 (1977).

Former Attorney General Griffin Bell contends, however, that a strong case can be made for extending contribution rights to price-fixers. He argues that a plaintiff may choose to settle with some defendants at a discount because he knows that he can recover the remainder of his damages from the nonsettling defendants. This results in a transfer of capital from nonsettling defendants to settling defendants and will force the nonsettling defendants to charge higher prices, which results in a reduction of consumer welfare. G. BELL, *supra* note 48, at 7–9.

This analysis, however, assumes that the cartel has made no supracompetitive profits with which to pay the judgment, and that supracompetitive profits obtained by price-fixing is a socially acceptable source of capital. If the firms which do not settle have not stored away their share of the cartel's supracompetitive profits, however, the excess liability which they have to bear may drive them out of the industry. Finally, as an empirical matter, the Staff of the House Subcommittee on Monopolies and Commercial Law found no cases "in which a small, relatively less culpable defendant has paid an outlandish judgment resulting in bankruptcy." STAFF REPORT, *supra* note 72, at 284.

[103]*See generally* Hay & Kelly, *An Empirical Survey of Price Fixing Conspiracies*, 17 J.L. & ECON. 13 (1974).

[104]H.R. 5794, 97th Cong., 2d Sess. § 4I(f)(2) (1982).

[105]H.R. 1242, 97th Cong., 1st Sess. § 4I(b) (1981).

some other method of computing damages in an antitrust case.[106] If a plaintiff chooses to prove damages by a "before and after" method, for example, but settles with one defendant before trial, the plaintiff may find his total claim reduced by "treble the actual damages attributable to the settling person's sales or purchases of goods or services."[107] It seems perverse to grant plaintiffs a choice of damage remedies and then remove that choice by reducing claims according to market share.[108]

Second, the use of market share as the method of determining contribution shares injects a great deal of complexity into antitrust actions.[109] The relevant product market must first be defined, which requires a showing as to the existence and effectiveness of substitutes. A delineation of the appropriate geographic market may then be necessary.[110] Evidence then

---

[106]Hoyt, Dahl & Gibson, *Comprehensive Models for Assessing Lost Profits to Antitrust Plaintiffs*, 60 MINN. L. REV. 1233 (1976).

There are essentially four methods of computing damages in a private antitrust case: "before and after" profits, *see, e.g.*, Bigelow v. RKO Radio Pictures, 327 U.S. 251 (1976); Eastman Kodak Co. v. Southern Photo Materials Co., 273 U.S. 359 (1927); a "yardstick" theory, *see, e.g.*, Joseph E. Seagram & Sons v. Hawaiian Oke & Liquors, 416 F.2d 71 (9th Cir. 1969), *cert. denied*, 396 U.S. 1062 (1970); foregone profits or loss of going concern value, *see, e.g.*, Farmington Dowel Prods. v. Forster Mfg., 421 F.2d 61 (1st Cir. 1970); and a market share theory, *see, e.g.*, Zenith Radio Corp. v. Hazeltine Research, 395 U.S. 100 (1969). *See generally* Parker, *Measure of Damages in Federal Treble Damage Actions*, 17 ANTITRUST BULL. 497 (1972); Note, *Measure of Damages for Destruction of All or Part of A Business*, 80 HARV. L. REV. 1566 (1967).

[107]H.R. 1242, 97th Cong., 1st Sess. § 4I(b) (1981).

[108]The imposition of market share analysis in a case where the plaintiff contemplates proof of damages by some other method presents the equitable problem of the plaintiff's potential loss of control over negotiation and settlement strategy and over the methods of discovery and proof of damages. There also may be concerns by antitrust plaintiffs that market share contribution and claim reduction computations may permit more numerous and well-financed defendants to overwhelm the plaintiff in the pretrial and trial process.

There are also transaction cost issues. A mandatory computation method may increase the possibility of error by plaintiffs in negotiating settlements, and it may increase discovery costs by the requiring additional efforts to determine relative market shares. Finally, process costs may be increased because a district court may have to sever the main antitrust case, in which damages and injury are demonstrated by one method, from the contribution case, in which damages must be reexamined using a market share methodology for allocation of damages.

To some extent, the establishment of a separate hearing on the computation of market shares, to the extent they are not determined in the main antitrust action, may ameloriate some of the transaction costs facing antitrust plaintiffs. For example, H.R. 5794 requires such a post-trial hearing by the courts. Such a hearing, however, would not clearly address the problems of ascertaining settlement options and could very likely increase process costs by the court in conducting two trials with two sets of evidence and methods of proving damages.

[109]*See* 4 P. AREEDA & D. TURNER, ANTITRUST LAW ¶¶ 908–22 (1980); Landes & Posner, *Market Power in Antitrust Cases*, 94 HARV. L. REV. 937, 939–52 (1981).

[110] P. AREEDA, ANTITRUST ANALYSIS ¶¶ 230–231 (3d ed. 1981).

must be gathered which pinpoints each defendant's involvement in each market.

The legislative proposals do little to resolve these questions. Only one proposal, Discussion Draft No. 2, defines "goods" and "purchases," and it appears that even those definitions cannot remove the fairly substantial problems left for judicial resolution.[111] It also seems probable that victims of antitrust conspiracies will be forced to address the definitional issues in the main antitrust action and that the courts will be faced with resolution of definitional disputes in adjudicating contribution shares.

There are some difficulties associated with use of the relative responsibility or comparative fault concept employed by H.R. 5794 and Discussion Draft No. 1. Those provisions advocate a relative responsibility method in all except horizontal price-fixing cases. The principal difficulty involves discerning each firm's complicity in a multimarket, vertical, national antitrust violation. Although a very strong case can be made for a comparative fault method of computing contribution shares,[112] it also appears that the method could increase complexity in the main antitrust case by requiring the parties, conceivably including the plaintiff, to litigate difficult issues of fault in a contribution hearing. It seems equally clear, however, that use of a relative fault method may be the fairest of all methods with respect to an allocation of responsiblity among defendants, may achieve the greatest deterrence effect of all contribution proposals, and may be the most familiar methodology because it is premised on traditional concepts of proof.[113]

One proposal, Discussion Draft No. 2, uses a pro rata contribution method in vertical agreement cases, but goes on to accord the district court latitude to allocate contribution shares in some other fashion where a pro rata method proves "mani-

---

[111] *See* S. REP. No. 359, 97th Cong., 2d Sess. 47–49 (1982).

[112] *See* Sullivan, *supra* note 5, at 416–23. The author details the equitable and policy considerations in a rule of comparative fault. Such a standard may require consideration of the actor's intent and should not be applied to traditionally per se antitrust violations. This selective application of a comparative fault contribution rule would maximize the deterrence effect of a contribution proposal and, if coupled with a reasonable claim reduction provision, would prove eminently fair to all parties.

[113] *Id.* Comparative fault or relative responsibility as a measure of contribution rights may achieve the greatest deterrence effect of the various proposals because of the closer relationship between the illegal conduct and the amount of the penalty. *Id.* at 421.

Secondly, federal judges are currently accustomed to making determinations of liability on the basis of fault. In this respect it appears that use of relative fault as a computation method may not increase process costs as much as other methods.

festly unjust." Some commentators on contribution legislation advocate pro rata contribution for the reason that such a rule is least expensive for the courts and the parties to administer and minimizes transaction costs.[114] A pro rata rule also may enhance the deterrent effect of the antitrust laws. If we assume that firms are risk averse,[115] the possibility that each will be forced to pay damages in excess of its responsibility will reduce the likelihood that it will violate the law.[116] On the other hand, the use of a pro rata method raises the issue of fairness among defendants, as each defendant must share liability equally, irrespective of culpability.[117]

H.R. 2244 advocates the use of relative magnitude and comparative fault methodologies to determine the floor on the amount of reduction or reallocation of damages. However, before these methods will be used, a culpable defendant must demonstrate that there is substantial injustice because he is paying a disproportionate amount of the awarded damages and that the plaintiff was somehow responsible for the injustice.[118]

C. *Claim Reduction*

Closely associated with the problems of computing contribution shares are problems of mandatory claim reduction. First, anything that reduces the absolute amount of liability facing each defendant may reduce the disincentive to engage in the proscribed conduct.[119] Therefore, the deterrence objective of the antitrust laws will be affected by any claim reduction legislation. Second, as the complexity of computing contribution shares increases, there may be a proportional increase in the complexity of negotiating and settling antitrust cases.[120]

---

[114]Easterbrook, Landes & Posner, *supra* note 5, at 365–66.

Transaction costs, under a pro rata computation method, are lower because individual market shares need not be ascertained, the judgment instead being shared in equal parts by the defendants. Nor would the fact finder need to make allocations of responsibility, as would be the case with a relative responsibility method. The court, after a finding of joint liability, would merely apportion the individual liability in equal shares.

[115]*See* Wilson P. Abraham Constr. Corp. v. Texas Indus., 604 F.2d 897, 901 n.8 (5th Cir. 1979), *aff'd. sub nom.* Texas Indus. v. Radcliff Materials, 451 U.S. 630 (1981).

[116]Easterbrook, Landes & Posner, *supra* note 5, at 365–66; Jacobson, *supra* note 5, at 242–45.

[117]*See supra* notes 68–74 and accompanying text.

[118]H.R. 2244, 98th Cong., 1st Sess. § 4J(a) (1983), *reprinted in* 44 ANTITRUST & TRADE REG. REP. (BNA) No. 1107, at 673 (Mar. 24, 1983).

[119]*See supra* notes 68–72 and accompanying text.

[120]*See supra* notes 73–84 and accompanying text.

Virtually all the contribution proposals provide that a settling defendant cannot be sued for contribution, but instead require a reduction in the plaintiff's claim against the remaining, nonsettling defendants. The provision forbidding contribution claims against settling defendants should facilitate the negotiation and settlement of antitrust claims by providing some finality to the prospective liability of a settling defendant, and therefore seems beneficial to antitrust legislation.[121]

Some difficulties are presented by the waiver or release provision contained in Discussion Draft No. 1.[122] That provision provides that the plaintiff may elect to release the settling defendant from future contribution claims by the amount of the settling defendant's contribution share. In that case, the amount of the settlement must equal or exceed the settling defendant's contribution share. Conversely, the plaintiff can withhold the waiver, which would permit nonsettling defendants to assert claims against the settling defendant for his contribution share. This provision obviously accords the plaintiff and settling defendant some latitude in negotiating settlements and accords some fairness among the defendants, but it also seems to inject potential impediments into the negotiation and settlement process. It appears that much of the impetus for contribution legislation stems from a few examples of plaintiffs extorting seriatim settlements from defendants.[123] Although these situations seem infrequent, it is apparent that the burden of accurately assessing a settlement strategy, and trading off the settlement price against the prospect of significant liability, is currently on antitrust defendants. The waiver principle in Discussion Draft No. 1 may shift the coercive impact onto the plaintiffs and may, as a practical matter, require a plaintiff to receive less in settle-

[121]A statutory release in an antitrust case bars subsequent recovery against a settling defendant, and it seems to accord with the common law rule. *See* Zenith Radio Corp. v. Hazeltine Research, 401 U.S. 321, 342–49 (1971). Obviously, such a rule is necessary in the context of contribution and claim reduction legislation where promotion of settlements and finality of litigation are important objectives. *See generally* Easterbrook, Landes & Posner, *supra* note 5, at 333–34.

[122]Discussion Draft No. 1, *supra* note 57, § 4I(e).

[123]*See* S. REP. No. 359, 97th Cong., 2d Sess. 3 (1982). In fact, the entire legislative history of the contribution and claim reduction legislation, beginning with S. 1468, 96th Cong., 1st Sess. (1979), through the current proposals, suggests intense pressure from defendants in a few significant antitrust cases. *See* S. REP. No. 359, 97th Cong., 2d Sess. 50 (1982) (supplemental views of Sens. Metzenbaum and Kennedy); *see also Antitrust Equal Enforcement Act of 1979: Hearings on S. 1468 Before the Subcomm. on Antitrust, Monopoly and Business Rights of the Senate Comm. on the Judiciary*, 96th Cong., 1st Sess. (1979) (reflecting significant interest from parties in cases partially affected by contribution and claim reduction legislation); G. BELL, *supra* note 48, at 2.

ment because a settling defendant wants to limit his exposure to a set sum, the settlement amount. It therefore seems that the waiver or release provision does not really add anything in resolving complexity and settlement problems; defendants will demand a full release and the plaintiff's recoverable damages will be reduced by the settling defendant's contribution share.

On the other hand, it may also be true that the waiver or release provision gives the plaintiff and settling defendant some flexibility in settlement decisions. Depending upon the strength of the plaintiff's case, the defendant's desire to avoid litigation, and the plaintiff's need for funds to advance the litigation, the choices may be freely exercised. Similarly, the waiver provision may allow the plaintiff to move the settlement price closer to the settling defendant's contribution share.[124]

The "substantial injustice" standard advocated in H.R. 2244, which reposes great discretion in the district court to reduce the plaintiff's damage award or to require contribution among defendants (the bill speaks in terms of "reallocation," but because "reallocation" and "contribution" are synonomous in this context, the latter will be used to maintain consistency) presents several problems. First, the House Subcommittee Report provides little guidance on how the substantial injustice standard is to be applied.[125] Furthermore, application of the standard may involve a separate procedural phase and thus may increase process costs in antitrust cases.[126] Insofar as the bill permits a

---

[124]*Cf.* G. BELL, *supra* note 48, at 13 (suggesting that the presence or absence of contribution and claim reduction rights may have less to do with the decision to settle than various other subjective factors).

[125]The House Subcommittee report accompanying the introduction of H.R. 2244 interprets the term "substantial injustice" to mean several "policy considerations" including:

> the amount of the judgment in relation to the role of the defendant, the strength of the evidence against the defendant, the seriousness of the violation, the degree of intent involved, the cooperation of the defendant in furnishing evidence and expediting the proceedings, [and] the amount of damages necessary to compensate the plaintiff.

STAFF REPORT, *supra* note 72, at 297.

[126]The Subcommittee Staff Report suggests that the court can make a finding of substantial injustice on the record after the determination of damages, and that no additional procedures will be necessary. *Id.* The report goes on to recognize that the defendant, in establishing a case of harm from disproportionate liability, could allege several factual matters, such as "relative size, lack of intent, weak evidence, [and] unclear legal standards." *Id.* at 298. The defendant, however, has the burden to make the necessary showing of disproportionality and injustice. *Id.* Such an difficult factual showing by defendants would necessitate a separate phase or proceeding to a concluded antitrust case, because parties may be reluctant to place these matters before the jury in the main action.

reduction in plaintiff's recovery below actual damages where plaintiff's conduct in the action significantly contributes to the substantial injustice, it injects greater uncertainty into the litigation. This provision seemingly serves the laudable goal of ameliorating situations of disproportionate liability when the plaintiff enters into "sweetheart" settlements with powerful, more culpable defendants and thereafter seeks to recover the full judgement from weaker defendants.[127] However, even the House Subcommittee Staff recognizes that instances of disproportionate damage awards are "rare," and this conclusion raises concerns about the applicability of the bill's contribution and claim reduction provisions in other antitrust cases.[128] Finally, H.R. 2244 provides that contribution or claim reduction may not decrease an individual defendant's liability below its relative culpability or market share, which grants power to the district court to alter the rule of joint and several liability in individual cases.[129] While the exercise of this provision may well affect the victim compensation objectives of the antitrust law, it may nevertheless advance its deterrence and fairness objectives. In the absence of a clearer demonstration of the bill's ramifications for antitrust litigation, and a better articulation of legislative objectives, it is difficult to assess the impact of H.R. 2244 on antitrust law and policy.

D. *The Right to a Jury Trial in Determining Contribution Shares*

Of the contribution proposals discussed in Section II, only H.R. 5794 and H.R. 2244 have special provisions regarding the use of jury trials in the apportionment of contribution shares. An earlier version of H.R. 5794 left the determination of the

---

[127]*See id.* Indeed, according to the Subcommittee Staff Report, the entire impetus of the bill is to protect defendants against inequitable and unfair results occasioned by settlement threats by allowing them to vindicate their innocence without coercion or fear of inordinate liability. *Id.* at 287–88.

[128]*Id.* at 288.

[129]The Subcommittee Staff Report recognized the conflicting aspects of the rule of joint and several liability. Initially, this rule threatens small defendants with damages judgments many times greater than their shares of the liability, while larger defendants face potential damage judgments only a few times greater than their actual liability. *Id.* at 288. However, any scheme that reduces damage judgments and increases litigation costs decreases the plaintiff's ability to obtain compensation. *Id.* at 298.

contribution issue to the judge alone.[130] Because of "possible Seventh Amendment problems,"[131] the following language was added to H.R. 5794: "*[u]nless inconsistent with the just and expeditious conduct of litigation*, contribution and claim reduction rights shall be determined by the court sitting without a jury."[132] H.R. 2244 adopts the same limitation on jury rights as H.R. 5794.[133]

Consistent with the Seventh Amendment's command that the right to a jury trial be "preserved,"[134] the courts decide whether or not to grant a jury trial on the basis of the common law practice as of 1791.[135] Intervening changes in our legal system have forced the courts to respond flexibly to new situations. Specifically, the merger of law and equity have forced the Court to reexamine the bases underlying the jury trial right. Under prior practice, for example, the "clean-up" doctrine allowed a chancellor to resolve legal issues in the process of resolving an equitable claim. In a pair of famous decisions, the Court disallowed this practice and held that legal issues should be determined by the jury, and that equitable issues should be resolved by the judge.[136]

The standard for determining whether a legal issue is presented was decided by the Supreme Court in *Ross v. Bernhard*.[137] The Court opined that "the 'legal' nature of an issue is determined by considering, first, the pre-merger [of law and

---

[130]Statement of William F. Baxter, Assistant Att'y Gen., Antitrust Div., Before the Subcomm. on Monopolies and Commercial Law, Comm. on the Judiciary, House of Rep., 97th Cong., Mar. 3, 1982, at 23.

[131]*Id.* at 23–24. Mr. Baxter further stated, "[Since] the submission of our original draft, we have made additional efforts to satisfy ourselves on the constitutionality of eliminating jury trials for contribution issues. Unfortunately, we have not been able to reach a firm conclusion."

[132]H.R. 5794, 97th Cong., 2d Sess. § 4I(g) (1982) (emphasis added).

[133]H.R. 2244, 98th Cong., 1st Sess. § 4J(b) (1983), *reprinted in* 44 ANTITRUST & TRADE REG. REP. (BNA) No. 1107, at 673 (Mar. 24, 1983).

[134]U.S. CONST. amend. VII.

[135]Baltimore & C. Line v. Redman, 295 U.S. 654, 657 (1935); NLRB v. Jones & Laughlin Steel Corp., 301 U.S. 1 (1937).

[136]Dairy Queen v. Wood, 369 U.S. 469 (1962); Beacon Theatres v. Westover, 359 U.S. 500 (1959). In *Beacon Theatres*, the Court rejected the so-called "clean-up" doctrine which permitted the court to hear legal claims without a jury when necessary in adjudicating equitable claims. 359 U.S. at 510–11. In *Dairy Queen*, the Court went further in establishing a nondiscretionary interpretation of the Seventh Amendment by holding that the issues need not be predominantly legal in nature before the jury right could be invoked. *Dairy Queen*, 369 U.S. at 473. If the claim was legal, regardless of its substantiality in relation to equitable claims, the right to a jury trial was recognized.

[137]396 U.S. 531 (1970).

equity] custom; second, the remedy sought; and, third, the practical abilities and limitation of juries."[138]

American courts seeking common law analogies for antitrust actions have looked to the law of torts.[139] The inquiry, therefore, is directed to the availability of a jury trial to determine contribution rights in a tort case in England in 1791.

Some American courts have disposed of the jury trial issue simply by stating that contribution is an equitable doctrine and therefore is not covered by the Seventh Amendment.[140] This analysis is, at best, too cursory. Admittedly there is little doubt that the doctrine of contribution had its *origins* in equity.[141] But in spite of its equitable origins, the doctrine of contribution also came to be recognized in the English *law* courts by the end of the eighteenth century.[142] A form of contribution among negligent joint tortfeasors was recognized as early as 1622 in the case of *Arundel v. Gardiner*.[143] Since tort cases in general, and contribution cases in particular, were decided by law courts, the historical analysis suggests that a jury trial would be available in contribution and claim reduction litigation.

---

[138]*Id.* at 538 n.10.

[139]*See, e.g.*, Washington v. American Pipe & Constr. Co., 280 F. Supp. 802, 804 (S.D. Cal.), *mandamus denied*, 393 F.2d 568 (9th Cir.), *cert. denied*, 393 U.S. 842 (1968). *But see* Note, *Contribution in Private Antitrust Suits*, 63 CORNELL L. REV. 682, 693, 696–97 (1978) (antitrust actions sound in quasi-contract rather than in tort). Technically, there was no premerger custom in the context of contribution and claim reduction rights in an antitrust case. Indeed, the adoption of H.R. 5794 would create such rights. The Supreme Court has rejected this purely historical method of analysis, however, and has made it clear that statutorily created rights are to be afforded Seventh Amendment protections, provided that the issues involved bear a close analogy to issues tried to English juries in 1791. Curtis v. Loether, 415 U.S. 189, 194 (1974).

[140]*See, e.g.*, Dawson v. Contractors Transp. Corp., 467 F.2d 727 (D.C. Cir. 1972); *see also* Jones v. Schramm, 436 F.2d 899, 901 (D.C. Cir. 1970):

> The doctrine of contribution originated in the courts of equity . . . . [W]hen contribution is sought against a defendant who was not sued by plaintiff, as is permitted by our decisions, the claim sounds in equity and the court acts as finder of the fact to determine whether the second tortfeasor from whom contribution is sought was negligent, and therefore liable to the victim.

[141]*See* Craythorne v. Swinburne, 14 Ves. Jun. 160, 33 Eng. Rep. 482 (1807); *see also* J. POMEROY, EQUITY JURISPRUDENCE § 1418 (1883).

[142]*See* Dawson v. Contractors Transp. Corp., 467 F.2d 727, 736–38 (D.C. Cir. 1972) (Fayh, J., dissenting); *see also* Cowell v. Edwards, 2 Bos. & Pul. 268, 126 Eng. Rep. 1275 (1800).

[143]Cro. Jac. 652, 79 Eng. Rep. 563 (1622). *Merryweather v. Nixon*, 8 T.R. 186, 101 Eng. Rep. 1337 (1799), is often cited for the proposition that there was no contribution among joint tortfeasors under the common law. However, *Merryweather* was an intentional tort case, and the later case of *Betts v. Gibbins* distinguished *Merryweather*: "*Merryweather* seems to me to have been strained beyond what the decision will bear. The present case is an exception to the general rule. The general rule is, that between wrongdoers there is neither indemnity nor contribution: the exception is, where the act is not clearly illegal in itself . . . ." 2 Ad. & E. 57, 74–77, 111 Eng. Rep. 22, 29–30 (1834).

The premerger custom is, however, only the beginning of the judicial analysis. Inquiry must also be directed to the nature of the remedy sought. If the suit is one in which legal rights are to be ascertained and determined, the Seventh Amendment applies regardless of the procedure chosen to bring suit.[144]

H.R. 5794 allows for considerable procedural flexibility in bringing the contribution action. The claim may be raised by cross-claim, counterclaim, or impleader, or in a separate action. Consider those situations where the plaintiff in the original antitrust suit chooses to proceed against some violators and the contribution issue is raised by cross-claim. The issue of joint liability must be litigated first because a contribution claim does not accrue until there has been a finding of joint or concurrent liability.[145] But the joint liability issue is usually a major part of the plaintiff's case. To establish a violation of section 1 of the Sherman Act it must be shown that the defendant conspired to engage in anticompetitive conduct. Because this is a legal issue, defendants would be afforded the right to a jury determination.

Consider next the case where only one of the conspirators is sued by the plaintiff. Suppose neither party demands a jury and the court determines that there has been a violation of the antitrust laws. The defendant, after satisfying the judgment, brings suit against the coconspirators for contribution. In the second suit a finding of joint liability is a critical part of the plaintiff's case.[146] Can the defending party in this second suit be denied a trial by jury on this issue? *Ross v. Bernhard*[147] instructs that the right to a jury should be determined according to the nature of the issue, irrespective of the procedure chosen. If the second defendant had a right to a jury trial in the case of a cross-claim, then he should not lose it simply because the original defendant chose not to implead additional defendants.

After a decision on joint liability is reached, the next issue for determination is the contribution shares.[148] Although the

---

[144]Ross v. Bernhard, 396 U.S. 531, 538 (1970).

[145]*See, e.g.* UNIF. CONTRIBUTION AMONG TORTFEASORS ACT § 1(a) (1955).

[146]In order to have a cause of action upon which relief can be granted, the plaintiff in the second suit must plead and prove that he is entitled to contribution from the defendant. Thus, the first determination to be made is whether the defendant is a tortfeasor at all.

[147]396 U.S. 531 (1970).

[148]Under H.R. 5794, the court, in fact, has very little discretion regarding the allocation method. Subsection 4I(f) of the bill prescribes a market share method in price-fixing cases, and a comparative fault method in all other cases.

method of allocating shares is a matter for the court to decide,[149] the proportion attributable to each violator could be a jury question in some circumstances. The most likely situation is where the allocation method chosen is that of comparative fault.[150] Fault determination has traditionally been a jury function. The issues involved in determining fault in the contribution case are identical to those tried in the main claim. The fact that the issues are cast in terms of contribution does not alter their basic nature. If the *Ross* "nature of the issue" test is interpreted literally, it would seem that the degree of comparative fault must be decided by a jury upon demand. With regard to the apportionment of contribution on a pro rata basis, there would be no need to convene a jury as the calculation would be a simple per capita division of damages. If the market share formula were used, structural issues such as product and geographic market definition would have to be determined. If these types of market structure issues existed at common law in England in 1791, they were determined by judges.[151] To the extent that the market share division cannot be stipulated by the parties, the common-law tradition strongly indicates that there is no right to a jury trial on the determination of market shares.

Also relevant to the nature of the remedy is the type of relief sought. Contribution is a demand for a money judgment. Money damages are traditionally a remedy granted by common-law courts.[152] Although money damages are the underlying relief sought, this is not a conclusive test of whether a jury trial is constitutionally required, though it does add support to the proposition.

---

[149]This assertion is based upon a two-stage historical analysis of the right to a jury trial as proposed by Professor Jorde. *See* Jorde, *The Seventh Amendment Right to Jury Trial of Antitrust Issues*, 69 CALIF. L. REV. 1 (1981). The history is clear that contribution, as a method of allocating damages, was never a matter of choice for the judge or jury. If the plaintiff were successful, the defendant would be assessed a pro rata share or be forced to bear the entire judgment (depending on the time period under consideration).

[150]Subsection 4I(f) of H.R. 5794 states in part: "In determining contribution shares with respect to all other claims, the court shall consider the relative responsibility of each party for the origination or perpetration of the violation for which damages have been awarded and the benefits derived therefrom." For an argument favoring the fault standard, see Sullivan, *supra* note 5.

[151]For a persuasive argument of this point, see Jorde, *supra* note 149.

[152]*See* Curtis v. Loether, 415 U.S. 189, 196 (1974) ("More important, the relief sought there—actual and punitive damages—is the traditional form of relief offered in the courts of law.").

The final factor to consider is complexity.[153] The central inquiry is whether contribution issues add so much complexity to an antitrust case that a jury could not competently decide it. As to the finding of joint liability, there appears to be no element of additional complexity. As described above, these issues must be litigated in the first instance to a jury in order to establish liability of any of the defendants. Extending a right to contribution should not add significantly novel litigation issues or strategies. Similarly, if contribution is based on a pro rata or market share standard and is decided by the judge, as discussed above, no increased complexity is added to the jury's burden.

The determination of contribution shares on a comparative fault basis could, however, add to the complexity. But before denying a jury trial on these grounds, the trial court should consider all the available procedural devices that might mitigate the complexity and weigh the possible results of such procedures against the likelihood that the remaining complexity would affect the jury's competence to render a rational decision. By exercising its discretion to invoke the rules of procedure, which permit separate trials of "any claim, cross-claim, or counterclaim or third party claims on issues,"[154] and which permit the court to isolate issues and claims by submitting special instructions, verdict forms, or interrogatories to the jury, the trial court could reduce the complexity of determining relative degrees of fault. Given judicial experience with comparative fault instructions and jury forms, the issue of contribution under a relative fault standard should not be unduly complicated.

E. *Retroactive Effect*

Two legislative provisions that have been considered by Congress expressly provide for retroactive application of contribution and claim reduction rights to those cases that are pending at the time of the bill's enactment. The Senate bills permit application of its claim reduction features to pending litigation if the party asserting a contribution or claim reduction right

[153]An extensive treatment of the Seventh Amendment problems in complex antitrust cases is beyond the scope of this Article. *See, e.g.*, Arnold, *A Historical Inquiry into the Right to Trial by Jury in Complex Civil Litigation*, 128 U. PA. L. REV. 829 (1980); Devlin, *Jury Trial of Complex Cases: English Practice at the Time of the Seventh Amendment*, 80 COLUM. L. REV. 43 (1980).

[154]FED. R. CIV. P. 21, 42(b), 49(b).

demonstrates that it would be inequitable not to grant the right retroactively. The bills further require a special finding of fact by the trial court on the retroactive application, and state that settlements and release provisions in settlement agreements signed before the effective date of the legislation shall not be disturbed.[155] H.R. 4072 simply provides that it shall apply to pending actions except where application "would result in manifest injustice."[156] The remaining proposals do not contain any provisions concerning the retroactive or prospective application of contribution and claim reduction rights.[157]

Before examining these bills to determine whether applying them retroactively would be unconstitutional, it may be helpful to consider discrete applications of the legislation. Initially, the claim reduction mechanism could reduce the total amount of damages that an antitrust plaintiff may recover against defendants. This situation would arise where the plaintiff settles with certain defendants, expecting to recover fully treble the actual damages proven at trial from the nonsettling defendants,[158] and

---

[155]*See* S. REP. No. 359, 97th Cong., 2d Sess. (1982); G. BELL, *supra* note 48, at 19.

[156]H.R. 4072, 97th Cong., 1st Sess. § 4I(e) (1981).

[157]There are essentially two issues: the effect of the statutory language in S. 380 and H.R. 4072, and the absence of any legislative indication in the remaining provisions concerning retroactive or prospective application.

The general rules that appear to operate in this area of constitutional concern are relatively indeterminate and are merely presumptive guidelines. Congress can, for example, provide that its legislation shall be applicable to pending cases, and such a provision is entitled to considerable weight. Further, in the absence of a statement that the legislation shall be prospectively applied, a statute may be applied to pending cases. *See* Bradley v. School Bd., 416 U.S. 696 (1974); Eikenberry v. Callahan, 653 F.2d 632 (D.C. Cir. 1981). Notwithstanding the congressional determination, however, the federal courts must always review specific instances of retroactive application for "manifest injustice." *Bradley*, 416 U.S. at 717.

Although the condemnation of retroactively applied legislation is not specifically articulated in the Constitution, the Supreme Court has interpreted the Contract Clause of Article I, Section 10 and the Due Process Clauses of the Fifth and Fourteenth Amendments as prohibiting retroactive application of legislation in certain situations. *See* Hochman, *The Supreme Court and the Constitutionality of Retroactive Legislation*, 73 HARV. L. REV. 692, 693–95 (1960). The Contract Clause, however, applies only to actions of the states.

[158]Under current standards a plaintiff in a private action is entitled to recover treble the actual damages demonstrated at trial less an offset from the trebled damages figure for any sums received in settlement. Zenith Radio Corp. v. Hazeltine Research, 401 U.S. 321, 348 (1971); Flintkote Co. v. Lysfjord, 246 F.2d 368, 397–98 (9th Cir.), *cert. denied*, 335 U.S. 835 (1957); Semke v. Enid Auto. Dealers Ass'n, 320 F. Supp. 445 (D. Okla. 1970).

In the recent case of Burlington Indus. v. Milliken & Co., 690 F.2d 380 (4th Cir. 1982), the court carefully examined the history of and justification for the rule that settlement payments are deducted after trebling actual damages, and concluded that the statutory history of section 4 of the Clayton Act and its underlying goals of deterrence and compensation for injured parties did not countenance claim reduction. *Id.* at 391–92.

the court orders a reduction in the plaintiff's recovery by treble the settling defendants' market share.[159] Because the plaintiff loses the amount of the "carve out," the plaintiff's expectations are defeated.

A second, although less likely, situation involving retroactivity problems concerns a disruption of settlement agreements where a plaintiff and some defendants settle for a stated figure and subsequently attempt to rescind the settlement agreement because the settlement figure was "inordinate" in light of the contribution legislation.[160] Although the right to rescind a settlement agreement for this type of mistake is unclear,[161] such a recission would require judicial consideration of the settling parties' claim that the policy behind the amendatory legislation is more important than the narrow interests of the contracting parties. The requirement that settling defendants involve themselves in the contribution and claim reduction phase of the case when they clearly expected to be released from the litigation raises similar issues. Although some or all of the information necessary to determine market shares and carve outs may already be in the possession of nonsettling defendants at the time of the settlements, it seems probable that settling parties cannot avoid further involvement in the contribution phase.[162]

---

[159]The contribution and claim reduction legislation under study commonly provides that in the event of prelitigation settlement, the plaintiff's claim for damages shall be reduced by the greatest of the stipulated amount, the consideration paid for settlement, or treble the actual damages attributable to the settling party's market share. *See* S. 380, 98th Cong., 1st Sess. § 4I(b) (1983); H.R. 4072, 97th Cong., 1st Sess. § 4I(b) (1981).

[160]For example, a defendant may enter into a settlement agreement for more than its proportionate share as determined under the market share or other formulas advanced by the contribution legislation. More frequently, however, a plaintiff may attempt to rescind a settlement agreement because of the effect of a subsequent carve out.

The drafters of S. 995, 97th Cong., 2d Sess. (1982), (now S. 380, 98th Cong., 1st Sess. (1983)), anticipated this problem and provided, in subsection 4I(f)(2), that settlements reached in pending cases may not be overturned or disturbed in any way. S. REP. No. 359, 97th Cong., 2d Sess. 26 (1982). It is clear that this provision is intended only to protect settling defendants against relitigation. The plaintiff still will suffer the effects of a reduction of his or her judgment, probably in the amount of treble the settlement amount. *Id.* at 51 (supplemental views of Sens. Metzenbaum and Kennedy); *see also* G. BELL, *supra* note 48, at 18–19.

[161]*See, e.g.*, G. BELL, *supra* note 48, at 18 (commercial frustration).

[162]The degree of involvement in post-settlement proceedings by settling defendants will depend upon a number of factors, such as the type of goods bought or sold, the pricing methodology for the goods, the settling defendants' complicity in the cartel, and the duration and complexity of the conspiracy. *See* S. REP. No. 359, 97th Cong., 2d Sess. 28 (1982). This is precisely the sort of situation that the drafters of S. 380 have attempted to consider in protecting settling defendants from the risk of relitigation. *See supra* note 160.

Finally, the courts may be presented with arguments from impleaded defendants in pending cases that the amendatory legislation created new rights of action against them.[163] Although this argument has more particular reference to the statute of limitation feature of contribution legislation than to the right of action feature,[164] it appears that such impleaded defendants may claim that the effect of legislation is to permit the assertion of new causes of action in the pending litigation.

The principal flaw in the retroactive application of legislation is that it impinges upon fundamental notions of the primacy of private expectations—either in the marketplace or as it affects individual freedoms—and public fairness—the notion that the government will be impartial in its dealings with its citizens and will provide neutral processes to implement impartial norms.[165] It is equally clear, however, that there may be public norms and interests that transcend the more secular concerns of individual expectations in due process and economic liberty.[166] Therefore, the Supreme Court has held that civil legislation may be retrospectively applied unless to do so would result in a clear constitutional violation.[167]

In a recent examination of the due process constraints on retroactive application of legislation, the Court structured the balancing analysis by requiring consideration of three factors: "(a) the nature and identity of the parties, (b) the nature of their

---

[163]For example, suppose a plaintiff files an antitrust action in 1979 against some coconspirators in a cartel. The plaintiff recovers a judgment against the named conspirators, but, prior to satisfaction of the judgment, a contribution and claim reduction bill is passed giving the named conspirators a right of action against the unnamed parties. In such a situation the named defendants would be entitled to file an action against the unnamed coconspirators seeking contribution. *Cf.* Professional Beauty Supply v. National Beauty Supply, 594 F.2d 1179 (8th Cir. 1979).

[164]For example, suppose a contribution and claim reduction bill is enacted one year after the defendant pays the judgment. The issue is whether or not the antitrust defendant may file a new action against a coconspirator who avoided the earlier litigation and assert a right to contribution.

Only three contribution proposals, H.R. 5794, 97th Cong., 2d Sess. (1982), Discussion Draft No. 1, *supra* note 57, and Discussion Draft No. 2, *supra* note 60, articulate a statute of limitations for asserting a contribution claim. The absence of statutory limitation provisions in the remaining bills presents substantive and procedural questions concerning the applicability of state law analogues to contribution claims and repose in antitrust cases. *See* Jacobson, *supra* note 5, at 240–41. Clearly, responsible legislative action would include a clear and short statute of limitations for the assertion of contribution claims.

[165]Hochman, *supra* note 157, at 692–97.

[166]L. TRIBE, AMERICAN CONSTITUTIONAL LAW § 10-1 (1978).

[167]*See, e.g.*, Usery v. Turner Elkhorn Mining Co., 428 U.S. 1 (1976); Welch v. Henry, 305 U.S. 134 (1938).

rights, and (c) the nature of the impact of the change in law upon those rights."[168] The nature of the parties, according to the Court, contemplates an assessment of the parties' status, incentives and ability to litigate, and, if they are private parties, whether or not they are attempting to vindicate public policies.[169] Thus, for example, where private parties attempt to vindicate rights under statutes proclaimed for their "great national concern," considerable weight should be placed on the side of retrospective application of the statute.[170]

In considering the nature of the preenactment rights, the Court has indicated its reluctance to retroactively apply statutory provisions where it would deprive parties of rights that had "matured or become unconditional."[171] Although this seems to contemplate the attachment of analytic importance to largely archaic notions of vested property rights, this inquiry looks both to whether or not the legislation in question attempts to fulfill or deny the parties' expectations and to the degree to which the parties' expectations have been realized.[172] For example, in *Bradley v. School Board*,[173] the plaintiff class of school children sued the local school board to compel desegregation of the public school system, and requested attorneys' fees under a federal statute. After pointing out that the plaintiffs were vindicating important matters of national policy, the Court stated

---

[168]Bradley v. School Bd., 416 U.S. 696, 717 (1974). *See generally* Hochman, *supra* note 157.

[169]*Bradley*, 416 U.S. at 717–18.

[170]*Id.* at 719 (citing United States v. Schooner Peggy, 5 U.S. (1 Cranch) 103, 110 (1801)).

[171]*Id.* at 720.

[172]*Id.*; *see also* Hochman, *supra* note 157, at 717–26.

In Usery v. Turner Elkhorn Mining Co., 428 U.S. 1 (1976), the Supreme Court considered a due process attack on federal black lung legislation that would have imposed certain financial responsibilities on coal mining firms for black lung benefits claimed by former employees. The mine operators complained that imposition of liability was impermissible because they could not have anticipated the imposition of liability at the time of employment. *Id.* at 15. The Court held that notwithstanding the possible reliance on the existing state of the law by the mine operators, retroactive application of legislation was "justified as a rational measure to spread the costs of the employees' disabilities to those who have profited from the fruits of their labor—the operators and the coal consumers." *Id.* at 18.

A similar method of analysis is used under the Contract Clause. U.S. CONST. art. I, § 10. Contracting parties, as a general rule, are entitled to rely on the law as it exists at the time of contracting. Home Bldg. & Loan Ass'n v. Blaisdell, 290 U.S. 398, 429–30 (1934). However, the legislature may vary or otherwise impair the substantive or remedial provisions of law that are relied upon by contracting parties where the legislative modification serves a legitimate public purpose. United States Trust Co. v. New Jersey, 431 U.S. 1, 21–22 (1977).

[173]416 U.S. 696, 720 (1974).

that the school board, which was appealing the retroactive application of the federal statute, had no "matured or unconditional right" in the fund provided by taxpayers.[174] According to the Court, the funds from which the attorneys' fees would be paid "were essentially held in trust for the public" and the school board "was subject to such conditions or instructions on the use of the funds as the public wished to make through its duly elected representatives."[175]

The third consideration, the influence of the change of law upon existing rights, concerns the degree to which retroactive application affects the parties' expectations and obligations, their opportunities for notice and a hearing, and the likelihood that the parties relied to their detriment on the preenactment state of affairs.[176] The closer the statutory provision comes to extinguishing the substance of a preexisting right, the less likely the statute is to be constitutional.[177]

The contribution bills that apply to pending cases present challenging questions of constitutionality. First, the plaintiffs are vindicating important governmental concerns by serving as private attorneys general.[178] Plaintiffs also will lose preexisting rights concerning the management of the litigation and the amount of the recovery. To the extent that settling defendants and newly impleaded defendants contend that retroactive application of contribution rights impair their rights under preenactment law, it seems that neither their status as antitrust wrongdoers nor their asserted interests provide compelling justifications for refusing retrospective application.[179] Similarly,

---

[174]*Id.*

[175]*Id.*

[176]*Id.* at 720–21; *see also* Hochman, *supra* note 157, at 711–17.

[177]Hochman, *supra* note 157, at 712.

[178]Hawaii v. Standard Oil Co., 405 U.S. 251 (1972); Hanover Shoe v. United Shoe Mach. Corp., 392 U.S. 481 (1968).

[179]*Cf.* Usery v. Turner Elkhorn Mining Co., 428 U.S. 1, 17–18 (1976).

The confluence between the status of the party asserting the right to statutory benefits and permissible governmental purposes was demonstrated in Simpson v. Union Oil Co., 396 U.S. 13 (1969). The plaintiff sued, claiming that the defendant's consignment agreements violated section 1 of the Sherman Act as a form of price-fixing. The Supreme Court agreed. Simpson v. Union Oil Co., 377 U.S. 13 (1964). In that decision, the Court reserved "the question whether, when all facts are known, there may be any equities that would warrant only prospective application in damage suits of the rule governing price-fixing by consignment device which we announce today." *Id.* at 24–25. On remand, the district court held that the rule announced in the Supreme Court's decision should be given prospective application because the defendant had a "reasonable basis for believing its actions were entirely lawful." 396 U.S. at 14. The Supreme Court reversed, holding that "[t]he question we reserved was not an invitation to deny the fruits of

an impleaded defendant cannot readily demonstrate any matured rights or expectations of immunity from suit for its complicity in an antitrust violation.

Settling parties, however, such as antitrust plaintiffs and defendants, may well have legitimate expectations as a result of a negotiated settlement, and more particularly, those parties may rely on the state of the law at the time they entered into their agreement.[180] It is further likely that retrospective application of the contribution bill will disrupt settled contractual and substantive obligations where a claim reduction provision requires a diminution of a plaintiff's prospect for damages and requires further involvement by settling parties in the contribution phase of the case.[181] In balancing the effects of a carve out on plaintiffs' expectations and on disruptions of settlement agreements against legislative purposes, it is clear that application of the claim reduction provisions to pending cases will not advance the deterrence function of the statute because the illegal conduct already has occurred. Moreover, retrospective application will frustrate legitimate compensatory interests of antitrust plaintiffs and endanger expectancy interests in settlement agreements.

Therefore, retroactive application of the contribution provisions to newly impleaded defendants will not present any substantial questions of unconstitutionality. Settling defendants who claim that the claim reduction provision is unconstitutional because it requires their additional effort in the contribution

---

successful litigation to this petitioner. Congress has determined the causes of action that arise from antitrust violations; and there has been an adjudication that a cause of action against respondent has been established." *Id.*

[180]As previously discussed, the reasonableness of plaintiffs' reliance is directly related to the permissibility of their expectations, the duration of the belief, and the legitimate government objective. *See supra* notes 172 & 179. In the case of contribution legislation, the objectives of the antitrust laws militate against retroactive application of the contribution legislation where such an application will disrupt previously negotiated settlements and will require settling defendants to participate in the relitigation of issues concerning contribution rights and shares. *See supra* notes 79–84 and accompanying text.

[181]S. 380, 98th Cong., 1st Sess. (1983), attempts to amelioriate some of the impact of the claim reduction provision on settling plaintiffs by providing procedural requirements in a claim reduction hearing and by imposing the burden of demonstrating the propriety and fairness of claim reduction on the nonsettling defendant requesting claim reduction. S. 380, § 4I(f)(1). The provision of hearing rights and due process standards is important in considering the constitutionality of retroactively applied legislation. L. TRIBE, *supra* note 166, § 10-1, at 476. However, in many situations, if not most, the effect of retroactive application will be to deprive antitrust plaintiffs of damages which they expected to receive and which provided the statutory incentive to litigate, as contemplated by the private right of action in section 4 of the Clayton Act. In the face of such a blatant infringement on the substantive rights accorded under preenactment law, the right to a hearing is small solace.

phase of the case may be successful if they demonstrate that their further involvement in the case was unanticipated and will be substantial. However, the claim reduction provisions will be unconstitutional where plaintiffs suffer carve outs from their damages, or where previously negotiated settlement agreements are substantially impaired by claim reduction principles. The rather intricate provisions in S. 380, particularly those creating a right to a hearing, cannot remedy the unconstitutionality of applying claim reduction to these situations.[182]

V. CONCLUSION

This Article considers the form that contribution and claim reduction legislation is taking and various implications of that form; it has not considered the propriety of a contribution rule.[183] The conclusion is that Congress, in considering any proposal for contribution and claim reduction, must give careful thought to the effect of the legislation on antitrust policy and practice, including fairness, deterrence, and transaction costs. These fundamental antitrust objectives and economic consequences are implicated in each of the pending proposals. When these considerations are evaluated in the context of each of the proposals, substantial flaws are discovered. The relative weights assigned and importance given to these values have varied throughout the judicial decisionmaking process. As Congress contemplates the pending legislation, it must resolve the conflicts between these goals. In balancing the values and competitive effects of the legislation, Congress can bring clarity to the law.

It is our firm conviction that if Congress adopts a position favoring contribution, then it should recognize distinctions in the severity of the proscribed activity and the actors' conduct in the conspiracy. The preferable approach of adopting a contribution rule should be recognized in cases involving rule of reason analysis, where there are no "bright line" demarcations

---

[182]*Cf.* G. BELL, *supra* note 48, at 23 (concluding that "a better solution would be to leave the proposed legislation silent with respect to its applicability").

[183]The authors have reached different conclusions concerning the merits of a contribution rule vis-a-vis a no-contribution rule. *See Antitrust Equal Enforcement Act of 1979: Hearings on S. 1468 Before the Subcomm. on Antitrust, Monopoly and Business Rights of the Senate Comm. on the Judiciary*, 96th Cong., 1st Sess. 142 (1979) (statement of Donald J. Polden); Sullivan, *supra* note 5.

between acceptable and proscribed conduct and where the anticompetitive nature of the challenged conduct is equivocal.[184]

Contribution rights should not be accorded in traditional per se offenses, such as horizontal price-fixing. Denying the right of contribution in such cases would best serve the deterrence objectives of the law. Little justification exists for allowing an antitrust violator to recover contribution from others where his conduct is anticompetitive. Similarly, the invocation of contribution should be within the trial court's discretion so that an economically influential violator cannot escape liability through coercion, collusion, or the plaintiff's caprice. The relative bargaining power of the wrongdoers and the extent of their participation in the illegal conduct are highly relevant in determining the propriety of contribution on a case-by-case basis. In these respects, the bills considered by the House are flawed, particularly those recognizing a right to contribution only in horizontal price-fixing cases.

Should Congress adopt a procontribution approach, it will need to resolve the method of apportioning liability and of determining the respective contribution shares. A pro rata basis, while clearly the easiest to administer, suffers from the same unfairness that results from a no-contribution rule, in that each violator would bear an equal share of the damages without regard to its relative participation in or benefits from the proscribed conduct. In this respect, deterrence also would be diluted, because the severity of the punishment would be apportioned equally. For the same reasons, a "market share" standard does not advance deterrence objectives, because the market share data are not necessarily related to complicity in the antitrust violation and are likely to raise problems of complexity.

Contribution on the basis of relative responsibility may be the most appropriate method of computing contribution rights and claim reduction, though it is perhaps the most complicated method proposed. Its value lies in elimination of the inequities caused by the no-contribution approach, while at the same time it recognizes distinctions in the severity of proscribed behavior and the relative culpability of each antitrust violator. In contrast to pro rata or market share standards, it may protect less culpable defendants with relatively slight economic power from dam-

---

[184]*See* Easterbrook, Landes & Posner, *supra* note 5, at 367–68; Sullivan, *supra* note 5, at 417, 421.

age exposure that does not correspond to the extent of participation in the illegal conduct. It also furthers the policy of the antitrust laws by predicating severity of liability on the extent of the responsibility for the proscribed activity. In doing so, it adds a measure of certainty and predictability to the application of a contribution rule. Although it may significantly affect litigation strategies, counseling methods, and settlement negotiation, a comparative contribution rule may best advance the policies underlying the law. The deterrent goals of certainty and severity of punishment will be advanced within the context of contribution, while some unfairness caused by the no-contribution tradition will be eliminated.

Before Congress adopts a position against the right to a jury trial on the contribution issue, it needs to consider the premerger custom, the remedy sought, and the ability of juries to rationally decide contribution issues. Historical research on these issues does not suggest a clear resolution, although it would seem to favor a discretionary approach by the trial court in weighing the relevant factors in each instance. The historical development of the Seventh Amendment right indicates that in determining contribution shares the right to a jury trial lacks historical precedent when the means used to reallocate the liability is based on a pro rata or market share determination. Stronger arguments are available for the invocation of jury trials when the damage distribution is based on relative degrees of fault.

Finally, the bills present real issues of constitutional adequacy with respect to attempts to make claim reduction principles applicable to pending antitrust cases. Current constitutional analysis strongly suggests that retrospective application of the claim reduction features will be unconstitutional under the Due Process Clause where settlements entered into before the effective date of the legislation are disrupted and where settlement carve outs reduce plaintiffs' settlement expectations, as conditioned by preenactment law. More important, however, is the notion that the primary objectives of the proposed legislation—enhanced deterrence and greater fairness and balance in compensatory goals—will be fairly and efficiently advanced by prospective application, and not by retrospective windfalls to defendants in a few pending cases.

The model legislation that follows addresses the concerns expressed in this Article. It attempts to advance the antitrust

objectives of deterrence, compensation, and fairness, in the context of traditional concepts of proof, and weighs those factors against transaction costs and complexity. It rejects a right of contribution in horizontal price-fixing cases, as envisioned under H.R. 4072, H.R. 1242, H.R. 5794, S. 995, and Discussion Drafts Nos. 1 and 2. Contribution, moreover, is not permitted under the proposal for antitrust conduct that is per se unlawful, as in classic cartel-like conduct, but rather only where the conduct is scrutinized under a rule of reason analysis. A further limitation places discretion in the trial court as to the invocation of the right to contribution. Finally, should the trial court invoke its discretion in the apportionment of damages, the model legislation would require contribution shares to be determined on a relative fault standard. In no event is the model legislation intended to alter the substantive antitrust law.

A BILL

To amend the Clayton Act to establish a right of contribution with respect to damages in certain actions brought under such Act.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

That the Clayton Act (15 U.S.C. §§ 12–27 (1980)) is amended by inserting after section 4H the following new section:

"SEC. 4I. (a) Any person who is liable for damages in an action brought under section 4, 4A, or 4C of this Act may claim contribution, in accordance with this section, from any other person jointly liable for such damages.

"(b) A claim for contribution may be asserted by crossclaim, counterclaim, or third-party claim in the same action as that in respect of which contribution rights are claimed, or in a separate action, whether an action has been brought or a judgment has been rendered against the person from whom contribution is sought.

"(c) A claim for contribution shall be forever barred unless filed within six months after the entry of the final judgment for which contribution is sought.

"(d) Contribution may not be claimed by or from a person who, pursuant to a settlement agreement entered into in good faith with a plaintiff in the action in respect of which contribution rights are claimed, has been released from liability or potential liability for the underlying claim.

"(e) Rights to contribution granted under subsection (a) shall be within the discretion of the trial court.

"(f) In any action under section 4, 4A, or 4C of this Act, the court shall reduce the claim of any person releasing any person from liability or potential liability for damages by the greatest of—

"(1) any amount stipulated for this purpose;

"(2) the amount of the consideration paid for the release; or

"(3) the contribution share of the person released.

"(g) (1) With respect to claims based upon horizontal per se violations, or price-fixing, contribution within this section shall not be permitted.

"(2) In determining contribution shares with respect to all other claims, the court shall consider the relative responsibility of each party for the origination or perpetration of and participation in the violation for which damages have been awarded and the benefits derived therefrom.

"(h) Nothing in this section should be construed to deny any person the right to a jury trial otherwise permitted for a claim asserted under subsection (a).

"(i) Nothing in this section shall affect the joint and several liability of any person.

"(j) This section shall apply only to actions under section 4, 4A, or 4C of this Act commenced after the date of enactment of this section.

"(k) The enactment of this section shall not invalidate any settlement entered before the date of this section's enactment, whether or not approved by the court."

STATUTE

REGULATING CATASTROPHES THROUGH FINANCIAL RESPONSIBILITY REQUIREMENTS: A MODEL STATE STATUTE

MICHAEL B. MEYER*

Activities producing catastrophic loss risk are defined as those human enterprises which are theoretically capable of producing some credible event which entails extremely large losses of human life, health or property. Two examples of catastrophic loss risk producing facilities, commercial nuclear power plants and LNG terminals, appear to produce a type of externality by imposing uncompensated loss risk costs on neighbors. Further, these facilities may be dependent upon the subsidies implicit in these externalities for their continued economic operation. A Model State Statute is proposed which would use insurance premiums as an unbiased source of probability and outcome estimates in order to eliminate this externality and the resulting subsidy, and to create increased economic efficiency.

One of the most studied and most troublesome problems in regulation is the problem of external diseconomies of production. External diseconomies may result when a cost of production is neither borne by the producer nor passed on to the consumer through price increases, but rather is transferred to the public at large or to some sub-class of the public. A classic example of an external diseconomy occurs when air or water pollutants are freely dumped, which causes the real costs attached to these pollutants to be borne by the users of the air or water. The traditional solution advocated by economists is to place a tax on the dumping roughly equivalent to the costs of the environmental degradation. The manufacturer pays the tax and eventually passes it on to the consumer of the product through price increases. A second example of an external dis-

---

*Senior Research Associate, Analysis and Inference, Inc. B.A., Mathematics, Harvard University, 1967; J.D., Boston College Law School, 1973.

William B. Fairley, Paul L. Chernick, and Garrick F. Cole provided many helpful comments on earlier drafts of this Article. All errors and oversights are, of course, the responsibility of the author. An earlier version of this Article was presented at a Workshop on Low Probability/High Consequence Events, sponsored by the Society for Risk Analysis, the U.S. Department of Energy, the U.S. Environmental Protection Agency, and the U.S. Nuclear Regulatory Commission, at the Hyatt Regency Hotel, Arlington, Va., June 1982, and will appear in the Conference Proceedings to be published in 1983 by Plenum Press.

economy is the increased public transit, water, sewer, school, and other municipal service costs that result from additional office or apartment building construction in congested areas.[1]

External diseconomies are troublesome because they represent a failure of the market mechanism. To the extent that a production cost can be shifted from the producer to the public at large, that cost is not reflected in the product's price. This results in overproduction and overconsumption of the product, as compared to the most efficient levels of production and consumption. This also imposes a tax on the public, or on some subclass of the public, who bears the real cost of the external diseconomy. External diseconomies are subsidies that run from society as a whole, or from some subclass of society, to the class of consumers of the product. These subsidies generally cause lower-than-optimal prices and higher-than-optimal levels of purchase and consumption of the product.[2]

A less-studied example of an external diseconomy results from the imposition of catastrophic loss risks upon the entire society or upon the neighbors of a facility potentially subject to a catastrophic accident. If such a facility is operated by a firm without sufficient assets or sufficient casualty liability insurance coverage, then the public or the neighbors may directly bear the costs of the loss risk. The risk should not be thought of as cost-free, no matter where it lies or how poorly understood the probability and the outcome of the catastrophic loss may be. If the risk is retained by the facility, the associated costs may be reflected by higher insurance premiums or by higher costs of capital and higher total capital requirements of the operator of the facility. If the risk is not retained by the facility, but rather is transferred to the neighbors, its costs may be reflected by the neighbors' lower property values, higher property insurance premiums, increased difficulty of sales of the property, or decline of the neighborhood.

Catastrophic loss risks and their external diseconomies are not merely a theoretical or academic problem. Liquefied natural

---

[1]There is a large body of economic literature on the problem of external diseconomies. *E.g.*, 1 A. KAHN, THE ECONOMICS OF REGULATION: PRINCIPLES AND INSTITUTIONS 194–95 (1970); E. MANSFIELD, MICROECONOMICS, THEORY AND APPLICATIONS 450–52 (1975); E. MISHAN, COST-BENEFIT ANALYSIS 101–80 (1971).

[2]"Optimal" is used here in the sense of Pareto optimal. Pareto optimality exists whenever no one can be made better off by a redistribution of goods without making others worse off. Conversely, if one person's lot can be improved by redistribution, without making others worse off, the situation before the redistribution was not Pareto optimal.

gas (LNG) storage and handling facilities and commercial nuclear power plants are two examples of facilities that are generally believed to produce catastrophic loss risks that are not fully assumed by the facility or by the facility's insurers. Without these external diseconomies, the economics of these activities would probably be substantially altered. Indeed, spokesmen for these industries have claimed that these subsidies are required for the existence of these two industries.[3]

This Article examines the problem of external diseconomies that result from the operation of commercial nuclear power plants and large LNG storage and handling facilities. Sections I and II outline the economic and legal implications of these external diseconomies. In Sections III and IV, a tentative conclusion is reached that catastrophic loss risks are improperly allocated in the United States today, which results in a wide array of undesirable effects. Finally, in Section V, a Model State Statute is presented that creates a regulatory scheme that should result in substantial improvements over the status quo, at least in areas that are not preempted by federal law.

I. Nuclear Power Plants

A. *Economic Issues in Catastrophic Loss Risk Sharing*

1. Probabilities and Outcomes of Various Credible Accidents

The most thorough current study of the probabilities and the outcomes of various accidents at commercial nuclear power plants is contained in the Nuclear Regulatory Commission's Reactor Safety Study (RSS).[4] Employing fault-tree and event-tree analysis[5] to estimate the probabilities of certain accidents at two types of nuclear power plants (a pressurized water reactor and a boiling water reactor), the RSS went on to estimate the likely consequences of these accidents. The RSS undoubtedly constituted a major advance in the understanding of accidents

---

[3]*See infra* note 36.

[4]NRC, Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants, WASH-1400, NUREG-75/014 (Oct. 1975) [hereinafter cited as RSS].

[5]Event trees are logical sequences, running forward in time, showing by branches the consequences which can result from any event. Fault trees are similar, but run backwards in time, showing what types of events can logically cause the failure being studied. Of course, for fault trees and event trees to disclose all paths to any particular failure, one needs perfect information about the system under examination.

at nuclear power plants; however, both the methodology and the data base employed by the RSS have been subject to severe criticism in the professional literature.[6] An official Nuclear Regulatory Commission (NRC) review determined that, although it was impossible to determine whether the core melt probabilities contained in the RSS were too high or too low, it appeared highly likely that the error bands, or confidence intervals, around the probabilities were understated and that true error bands would be wider.[7] Nevertheless, the RSS provides a starting point for nuclear power plant accident probability and outcome assessments.

The RSS produced estimates for the probabilities of accidents of varying severity, both for immediate and delayed consequences, which can generally be described as extremely small for the accidents of catastrophic proportion (see Tables I and II, reproduced below).[8] The RSS also produced similar esti-

---

[6] AD HOC RISK ASSESSMENT REVIEW GROUP, NRC, RISK ASSESSMENT REVIEW GROUP REPORT TO THE U.S. NUCLEAR REGULATORY COMM'N, NUREG/CR-0400 (Sept. 1978) [hereinafter cited as LEWIS REPORT]; Yellin, *The NRC's Reactor Safety Study*, 7 BELL J. ECON. 317–39 (1976); EPA, REACTOR SAFETY STUDY [WASH-1400], A REVIEW OF THE REPORT, EPA-520/3-76-009 (1976); UNION OF CONCERNED SCIENTISTS, THE RISKS OF NUCLEAR POWER REACTORS: A REVIEW OF THE NRC REACTOR SAFETY STUDY (Aug. 1977); *Report to the Am. Physical Soc'y by the Study Group on Light Water Reactor Safety*, 47 REV. MOD. PHYSICS, Supp. No. 1 (Sum. 1975). Subsequent estimates of accident probabilities indicate that actual reactor experience produces probability estimates which are substantially more pessimistic than those contained in the RSS. *See* J. MINARICK & C. KUKIELKA, PRECURSORS TO POTENTIAL SEVERE CORE DAMAGE ACCIDENTS: 1969–1979, A STATUS REPORT, NUREG/CR-2497, ORNL/NSIC-182/VI (NRC, June 1982); J. MINARICK & C. KUKIELKA, PRECURSORS TO POTENTIAL SEVERE CORE DAMAGE ACCIDENTS: 1969–1979 (1982).

[7] LEWIS REPORT, *supra* note 6, at vi.

[8]

Table 1 *Consequences of Reactor Accidents For Various Probabilities For One Reactor*

Chance Per Reactor-Year	*Early Fatalities*	*Early Illnesses*	*Total Property Damage (Billions $)*	*Decontamination Area (Sq. Miles)*	*Relocation Area (Sq. Miles)*
1/20,000 (5×10^{-5})	<1.0	<1.0	<0.1	<0.1	<0.1
1/1,000,000 (1×10^{-6})	<1.0	300	0.9	2,000	130
1/10,000,000 (1×10^{-7})	110	3,000	3	3,200	250
1/100,000,000 (1×10^{-8})	900	14,000	8	—	290
1/1,000,000,000 (1×10^{-9})	3,300	45,000	14	—	—

Source: RSS, *supra* note 4, at 83 (Table 5-4 in original).

mates for 100 reactors to demonstrate the synergistic effect of reactors in use across the country; these probability estimates were in each case merely 100 times higher. Thus, the 1/1,000,000,000 probability of an accident causing 3,500 early fatalities and $14 billion property damage for one reactor-year was interpreted as a 1/333,333 probability for one such accident over the thirty-year lives of 100 reactors.[9]

The RSS has been severely criticized by reviewers, who have asserted that the RSS probably has underestimated the probabilities of the various accidents, and certainly has overestimated the precision of its own results.[10] There is considerable doubt as to whether the RSS provides an adequate basis for public policy decisionmaking. The official NRC review of the RSS concluded that "the absolute values of the risks presented by the [RSS] should not be used uncritically either in the regulatory process or for public policy purposes."[11] In addition, some insurance industry sources view the RSS probability and outcome estimates as too optimistic.[12] Despite these problems, the RSS estimates will be used here as the basis for analysis of cata-

Table 2 *Consequences of Reactor Accidents for Various Probabilities For One Reactor*

Chance Per Reactor-Year	*Latent Cancer Fatalities (Per Year)**	*Thyroid Nodules (Per Year)**	*Genetic Effects (Per Year)***
1/20,000	<1.0	<1.0	<1.0
1/1,000,000	170	1,400	25
1/10,000,000	460	3,500	60
1/100,000,000	860	6,000	110
1/1,000,000,000	1,500	8,000	170
Normal Incidence	17,000	8,000	8,000

*This rate would occur approximately in the 10 to 40 year period following a potential accident.

**This rate would apply to the first generation born after a potential accident. Subsequent generations would experience effects at a lower rate.

Source: RSS, *supra* note 4, at 83 (Table 5-5 in original).

[9]RSS, *supra* note 4, at 84.

[10]*See supra* note 6 and accompanying text.

[11]LEWIS REPORT, *supra* note 6, at 3.

[12]H. Denenberg, Testimony Before the Atomic Safety and Licensing Board, Licensing Hearing for Three Mile Island Nuclear Power Plant (Nov. 7, 1973), *reprinted in Possible Modification or Extension of the Price-Anderson Insurance and Indemnity Act: Hearings before the Joint Comm. on Atomic Energy*, 94th Cong., 2d Sess. 226–40 (1974); Wood, *Nuclear Liability After Three Mile Island*, 48 J. RISK & INS. 450–64 (1981); *see also* P. CHERNICK, W. FAIRLEY, M. MEYER & L. SCHARFF, DESIGN, COSTS, AND ACCEPTABILITY OF AN ELECTRIC UTILITY SELF-INSURANCE POOL FOR ASSURING THE ADEQUACY OF FUNDS FOR NUCLEAR POWER PLANT DECOMMISSIONING EXPENSE 63–84, NUREG/CR-2370 (NRC, Dec. 1981) [hereinafter cited as P. CHERNICK].

strophic loss risk sharing, on the dual grounds that no alternative estimates have been accepted by the scientific community and that, by using estimates of probabilities that some observers believe are more likely to be too low than too high, an element of conservatism will be introduced into the following discussion.

Thus, estimates based on the RSS indicate that credible accidents from nuclear power plants could result in property damage of about $14 billion, in about 3,500 early fatalities, in approximately 1,500 latent fatalities per year for about a thirty-year period, and in contamination of about 3,200 square miles. The total fatalities, [3,500 + (1,500 × 30) =] 48,500, might be equivalent (in dollar terms) to economic losses of about $50 billion.[13] Total damages might thus be on the order of magnitude of $60 billion. (This statement means no more than that we can conclude that total damages would be above $6 billion and below $600 billion.) The probability of such an accident, according to the RSS, is about 1/1,000,000,000 per reactor-year, which equates to about 1/333,333 for 100 reactors for thirty years.

Other estimates have produced substantially larger probabilities and substantially more severe outcomes.[14] Of course, the actual outcome of any such accident would depend on variables such as the severity of the accident and the population density downwind of the plant at the time of the accident.

An extreme example of the destructiveness of a hypothetical accident at a nuclear power plant is provided by the Indian Point site. Indian Point is located in Buchanan, New York, about thirty-five miles from mid-town Manhattan. Estimates of a "worst case" accident at Indian Point, which would entail both a substantial radiation release at the site and a steady wind to the south and east, could contaminate substantial portions of New York City and Long Island. This could cause total deaths in the 10,000 to 100,000 range, property damage in the $50 billion range, and health care costs well over $1 billion.[15]

---

[13]This assumes a valuation of a human fatality of $1 million. Settlements in the $0.5 million to $3 million range are relatively common in wrongful death actions, and regulatory actions imply roughly similar valuations of human lives. *See* Cohen, *Society's Valuation of Life Saving in Radiation Protection and Other Contexts*, 38 HEALTH PHYSICS 33–51 (1981); Graham & Vaupel, *Value of a Life: What Difference Does It Make*, 1 RISK ANALYSIS 89–95 (1981).

[14]*See, e.g.*, UNION OF CONCERNED SCIENTISTS, *supra* note 6, at 113–30; Wood, *supra* note 12, at 450–64.

[15]D. STRIP, ESTIMATES OF THE FINANCIAL CONSEQUENCES OF NUCLEAR POWER REACTOR ACCIDENTS 4–11, NUREG/CR-2723, SAND82-1110 (Sandia Nat'l Labs. 1982); L. RITCHIE, J. JOHNSON & R. BLOND, CALCULATIONS OF REACTOR ACCIDENT CON-

2. Financial Resources for Accident Compensation

Total compensation for both third-party property damage and personal injury resulting from an accident at a nuclear power plant is limited by the Price-Anderson Act to $560 million.[16] This limitation on total liability has been held constitutional by the Supreme Court.[17] Under the Act, a facility is required to obtain whatever private insurance is available, and to contribute $5 million to a fund, with the federal government providing the remainder of the total fund of $560 million available for third-party compensation. The Price-Anderson Act has been extended through 1987 for its third ten-year period. No other resources will be available unless Congress chooses to provide additional funds or to require additional contributions after a catastrophic accident has occurred.[18] The Price-Anderson Act has thus partially vitiated the ability of liability insurance premiums to internalize externalities.[19]

B. *Legal Considerations*

State regulation of radiation limits for public health and safety purposes has been clearly preempted by federal statute. In *Northern States Power Co. v. Minnesota*,[20] the Eighth Circuit held that a state nuclear waste law was preempted by the Atomic Energy Act, and that the federal government has exclusive authority to regulate the construction and operation of nuclear power plants.[21]

State regulation imposing financial requirements on nuclear power plants also appears to be preempted by the Price-Anderson Act if motivated by safety concerns.[22] Thus, state regulation

---

SEQUENCES, VERSION 2, USER'S GUIDE, NUREG/CR-2326, SAND81-1994 (Sandia Nat'l Labs. 1982). *See generally* RSS, *supra* note 4, at 83–84.

[16]42 U.S.C. § 2210 (1976). For a legislative history and statement of the legislative purpose of the Price-Anderson Act, see 1957 U.S. CODE CONG. & AD. NEWS 1803.

[17]Duke Power Co. v. Carolina Envtl. Study Group, Inc., 438 U.S. 59 (1978), *rev'g* Carolina Envtl. Study Group, Inc. v. AEC, 431 F. Supp. 203 (W.D.N.C. 1977).

[18]Congress has reserved the right to review any incident in which damages exceed $560 million and "take whatever action is deemed necessary and appropriate." 42 U.S.C. § 2210(e) (1976).

[19]Ralston, *Pollution Liability and Insurance: An Application of Economic Theory*, 46 J. RISK & INS. 497, 505 (1979).

[20]447 F.2d 1143 (8th Cir. 1971), *aff'd per curiam*, 405 U.S. 1035 (1972).

[21]*Id.* at 1154.

[22]*See* Duke Power Co. v. Carolina Envtl. Study Group, Inc., 438 U.S. 59, 82–94 (1978) (upholding constitutionality of the Price-Anderson Act with discussion of congres-

of liability insurance coverage, and other similar state-imposed financial responsibility requirements probably could not be undertaken without federal authorization.

The Supreme Court recently has affirmed the existence of a degree of state control over nuclear power plants in conjunction with "their traditional responsibility in the field of regulating electrical utilities for determining questions of need, reliability, cost, and other related state concerns."[23] In *Pacific Gas & Elec. v. Energy Resources Comm'n*, Justice White wrote that the federal role does not preempt state economic regulation of utilities. Thus, state legislation imposing financial responsibility requirements would not be preempted if it is part of the state's regulation of the financial feasibility of construction of a nuclear power plant.

II. LNG Storage, Transportation, and Transshipment Facilities

A. *Economic Issues in Catastrophic Loss Risk Sharing*

1. Probabilities and Outcomes of Various Credible Accidents

The shipment and storage of liquefied natural gas is also a possible source of catastrophic loss risks of a severity comparable to nuclear power plants. As much uncertainty exists concerning the probabilities and the outcomes of various credible LNG accidents as for commercial nuclear power plants. The most widely cited summary of LNG accident probabilities and outcomes in the United States is the 1978 study prepared by the General Accounting Office (GAO).[24] In contrast to the RSS, the GAO Report has been criticized for being too alarmist in tone and for overstating risks.[25]

Four separate phases of LNG operations in the United States are important for assessment of catastrophic risk: import ship-

---

sional authority over, and interest in, area of liability insurance coverage, and implying that area is preempted, although no specific ruling made on preemption).

[23]Pacific Gas & Elec. v. Energy Resources Comm'n, No. 81–1945, slip op. at (Apr. 20, 1983). [Available in 43 S. Ct. Bull. (CCH) ¶B2076 (1983).]

[24]GAO, Report to the Congress of the United States, Liquefied Energy Gases Safety, EMD-78-28 (July 31, 1978) [hereinafter cited as GAO Report].

[25]*See id.* at vol. III (containing comments on the GAO Report from other federal agencies).

ping via marine tankers, marine transport via barges, land transport via trucks and rail cars, and storage and transshipment facilities. There do not appear to be acceptable estimates that are in agreement as to the proper order of magnitude for probabilities of accidents involving land transport via truck and rail cars or for accidents involving storage facilities.[26] Better estimates of probabilities for accidents involving tanker and barge operations, however, have been developed in regulatory proceedings and in the professional literature (see Table III below).[27]

Questions have been raised concerning the assumptions, data, and methods of calculation used to generate these probability estimates. As with the RSS, there appears to be substantial reason to think that these estimates have wide uncertainty bands

---

[26]J. Fay, Risks on LNG and LPG (1979) (unpublished manuscript) (land transportation of LNG and liquefied propane gas (LPG) by rail and truck may comprise the largest contribution to total societal risk of LNG and LPG operations).

[27]Such estimates have to be increased for number of trips per year and for number of years in order to calculate the estimated probabilities for a period of time in one harbor due to one ongoing operation.

Table 3 *Estimates of Odds of LNG Spills*

Type of Ship	*Source of Estimate*	*Odds of Spill Per Trip*
tanker	A.D. Little*	2.665×10^{-7}
tanker	Dr. T.W. Horner**	2.144×10^{-7}
tanker	A.D. Little***	1.45×10^{-7}
barge	Dr. T.W. Horner****	1.8×10^{-6}
barge	A.D. Little*****	1.1×10^{-6} to 2.6×10^{-7}

*Testimony of A.A. Brown of A.D. Little Co. in *In re* Distrigas Corp., F.P.C. Docket No. CP73-78, at 1.

**2 BUREAU OF NATURAL GAS, FPC, FINAL ENVIRONMENTAL IMPACT STATEMENT FOR THE CONSTRUCTION AND OPERATION OF AN LNG IMPORT TERMINAL AT STATEN ISLAND, NEW YORK (Docket Nos. CP73-47, *et al.*), Attachment 2, at 4 (1974) (Attachment 2 consists of T. Horner, Probability Assessment of LNG Ship Accidents in the New York and Providence Harbors, A Report to the FPC).

****Id.* (citing D. Allan, Probabilities of Collision and Damage Affecting the General Dynamics 124,000M³ LNG Ship (Arthur D. Little Co. Report, 1974)).

*****Id.* at Attachment 3, at 4 (Attachment 3 consists of T. Horner, Probability Assessment of LNG accidents on New York Barge Transits from Staten Island to Newtown Creek and Steinway Creek, A Report to the FPC).

******Arthur D. Little Co., Analysis of Probability Collisions, Rammings, and Groundings of the LNG Barge *Massachusetts*, at C-2 (Oct. 1974). Estimates vary depending upon barge route taken.

around them, and some reason to think that they are too optimistic.[28]

LNG spills could conceivably result in large conflagrations and explosions. Whether these events occur after a spill depends upon the rate of spill, the rate of evaporation, the wind speed and direction for moving and dispersing the cloud, and the method of ignition or detonation of the cloud. It is difficult to attach appropriate probabilities to the different outcomes. Credible accidents of unknown but small probabilities could occur. For example, the Distrigas LNG tanks in Everett, Massachusetts theoretically could produce a spill that, under certain weather and ignition conditions, could destroy downtown Boston two miles away. A tanker accident several miles off the coast theoretically could produce a spill that, under certain weather and ignition conditions, could cause extensive onshore damage.

One major source of uncertainty stems from disagreement by experts over how far an evaporated cloud of gas/air mixture could spread and still remain flammable. The longer a gas/air cloud can hold together while being transported, the greater the area at risk. A methane/air mixture will burn if the methane comprises between 5.3% and 15% of the total, while an ethane/air mixture will burn if the ethane comprises between 3.0% and 12.5% of the total.[29] To illustrate the practical importance of the disagreement over how rapidly gas/air clouds dissipate as they move, widely varying estimates have been made for the distance that a gas/air cloud could travel and still remain flammable (see Table IV below).[30]

[28]Fairley, *Evaluating the "Small" Probability of a Catastrophic Accident from the Marine Transportation of Liquefied Natural Gas*, in STATISTICS AND PUBLIC POLICY 331–53 (W. Fairley and F. Mosteller eds. 1977); Fairley, *Assessment for Catastrophic Risks*, 1 RISK ANALYSIS 197–204 (1981).

[29]1 GAO REPORT, *supra* note 24, at 2-4.

[30]

Table 4 *Estimates of Distances Gas/Air Clouds Can Travel and Still Remain Flammable*

Source of Estimate	*Miles*
Federal Power Commission	0.75
American Petroleum Institute	5.2
Cabot Corporation	11.5
U.S. Coast Guard	16.3
Prof. James Fay	17.4
U.S. Bureau of Mines	25.2 to 50.3

Sources: J. HAVENS, PREDICTABILITY OF LNG VAPOR DISPERSION FROM CATASTROPHIC SPILLS INTO WATER: AN ASSESSMENT (U.S. Coast Guard, Apr. 1977); 1 GAO REPORT, *supra* note 24, at 11-12.

Obviously, because these estimated distances vary greatly, and the "correct" answer must depend upon variable factors which are unknown before the accident, substantial uncertainty necessarily exists in the assessment of these potential risks. Substantial uncertainty also exists over whether such gas/air mixtures can detonate or explode, and over the chances of a fire ignited on the downwind edge of such a cloud burning back to the source of the leak or spill.

The most accurate summary of the probabilities and consequences of various LNG accidents is that very severe outcomes are theoretically possible, which would involve billions or tens of billions of dollars of damages. While the probabilities of such catastrophic accidents are obviously small, they warrant our attention as a potential source of diseconomies.

2. Financial Resources for Accident Compensation

Most large LNG operations present the possibility of accidents with resultant damages in excess of insurance coverage or available corporate assets. For example, tankers that carry LNG usually are owned by corporations specially created for the sole purpose of operating that tanker; often such a corporation has no other assets. Additionally, federal admiralty law usually limits the liability of the vessel's owner for a maritime accident to the salvage value of the vessel and its cargo.[31] In the case of a severe LNG accident involving a tanker, positive salvage value might well be close to zero.

Similarly, large LNG storage and unloading facilities typically are owned by corporations that have no other substantial assets and that are subsidiaries of the parent corporations involved in the LNG import, storage, and sale business. For example, the LNG storage tanks in Everett, Massachusetts are approximately two miles from downtown Boston and are adjacent to heavily populated inner-city residential neighborhoods. These tanks are owned and operated by Distrigas of Massachusetts Corporation, which has no other substantial assets beyond the tanks, piers, and transshipment facilities. Distrigas of Massachusetts Corporation is separate from the importer, Distrigas Corporation; both are subsidiaries of Eastern Energy Corporation, which is in turn a subsidiary of the Cabot Corporation. Thus, the relatively substantial assets of the Cabot Corporation and of the

---

[31] 46 U.S.C. §§ 181–189 (1976); *see* 3A W. BENEDICT, BENEDICT ON ADMIRALTY § 1 (7th ed. 1980).

Distrigas Corporation would not be available to pay for damages resulting from a large accident suffered by Distrigas of Massachusetts Corporation at the Everett facilities.[32] Of course, insurance coverage may ensure financial responsibility substantially beyond the specific corporate assets. This is difficult to assess, however, because insurance coverage is generally not public information.[33]

Similarly, LNG shipment by truck and rail tank car probably entails accident consequences far more severe than available corporate assets and insurance coverage. The fragmented and varied situation in this area is not well documented and does not lend itself to generalization.

B. *Legal Considerations*

The nature of LNG use and shipment lends itself to state regulation. The clear preemption of state regulation of commercial nuclear power radiation accident protection is not present in the LNG area.[34] States have sufficient power under the police power doctrine to require certain amounts of corporate assets or levels of insurance coverage for many intrastate, or partially intrastate, operations.[35] The extent to which certain LNG operations involve shipments in interstate commerce, and hence are subject to federal regulation, is dependent upon the specific facts of each operation. Nevertheless, all LNG facilities and operations in the United States remain subject to some level of state control. This control should be used to ensure the existence of certain corporate assets or levels of insurance coverage.

III. Policy Implications

Despite the legitimate dispute over the probabilities and the consequences of various possible accidents, no doubt exists that

[32]*See* Lucas v. Mobil Oil Corp., 331 F. Supp. 957 (N.D. Tex. 1971). For a discussion of the financial responsibility of a parent for its subsidiary, see 13A W. Fletcher, Cyclopedia of the Law of Private Corporations § 6222 (perm. ed. 1961).

[33]One of the few sources discussing the corporate structure of companies that transport and store LNG and of the inadequacy of their insurance coverage is 1 GAO Report, *supra* note 24, at 11-1 to 11-38.

[34]The operation of admiralty law is largely or completely beyond the control of state legislatures. *See* Moragne v. States Marine Lines, 398 U.S. 375 (1970); Chelentis v. Luckenback S.S. Co., 247 U.S. 372 (1918); Southern Pac. Co. v. Jensen, 244 U.S. 205 (1917). Similarly, some aspects of interstate, and import and coastal shipping are within the federal domain.

[35]*See supra* note 23.

commercial nuclear power plants and LNG facilities impose catastrophic loss risks of some magnitude upon their neighbors. Some credible accidents would not be adequately covered by insurance or capital. Additionally, the magnitude of these transferred catastrophic loss risks may be sufficient to affect the economics of the industries involved if they are transferred in whole back to the industry.

Government and industry representatives in both cases have indicated that if the costs of these catastrophic loss risks were placed upon the facility generating the risks, the industry in question might not exist. The force and breadth of these assertions made by supporters of the industries indicate the magnitude of the problem. In testimony supporting passage of the Price-Anderson Act, representatives of publicly-owned utilities, investor-owned utilities, reactor manufacturers, government regulators, and electric utility industry associations testified that passage of the Price-Anderson Act was a necessary prerequisite to the manufacture, sale, and operation of commercial nuclear power reactors.[36] Both the district court and the Supreme Court found that these assertions were crucial in obtaining the passage of the Price-Anderson Act.[37] Similarly, both courts found that the Price-Anderson Act was instrumental in, and indeed that it was a prerequisite to, the existence and growth of the commercial nuclear power industry.[38]

The LNG industry also owes its existence, at least partially, to the limitations on liability presently governing LNG operations. For example, the United States Department of Commerce, in making formal comments upon the GAO Report, stated categorically that:

> The conclusions and recommendations with respect to statements requiring that the companies involved in supplying LNG and LPG [liquefied propane gas] be liable to the full extent of the assets of their parent corporations would be totally unacceptable to any corporate board of directors. If the liability conditions suggested by the GAO were to be imposed on the industry, the effect would surely be the total withdrawal and disinvolvement of private enterprise with

---

[36]*Hearings on Government Indemnity for Private Licensees and AEC Contractors Against Reactor Hazards Before the Joint Comm. on Atomic Energy*, 84th Cong., 2d Sess. 5 (1956).

[37]Carolina Envtl. Study Group, Inc. v. AEC, 431 F. Supp. 203, 215–18 (W.D.N.C. 1977), *rev'd on other grounds sub. nom.*, Duke Power Co. v. Carolina Envtl. Study Group, Inc., 438 U.S. 59, 63–67, 75–78 (1978).

[38]431 F. Supp. at 215–18; 438 U.S. at 63–67, 75–78.

> LEG [liquefied energy gases] projects or any project that has unlimited risk.[39]

The Commerce Department went on to indicate, however, that the consumer of LNG should bear some of the costs of LNG accident liability insurance, and that the federal government should bear the remainder of the expense. No recommendation was made that neighbors should continue to bear part of the risk.[40]

These pessimistic opinions that the commercial nuclear power industry and the LNG industry would not survive if forced to be financially responsible for catastrophic accidents appear to be contradicted by the very low probabilities for accidents cited by these two industries in licensing cases. If the accident probability and consequence estimates of the RSS or the LNG transport studies are believable, then the logical conclusion is that liability insurance premiums for such coverage should be quite modest. Pessimistic industry opinions may reflect underlying industry skepticism concerning these optimistic probability and consequence estimates.

There is no obvious economic or policy justification for requiring neighbors to absorb part of a facility's catastrophic loss risk costs. This market failure can be corrected by making the facility bear the costs and passing them on to the consumer through price increases. If a legislature finds that operation of these facilities is in the public interest, and that they could not exist without some subsidies for these catastrophic loss risk costs, then the taxpayers should bear the burden of the subsidies. No justification exists for the present implicit legislative finding that subsidies should be created, but that they should be supported by the facility's neighbors rather than by taxpayers in general.

The existence of these industries suggests that there is a substantial public interest in the continued safe operation of commercial nuclear power plants and LNG facilities. There is also a substantial public interest in correcting the market failures that result from not passing on to the consumer the costs of catastrophic loss risks, and of not creating safety incentives of the proper magnitude for the facility owners and operators. The crucial issue is to provide economically efficient assessment and

---

[39]3 GAO REPORT, *supra* note 24, at 9.
[40]*Id.* at 9–10.

allocation of these catastrophic loss risk costs by ensuring that adequate, but not excessive financial protection exists to compensate potential victims of such an accident.

IV. Market Risk Assessments: Strengths and Weaknesses

Market risk assessments can provide a reasonably independent and unbiased source of both probabilities and outcomes of catastrophic loss risks.[41] The idea of obtaining risk assessments from independent parties, such as insurance companies or self-insurance pools, is attractive because an insurer has a substantial financial stake in the accuracy of both probability and consequence estimates. Although such providers of market risk assessments would not necessarily have any better data or any more accurate methodologies than other risk assessors, their financial stake in accurate estimates should provide reasonably independent and unbiased estimates, if insurance markets enjoy effective competition and are not limited by price collusion or other anti-competitive problems.

Two possible roles could be fulfilled by market risk assessments in regulating catastrophic loss risk costs:

1. market risk assessments can be used directly, through insurance or other risk-of-loss shifting mechanisms, to transfer risk costs to the appropriate parties; or
2. market risk assessments can be used indirectly to estimate costs in ratesetting contexts, or as sources of probablility and outcome estimates in health and safety regulation, whether or not transfer of risk costs occurs.[42]

The Model State Statute adopts a regulatory scheme to further these roles by forcing facilities that produce catastrophic loss risks to transfer these costs to appropriate parties and, where

---

[41] W. Fairley, *Market Risk Assessment of Catastrophic Risks*, in The Risk Analysis Controversy: An Institutional Perspective (H. Kunreuther and E. Ley eds. 1982); R. Field, Statement Before the Committee on the Institutional Means for Assessment of Risks to Public Health of the National Research Council (Feb. 10, 1982) (on file at Harv. J. on Legis.); Wood, *supra* note 12; Denenberg, *supra* note 12. Several investigators have made the point that government specification of risk reduction measures has reduced incentives to develop more efficient methods of risk control. Advisory Committee on Reactor Safeguards, An Approach to Quantitative Safety Goals for Nuclear Power Plants 32, NUREG-0739 (1980); Starr and Whipple, *Coping with Nuclear Power Risks: The Electric Utility Incentives*, 23 Nuclear Safety 1, 1–7 (Jan.–Feb. 1982).

[42] I am indebted to Paul L. Chernick for this observation.

such a transfer of costs is impossible, by generating independent cost estimates or probability and consequence estimates. In addition, the proposed regulatory scheme is intended to improve market sensitivity to safety issues through the use of the insurance mechanism. The Statute will improve the liability system through clarification of the costs of safety, or will facilitate creation of mechanisms for coping with risks that cannot be borne by the risks' creator.[43]

A. *Strengths*

Market risk assessments and the financial responsibility requirements set forth in the Model State Statute have two attractive features. First, in areas where catastrophic loss risk cost subsidies are neither proper nor necessary, market risk assessments provide a direct, simple, and unbiased source of information on the optimal costs of safety and on the optimal level of production of safety-enhancing devices. In addition, they provide a simple method for internalizing the catastrophic loss risk costs. Second, in areas in which catastrophic loss risk cost subsidies are proper and necessary, market risk assessments can provide relatively unbiased information to measure the externalities imposed by the activity. These two advantages will be discussed briefly in this Section.

First, a regulatory scheme that forces producers of catastrophic loss risks to obtain market risk assessments for the costs of these risks, and that forces internalization of the resulting externalities through financial responsibility requirements, is more efficient and dependable than traditional health and safety regulation. "Regulatory capture" by special interest groups representing the regulated industry is not a concern when regulation functions through market risk assessments. Unlike traditional health and safety regulation, regulators need not face the extraordinarily difficult choice of determining how safe is safe enough. Decisions concerning optimal total levels of safety or the relative cost-effectiveness of comparative investments in safety are irrelevant in a market risk assessment regulatory scheme. By mandating the use of insurance and insurance-based

---

[43]B. FISCHHOFF, S. LICHTENSTEIN, P. SLOVIC, S. DERBY & R. KEENEY, ACCEPTABLE RISK 150–53 (1981).

estimates, the proposed statutory scheme would internalize loss risk costs and would increase the likelihood that loss risk estimates were objective and unbiased. Regulated industries no longer would be able to argue for licensing purposes that they are so safe as to make accident risks trivially small, while also arguing for insurance indemnification purposes that insurance costs would be so great as to cripple the industry. The same estimate, whether high or low, would be used for both purposes, and would be the one on which insurers would be willing to rely. Regulators would be partially released from their current dependency on engineering and epidemiological risk estimates, with their attendant uncertainty and result-orientation. Most importantly, regulators would be able to make public policy decisions based on risk assessments generated by truly independent and financially responsible competing sources.

Second, even where subsidization of the product's consumers by the neighbors of the facility is deemed necessary and appropriate, market risk assessments can achieve the more limited goal of informing regulators of the dimensions of the problem. Market risk assessments can supply regulators with independent sources of probability and consequence estimates even when the actual financial responsibility requirements of the statute would not be applied. Regulators presumably would be better informed concerning their actions if they had available a "catastrophic loss risk cost subsidy budget" (analogous to a tax-expenditure budget). Such information would reveal the existence and the magnitude of subsidies created by government action.

B. *Weaknesses*

Despite the obvious attractiveness of a regulatory scheme that uses market risk assessments as a relatively unbiased source of probability and consequence estimates and as a vehicle for internalizing externalities, at least five limitations deserve recognition before any such regulatory scheme is implemented. The five limitations discussed in this Section are: (1) market risk assessments can only expose the subsidy for examination, and cannot regulate the activity, where subsidization is necessary or desirable; (2) for the market risk assessments to be unbiased, adequate competition must exist in the insurance industry;

(3) the insurance industry should have approximately the same discount rate and degree of risk aversion as society; (4) the proposed scheme can operate at the state level only in areas that are not preempted by the federal government; and (5) the proposed scheme can operate only insofar as it is justified by the police power of the states and does not burden interstate commerce.

First, public policy may make some subsidization unavoidable and necessary. For example, it often is argued that commercial nuclear power provides a valuable hedge against several non-economic costs of excessive dependence upon imported oil. If this is true, then commercial nuclear power should be subsidized as if there were a "shadow" price on imported oil, because the actual value of additional nuclear generation might exceed the direct costs of the displaced oil generation.[44] Taxpayer subsidies are preferable, but if they are politically impossible, it may be necessary, though inequitable, to extract this subsidy in part from the facility's neighbors.

Second, the insurance industry (or competitor industries, such as the self-insurance pool industry) must be competitive if their market risk assessments are to be unbiased. This is purely an empirical economic question, and the current evidence seems to justify the finding that effective competition does indeed exist.[45] This limitation must be kept in mind for any particular application of the suggested regulatory scheme.

---

[44]National defense costs may rise due to America's perceived excess dependence upon imported oil from the Persian Gulf and other unstable areas. If this is the case, the actual (cash) price per barrel of imported oil can be thought of as being below the "true" total price of the imported oil; the difference between the "true" (unknown) total price and the actual (cash) price is often referred to as the "shadow" price on imported oil.

[45]Evidence that the property-liability insurance industry enjoys workable competition is summarized in Joskow, *Cartels, Competition and Regulation in the Property-Liability Insurance Industry*, 4 BELL J. ECON. 118–40 (1973); NATIONAL COMM'N FOR REVIEW OF ANTITRUST LAWS & PROCEDURES, REPORT TO THE PRESIDENT AND ATTORNEY GENERAL (1979); U.S. DEP'T OF JUSTICE, THE PRICING AND MARKETING OF INSURANCE (1977); STAFF OF THE NAT'L ASS'N OF INS. COMM'RS, MONITORING COMPETITION: A MEANS OF REGULATING THE PROPERTY AND LIABILITY INSURANCE BUSINESS (1974). Workable competition probably exists even in such high-variance outcome markets as the nuclear insurance market; this can be seen by the creation of competing self-insurance pools for first-party property damage coverage (Nuclear Mutual Limited and Nuclear Electric Insurance Limited) to compete with the insurance industry's pool (American Nuclear Insurers/Mutual Atomic Energy Reinsurance Pool). *See* P. CHERNICK, W. FAIRLEY, M. MEYER & L. SCHARFF, *supra* note 12, at 49–58, 67–82; J. LONG, NUCLEAR PROPERTY INSURANCE: STATUS AND OUTLOOK 19, NUREG-0891 (NRC, 1982).

Third, the insurance industry (and its competitor industries) must have roughly similar discount rates and risk aversion curves as society before privately generated premium estimates can be assumed to approximate publicly acceptable estimates of costs. Discount rates are the percentages by which future economic events are discounted to make them comparable to present economic events; risk aversion curves measure how much various parties prefer certainty to risk given certain economic risks and benefits. Fortunately, it is again an empirical economic fact that these discount rates and risk aversion curves seem to be close enough to one another to permit the insurance premiums to convey useful information.[46] There are extreme situations, however, where these considerations might become crucial. For example, there is sensitivity even to small differences in the relevant discount rate when considering time periods far in the future. In addition, for particularly extreme situations of physical risk, moderate differences in degrees of risk aversion would become important. Attention would have to be devoted to this possibility in administering the proposed regulatory scheme.

Fourth, some areas of potential state regulation, such as those involving radiation and radioactive effluent from commercial power plants, are preempted by federal statutory and administrative regulation. The application of the proposed regulatory scheme still would be possible in this situation. For nuclear power, the proposed statute would have to be enacted at the federal level or at the state level pursuant to a specific federal statutory grant of authority to the states.

---

[46]The economic literature on the subjects of risk aversion and public and private discount rates is voluminous; a few highly selective references to that literature will be provided here. For a discussion of property-liability insurance industry costs of capital, and hence discount rates, see Fairley, *Investment Income and Profit Margins in Property-Liability Insurance: Theory and Empirical Results*, 10 BELL J. ECON. 192–210 (1979); Hill, *Profit Regulation in Property-Liability Insurance*, 10 BELL J. ECON. 172–91 (1979). On discounting and discount rates for public and private projects, see P. DASGUPTA, A. SEN & S. MARGLIN, GUIDELINES FOR PROJECT EVALUATION 154–200 (1972); E. MISHAN, *supra* note 1, at 181–267. On the subject of risk aversion generally, and risk aversion functions, see J. HENDERSON & R. QUANDT, MICROECONOMIC THEORY: A MATHEMATICAL APPROACH 52–60 (3d ed. 1980); R. KEENEY & H. RAIFFA, DECISIONS WITH MULTIPLE OBJECTIVES: PREFERENCES AND VALUE TRADEOFFS 148–88 (1976); and H. RAIFFA, DECISION ANALYSIS: INTRODUCTORY LECTURES ON CHOICES UNDER UNCERTAINTY 90–97 (1968). On the subject of utility or preference functions, and utility theory and game theory in general, see R. LUCE & H. RAIFFA, GAMES AND DECISIONS: INTRODUCTION AND CRITICAL SURVEY 12–38 (1957); J. VON NEUMANN & O. MORGERNSTERN, THEORY OF GAMES AND ECONOMIC BEHAVIOR 15–45 (3d ed. 1953).

Fifth, the proposed regulatory scheme could operate at the state level only insofar as it was founded upon the state police power doctrine and did not burden interstate commerce. No obstacle should be presented by the Commerce Clause as long as there is no discrimination in favor of intrastate activities[47] and the state action is reasonably related to the nexus of activities taking place within the state.[48] The proposed statute is designed to meet these criteria. Thus, the Commerce Clause should not constrain the proposed scheme.

V. REMEDIES: A MODEL STATE STATUTE

Some states have considered legislation that would regulate loss risk allocation resulting from specific hazardous operations. These efforts have been too restrictive in their scope. For example, the Massachusetts legislature in 1979 considered a bill that would require minimal standards of financial responsibility for LNG and LPG facilities and operations.[49] The Massachusetts legislation, however, attempted to tackle the problem with reference to only two potentially hazardous materials or industrial processes; a long list of other potentially hazardous materials (such as gasoline, naptha, chlorine, benzene, herbicides, and insecticides) and industrial processes (such as large refineries, nuclear power plants, and petrochemical processing plants) present equally substantial catastrophic loss risks.

A superior legislative approach would consider an appropriate method of loss risk allocation. Some preliminary work has been done towards evaluating the relative merits of different methods of loss allocation, such as liability insurance or compensation funds.[50] State attempts to regulate in this area, however, are virtually nonexistent.

A proposed uniform state law would be premature at present,

---

[47]Dean Milk Co. v. Madison, 340 U.S. 349, 354 (1951).

[48]Pike v. Bruce Church, 397 U.S. 137, 142 (1970).

[49]S. 511, Mass. Senate, 1979 Sess. (entitled "Rules and Regulations Concerning the Safety and Liability Requirements for LNG and LPG").

[50]*See* J. FERREIRA & S. HILL, MECHANISMS FOR SHARING THE COSTS OF LARGE ACCIDENTS (Energy Impacts Project, Lab. of Architecture & Planning, M.I.T., Sept. 1980); J. FERREIRA, *Promoting Safety through Insurance*, in SOCIAL REGULATION: STRATEGIES FOR REFORM 267–88 (E. Bardach and R. Kagan eds. 1982).

due to the wide variety of problems faced by different states and the value of encouraging experimentation among states. A Model State Statute is presented here, however, to provide a starting point for discussion and consideration of the specific problems faced by different states. No attempt is made to limit the coverage of the statute by type of operation. Rather, the coverage of the statute is defined by the estimated probabilities and consequences that could result from the credible accidents. Substantial latitude is left to the state for exercise of administrative discretion by the agency responsible for administering and enforcing the statute.

The Model State Statute first determines whether a facility or activity is capable of causing a credible catastrophic accident. If this threshold test is satisfied, the owner or operator then has the burden of presenting a plan that allocates the resulting catastrophic loss risk costs to proper parties. The design of the plan, and the choice of reliance upon casualty insurance, self-insurance pools, loss channeling agreements, compensation pools, or corporate assets, is left to the owner or operator. The plan is to be approved if substantially all of the catastrophic loss risk costs are shifted away from the neighbors of the facility or activity. If the plan is not approved within a specified time period, the facility or activity is not permitted to continue operations. An exception is permitted in cases in which the cost of implementing the plan would make continued operation of the facility or activity impossible. In these cases, the entire plan still must be produced and its costs estimated, but the plan itself may be left only partially implemented.

Experience with traditional regulation has been unsatisfactory. Limited agency resources and expertise, expansive regulatory responsibility and authority, vague statutory authorizations to regulatory agencies, result-oriented risk estimates specifically designed for use in adversary proceedings, and the different agendas of the regulated industries and the regulators all have contributed to the failure of the traditional regulatory scheme. Economic theory and some empirical economic facts indicate that health and safety regulation based upon market risk assessments and resulting financial responsibility requirements would probably outperform the traditional regulatory scheme. The proposed Model State Statute is offered as a contribution to a timely resolution of this serious problem.

A MODEL STATE STATUTE FOR THE REGULATION OF THE ALLOCATION OF CATASTROPHIC LOSS RISK COSTS

Statement of Legislative Findings and Purpose

The legislature hereby finds that some facilities and activities currently existing and conducted in this state produce catastrophic loss risks. The legislature hereby further finds that the costs of these catastrophic loss risks should not be borne primarily or even substantially by the neighbors of the facilities or activities, but rather should, if possible, be borne primarily or completely by the owner or operator of the facility or activity. The legislature hereby further finds that facilities or activities that produce catastrophic loss risks confer substantial economic benefits upon the citizens of the state, and that wherever possible, protection against the catastrophic loss risks produced should be imposed in such a way as not to affect the economic viability of the facility or activity.

Section 1. Creation of the State Catastrophic Loss Risk Regulatory Commission

(a) There is hereby created a State Catastrophic Loss Risk Regulatory Commission. The Commission shall be headed by a Director. The Director shall be appointed by the Governor, and shall serve coterminously with the Governor. The Director, upon appointment, shall continue in his term of office unless removed for cause.

(b) The salary of the Director shall be 85% of the salary fixed by law for the Governor of the state.

(c) The annual budget of the Commission shall be thirty (30) times the salary of the Director.

(d) The Director shall appoint such professional and clerical staff as he shall see fit; providing, however, that the Director shall appoint at least the following:

(1) Three economist/statisticians as follows:

(a) one economist/statistician with a Ph.D. in economics, statistics, or mathematics, with a salary of 99% of the Director's salary;

(b) two economist/statisticians, with M.A. or M.S. degrees in economics, statistics, or mathematics, each with salaries of at least 90% of the Director's salary.

(2) Two engineers, as follows:

(a) one engineer, licensed to practice as a professional engineer in the state, with a masters degree in engineering, at a salary of 99% of the Director's salary.

(b) one engineer, licensed to practice as a professional engineer in the state, and with a bachelors degree in engineering, at a salary of at least 90% of the Director's salary.

(3) Three lawyers, as follows:

(a) one lawyer, licensed to practice law in the state, and with a J.D. or an LL.B. degree, at a salary of 99% of the Director's salary.

(b) two lawyers, licensed to practice law in the state, with J.D. or LL.B. degrees, each at a salary of at least 90% of the Director's salary.

(e) All staff appointed by the Director shall serve at the Director's pleasure, including those staff appointed to positions specified in Section 1(d).

Section 2. DEFINITIONS

(a) "Catastrophic Loss" shall mean a loss of more than 250 human lives, or a loss of more than $500,000,000 (in 1980 dollars) according to the following formula:
combined damages = [($2,000,000 in 1980 dollars)(X)] + [Y]
where X = number of human deaths and
Y = amount of property damage, in 1980 dollars.

(b) "Catastrophic Loss Risk Producing Facility or Activity" shall mean a facility or activity which the Director has found to be capable of having a credible accident which would result in a catastrophic loss, pursuant to Section 3 of this chapter.

(c) "Commission" shall mean the State Catastrophic Loss Risk Regulatory Commission.

(d) "Credible Accident" shall mean either:

(i) an accident which could happen, and which has sufficient probability or likelihood of occurring so as to cause reasonable and prudent business decisions to include consideration of such an accident; or

(ii) an accident which could happen which has a probability or likelihood of occurring of greater than 1 in 1,000,000,000.

(e) "Director" shall mean the Director of the State Catastrophic Loss Risk Regulatory Commission.

(f) "Plan" shall mean a Catastrophic Loss Risk Allocation Plan, as further described in Section 4 of this chapter.

Section 3. DESIGNATION OF FACILITY OR ACTIVITY AS A CATASTROPHIC LOSS RISK PRODUCING FACILITY OR ACTIVITY

(a) The Director shall conduct an adjudicatory hearing to determine whether any facility or activity which appears to the Director to be

a catastrophic loss risk producing facility or activity could produce a credible accident which would result in a catastrophic loss; this preliminary determination by the Director that a facility or activity appears to be a catastrophic loss risk producing facility shall not require any preliminary findings of fact or law by the Director or any preliminary hearing or any other administrative action. The Director may hold such an adjudicatory hearing upon his own motion, and the Director shall hold such an adjudicatory hearing upon motion of the state attorney general, the state public advocate, the commissioner of public health, the commissioner of public safety, the commissioner of insurance, the chairman of the public service commission, or the chairman of the energy facilities siting council, or upon the motion of the mayor, city manager, or board of selectmen of the city or town in which any such facility or activity exists, or upon the motion of any group of 100 citizens.

(b) An adjudicatory hearing held pursuant to this Section shall not be held upon less than thirty (30) days notice. Notice shall be served upon the owner or operator of the facility or activity in question, and upon the state attorney general, the state public advocate, the commissioner of public health, the commissioner of public safety, the commissioner of insurance, the chairman of the public service commission, and the chairman of the energy facilities siting council, and upon the mayor, city manager, and board of selectmen of the city or town in which the facility or activity in question is located.

(c) The staff of the State Catastrophic Loss Risk Regulatory Commission shall appear in any adjudicatory proceeding held pursuant to this Section. The staff of the State Catastrophic Loss Risk Regulatory Commission shall take a position as to whether the facility or activity is a catastrophic loss risk producing facility. The Director shall permit, upon motion, any of the following to intervene as parties to any adjudicatory proceeding held pursuant to this Section: the owner or operator of the facility or activity in question, the state attorney general, the state public advocate, the commissioner of public health, the commissioner of public safety, the commissioner of insurance, the chairman of the public service commission, the chairman of the energy facilities siting council, and the mayor, city manager or board of selectmen of the city or town in which the facility or activity in question is located. The Director may permit any other party, who makes a showing that it is substantially affected by the outcome of the proceeding and that it is capable of making a substantial and meritorious contribution to the record, to intervene as a party in any adjudicatory proceeding held pursuant to this Section.

(d) The Director shall issue his final decision within six (6) months of the date of the first day of hearings in any adjudicatory proceeding held pursuant to this Section. Such decision shall include a specific determination of whether the facility or activity in question is a catastrophic loss risk producing facility, which in turn shall be based

upon a determination of whether the facility or activity in question could produce a credible accident which could result in a catastrophic loss.

Section 4. Procedure Upon Designation of a Facility or Activity as a Catastrophic Loss Risk Producing Facility or Activity

(a) Upon a finding by the Director that a facility or activity is a catastrophic loss risk producing facility or activity pursuant to Section 3 of this chapter, the owner or operator of such facility or activity shall file with the Director within sixty (60) days of the finding a catastrophic loss risk allocation plan. This plan shall provide for allocating no less than 90% of the total catastrophic loss risk to the owner or operator of the facility or activity. This plan may additionally reallocate the catastrophic loss risks, in whole or in part, from the owner or operator of the facility or activity, by one or more of the following methods:

(1) property and casualty liability insurance policies from insurers, groups of insurers, or self-insurance pools;
(2) compensation funds or pools;
(3) risk shifting or channeling agreements among various owners or operators;
(4) assumption of responsibility by governmental agencies; or
(5) corporate assets.

(b) The Director shall hold an adjudicatory hearing upon the adequacy of the plan submitted pursuant to subsection (a) of this Section within thirty (30) days of the submission of the plan. All parties to the original hearing pursuant to Section 3 of this chapter shall be parties to such hearing. The Director shall conclude such hearing within thirty (30) days of the date of the first hearing day, and shall issue a decision as to whether or not the plan is adequate within thirty (30) days of the completion of the hearing. The Director shall approve the plan if the plan provides reasonable assurance that no less than 90% of the total catastrophic losses from any credible accident will be borne by the owner or operator or by some third party (other than the neighbors of the facility or activity) pursuant to subsection (a) of this Section. If the Director finds that the plan provides the reasonable assurance required by the preceding sentence of this subsection, but that the costs of implementing the plan are so great as to make the facility or activity economically nonviable, the Director may approve a modified plan which provides reasonable assurance that some percentage smaller than 90%, but in no case less than 50%, of the total catastrophic losses from any credible accident will be borne by the owner or operator or by some third party (other than the neighbors of the facility or activity) pursuant to subsection (a) of this Section.

(c) The owner or operator of the facility or activity shall place a plan approved by the Director pursuant to subsection (b) of this Section in full operation within 120 days of the Director's approval.

(d) The owner or operator of the facility or activity shall modify a plan disapproved by the Director pursuant to subsection (b) of this Section to comply in all respects with the changes in the plan specified by the Director as adequate to cause subsequent approval and shall place such modified plan in full operation within 120 days of the Director's decision disapproving the original plan.

(e) The Director shall order any facility or activity which does not have a fully approved catastrophic loss risk allocation plan in full operation within nine (9) months after a finding by the Director pursuant to Section 3 of this chapter that such facility or activity is a catastrophic loss risk producing facility or activity, to cease operation forthwith. Such order of the Director shall be enforceable upon petition by any party to the Section 3 proceeding in [here insert name of the appropriate state court].

(f) It shall be a criminal offense, punishable by a fine of up to $25,000 per day of ownership or operation, to own or operate a facility or activity more than nine (9) months after a finding by the Director pursuant to Section 3 of this chapter that such facility or activity is a catastrophic loss risk producing facility without having a fully approved catastrophic loss risk allocation plan in full operation.

(g) The Director shall without limitation maintain as public records, and may additionally disseminate to the public as he determines necessary, any explicit or implicit probability estimates and outcome estimates provided to the Commission pursuant to proceedings held pursuant to this Section or held pursuant to Section 3.

Section 5. PROCEDURE UPON DETERMINATION THAT A FACILITY OR ACTIVITY IS NOT A CATASTROPHIC LOSS RISK PRODUCING FACILITY OR ACTIVITY

Upon a determination by the Director that a facility or activity is not a catastrophic loss risk producing facility or activity after a hearing pursuant to Section 3 of this chapter, such facility or activity shall not be the subject of any additional hearing pursuant to Section 3 of this chapter within four years of such determination; provided, however, that such a facility or activity may be the subject of such a hearing within four years of such a determination if the Director first determines that the nature of the facility or activity has undergone substantial and relevant change since the time of the previous determination or that substantial additional evidence exists, which did not exist at the time of the previous determination, which is likely to affect the outcome of the proceeding. The burden of demonstrating that the nature of the facility or activity has undergone substantial and relevant change since the time of the previous determination, or that substantial additional evidence exists, which did not exist at the time of the previous determination,

which is likely to affect the outcome of the proceeding, is upon the party requesting that a hearing be held pursuant to Section 3 of this chapter.

Section 6. Regulations

The Director is authorized to promulgate permanent regulations which he deems to be necessary and convenient for the proper and efficient administration of all programs and duties undertaken pursuant to this chapter. The Director may promulgate such permanent regulations only after notice and hearing pursuant to the procedural requirements of [here insert statutory reference to state statute governing the issuance of regulations after a hearing and upon a record]. Permanent regulations promulgated by the Director shall be made upon the record of the hearing conducted by the Director pursuant to this Section, and no permanent regulations shall be promulgated by informal agency action. The Director may promulgate emergency regulations, without notice or hearing, which may remain in force for no more than sixty (60) days, if the Director makes a determination that such regulations are necessary to protect public health and safety and that delay in the issuance of such regulations would result in substantial risk to public health and safety. [If this last sentence contravenes the relevant state Administrative Procedure Act provision, replace this last sentence with the relevant standard from the state Administrative Procedure Act.]

Section 7. Judicial Review

(a) Any final order, ruling, or decision of the Director may be appealed by any party to the proceeding before the Director (if the proceeding was an adjudicatory proceeding) or by any party adversely affected by the final order, ruling, or decision (if the proceeding was a rulemaking proceeding) by filing a complaint in the nature of a petition for judicial review with the [here insert name of the appropriate state court]. Such complaint in the nature of a petition for judicial review shall be filed within twenty (20) days of the date of service of the final order, ruling, or decision on the appealing party (if the proceeding was an adjudicatory proceeding) or shall be filed at any time (if the proceeding was a rulemaking proceeding).

(b) The record on appeal of any final order, ruling, or decision of the Director may consist of either:

  (i) all written or documentary exhibits admitted into evidence at the hearing, all written transcripts of oral testimony at the hearings, the final order, ruling, or decision being appealed, and any other additional material which all parties to the appeal agree to incorporate; or

  (ii) an agreed statement of facts if all parties to the appeal certify that the agreed statement of facts recites all relevant factual matters necessary to a decision on the appeal and that the agreed statement of facts conforms to the truth.

(c) The [here insert name of the appropriate state court] shall uphold the final order, ruling, or decision of the Director in any adjudicatory proceeding unless the court finds that such final order, ruling, or decision was not made upon substantial evidence, was arbitrary and capricious, was an abuse of discretion, was made upon illegal procedure, or was otherwise erroneous as a matter of law. The [here insert name of appropriate state court] shall not hear any evidence, nor shall it make any *de novo* review of the factual record, nor shall it make any *de novo* findings of fact.

(d) The [here insert name of the appropriate state court] shall uphold the final order, ruling or decision of the Director in any rulemaking proceeding unless the court finds that such final order, ruling, or decision was made upon illegal procedure, exceeded the statutory authority of the Director, or was not based upon the record before the Director.

Section 8. SEVERABILITY

If any section, subsection, paragraph, sentence, phrase, clause, group of words, or word of this chapter is determined to be unconstitutional by any court of competent jurisdiction, such determination shall not be construed to effect the remainder of this chapter.

Section 9. SHORT TITLE

This chapter shall be referred to as the Catastrophic Loss Regulatory Act.

ARTICLE

BEDTIME FOR *BIVENS*: SUBSTITUTING THE UNITED STATES AS DEFENDANT IN CONSTITUTIONAL TORT SUITS

THOMAS J. MADDEN*
NICHOLAS W. ALLARD**
DAVID H. REMES***

During the past decade, Congress has struggled with the problem of sovereign immunity for constitutional torts committed by employees of the federal government. The Ninety-eighth Congress has proved to be no exception. Three bills, H.R. 595, S. 775, and S. 829, have been introduced during the current Congress to modify the immunity available to the federal government in the event of constitutional torts committed by its officials. For the most part, the current system has remained unchanged since the Supreme Court ruled in Bivens v. Six Unknown Named Agents *that victims of constitutional violations by federal officials may recover damages against wrongdoing officials under the Constitution.*

The authors, critical of the Bivens *remedy, find that those bills currently before Congress that attempt to rectify* Bivens *do not adequately reconcile the seemingly incompatible goals of providing adequate compensation for the tort victim and of ensuring the accountability of the wrongdoing official. According to the authors, it is not enough for the legislation to provide only for the substitution of the United States as the exclusive defendant in actions against its wrongdoing officials. The legislation should preclude the United States from asserting the qualified immunity available to the wrongdoing official, and should establish an alternative system of administrative sanctions against the individual official.*

For over ten years, Congress has considered, but failed to adopt, amendments to the Federal Tort Claims Act (FTCA)[1]

---

*Kaye, Scholer, Fierman, Hays & Handler; President, Federal Bar Association (D.C. Chapter); Adjunct Professor, American University; General Counsel, Law Enforcement Assistance Administration 1972–1979. J.D., Catholic University, 1968.

**Kaye, Scholer, Fierman, Hays & Handler. B.A., Princeton University, 1974; B.A., Oxford University, 1976; J.D., Yale Law School, 1979.

***Kaye, Scholer, Fierman, Hays & Handler. B.A., Columbia University, 1976; J.D., Harvard Law School, 1979.

The authors acknowledge the indispensable efforts and talents of Thaddeus J. Plante.

This Article follows an extensive study by the authors on the subject of executive-branch official liability and immunity at the federal, state, and local levels of government prepared for the Administrative Conference of the United States. T. MADDEN & N. ALLARD, ADVICE ON OFFICIAL LIABILITY AND IMMUNITY (1983); *see* Federal Officials' Liability for Constitutional Violations, 47 Fed. Reg. 58,207 (1982) (to be codified at 1 C.F.R. § 305.82-6). The analyses, comments, and views expressed in this Article are those of the authors and do not necessarily represent the policies, positions, or recommendations of the Administrative Conference.

[1]28 U.S.C. §§ 1346(b), 2671–2680 (1976 & Supp. V 1981). The FTCA provides that the United States generally will be liable for its torts in the same manner as a private

that would provide a damage remedy against the United States for constitutional injuries caused by federal officials.[2] These proposed amendments would have provided for the substitution of the federal government as the defendant in suits that can now be brought against only individual officials. The *Harvard Journal on Legislation*'s last look at this legislative activity was a 1979 article by Attorney General Griffin Bell.[3] Now, with the Ninety-eighth Congress in session, the long-running controversy over liabilities and immunities for constitutional torts caused by federal officials is again on the national agenda.

This Article analyzes the issues that Congress faces as it considers whether to authorize a new waiver of sovereign im-

---

individual at common law. The United States may be held liable for such torts only to the extent that it has waived its immunity and has consented to being sued.

[2]*See* S. 829, 98th Cong., 1st Sess. (1983); S. 775, 98th Cong., 1st Sess. (1983); H.R. 595, 98th Cong., 1st Sess. (1983); H.R. 7034, 97th Cong., 2d Sess. (1982); H.R. 6359, 97th Cong., 2d Sess. (1982); S. 1775, 97th Cong., 1st Sess. (1981); H.R. 1696, 97th Cong., 1st Sess. (1981); H.R. 24, 97th Cong., 1st Sess. (1981); S. 695, 96th Cong., 1st Sess. (1979); H.R. 2659, 96th Cong., 1st Sess. (1979); H.R. 193, 96th Cong., 1st Sess. (1979); S. 3314, 95th Cong., 2d Sess. (1978); S. 2868, 95th Cong., 2d Sess. (1978); S. 2117, 95th Cong., 1st Sess. (1977); H.R. 9437, 95th Cong., 1st Sess. (1977); H.R. 9219, 95th Cong., 1st Sess. (1977); H.R. 12,715, 93d Cong., 2d Sess. (1974); S. 2558, 93d Cong., 1st Sess. (1973); H.R. 10,439, 93d Cong., 1st Sess. (1973). These bills, with the exception of the pending S. 775 and S. 829, are collected in T. MADDEN & N. ALLARD, ADVICE ON OFFICIAL LIABILITY AND IMMUNITY (1983). S. 829, submitted to Congress on March 16, 1983 by President Reagan, is the administration's omnibus crime legislation. Title XIII of the bill would make the United States liable for constitutional torts committed by its employees, and would establish the FTCA as the exclusive means or remedy by which the actions of any federal official may be remedied in tort damages.

In 1974, Congress amended the FTCA to permit suits against the government for assault, battery, false arrest, abuse of process, and malicious prosecution by federal investigative or law enforcement officers. Act of Mar. 16, 1974, Pub. L. No. 93-253, § 2, 88 Stat. 50, 50 (codified at 28 U.S.C. § 2680(h) (1976 & Supp. V 1981)); *cf. Amendment of the Federal Tort Claims Act: Hearings Before the Subcomm. on Administrative Law and Governmental Relations of the House Comm. on the Judiciary*, 96th Cong., 1st Sess. 123 (1979) (statement of George Bermann, Professor of Law, Columbia Univ. Law School) (criticizing the statutory distinction between constitutional and nonconstitutional torts as unsound and unnecessary) [hereinafter cited as *1979 House Hearings*]. The 93d Congress, having enacted Pub. L. No. 93-253, failed to act on alternative legislation proposed by the Nixon administration that would have broadened the FTCA to cover claims of constitutional torts committed by government employees. *Federal Tort Claims Amendments: Hearings on H.R. 10439 Before the Subcomm. on Claims and Governmental Relations of the House Comm. on the Judiciary*, 93d Cong., 2d Sess. 14–15 (1974) (testimony of Irving Jaffe, Acting Ass't Att'y Gen., Civil Div., Dep't of Justice); *see also Intelligence Activities and the Rights of Americans: Final Report of the Senate Select Comm. to Study Governmental Operations with Respect to Intelligence Activities*, S. REP. No. 755, 94th Cong., 2d Sess. 336–38 (1976) (recommending a statutorily created federal cause of action against the government and individual federal intelligence officer or agent for actual or threatened injury caused by a violation of the Constitution committed under color of law).

As used in this Article, the term "official" encompasses officials, employees, and other executive-branch personnel of the federal government.

[3]Bell, *Proposed Amendments to the Federal Tort Claims Act*, 16 HARV. J. ON LEGIS. 1 (1979).

munity under the FTCA. Section I of the Article discusses the problems that recurrently trouble Congress in its attempts to pass legislation that would mitigate the tension between official liability and sovereign immunity. This background discussion is followed by a description in Section II of the principal bills in the House and the Senate: H.R. 595[4] and S. 775.[5] Both were introduced early in the Ninety-eighth Congress and are nearly identical to legislation that was pending at the close of the Ninety-seventh Congress.

Section III reviews in detail the numerous issues raised by H.R. 595 and S. 775. This Section is comprised of four parts. The first part discusses the retention of the "good-faith immunity" of the wrongdoing official. This is the most controversial issue raised by the two bills. The argument is presented that there is no justification for allowing the federal government to assert an official's qualified immunity or his reasonable good-faith belief in the lawfulness of his action. This argument is grounded both on the belief that the policy basis for good-faith immunity disappears when the official himself is no longer subject to suit and on the belief that allowing the United States to assert such immunity would undercut two major objectives of the legislation: compensation for the victim and accountability of the tortfeasor.

The establishment of an independent administrative disciplinary board is the subject of the second part of this Section. The establishment of such a board would supplement the current disciplinary system for official misconduct, thereby providing an effective deterrent against wrongful conduct by federal officials. In the third part of this Section, it is argued that victims of constitutional torts committed by government officials should be afforded the opportunity to a jury trial and to punitive or additional damages. While there is concern that juries may be overly sympathetic to plaintiffs and consequently may award excessive damages to victims, this problem could be averted by

---

[4]H.R. 595, 98th Cong., 1st Sess. (1983), was introduced by Rep. Sam B. Hall (D-Tex.), 129 CONG. REC. H94 (daily ed. Jan. 6, 1983). H.R. 595 is identical to H.R. 7034, 97th Cong., 1st Sess. (1981), the bill that was pending in the House Judiciary Committee at the conclusion of the 97th Congress, having been approved and forwarded by the Subcommittee on Adminstrative Law and Governmental Relations.

[5]S. 775, 98th Cong., 1st Sess. (1983), was introduced by Sen. Charles E. Grassley (R-Iowa), 129 CONG. REC. S2694-97 (daily ed. Mar. 11, 1983). S. 775 is pending in the Senate Committee on the Judiciary. The bill is nearly identical to legislation introduced during the last Congress by Senator Grassley that was reported to the Senate Judiciary Committee after hearings by the Subcommittee on Agency Administration.

setting a ceiling on the amount of punitive damages available for plaintiffs. The last part of Section III briefly reviews the problems of attorney fees, the legislative alternatives pertaining to the liability of former officials, the breadth of liability for constitutional torts committed by federal officials, and the ramifications of the distinction between constitutional and nonconstitutional torts. These matters are found to be critical to legislation dealing with immunity for officials who commit constitutional torts; it is urged that they be carefully considered by Congress.

The conclusion is reached in Section IV that the FTCA should be amended to substitute the United States as the exclusive defendant in all private damage actions for constitutional violations committed by officials while acting within the scope of their office. This legislative amendment should not permit the United States the opportunity to assert defensively the good-faith immunity of the wrongdoing official. In order to ensure the accountability of federal officials and to provide for the deterrence of unlawful and improper conduct, each federal agency should be specifically authorized to use existing administrative controls on officials who have violated the constitutional rights of any person. Moreover, Congress should evaluate the feasibility of establishing an independent disciplinary mechanism to supplement current administrative mechanisms. This would ensure that appropriate disciplinary action would be unfettered by an agency's interest in protecting its employees or its policies, and also would help to shield employees from unfairly being held responsible for agency misconduct.

I. BACKGROUND

There is wide dissatisfaction with the growing reliance on private damage actions against government officials for violations of constitutional rights.[6] The availability of such actions was proclaimed by the Supreme Court twelve years ago in *Bivens v. Six Unknown Named Agents*,[7] which "established that

[6] *See, e.g., infra* notes 14–17.

[7] 403 U.S. 388 (1971). In *Bivens*, the petitioner alleged that agents of the Federal Bureau of Narcotics made a warrantless entry and search of his apartment without probable cause. The Supreme Court held that petitioner had a federal cause of action under the Fourth Amendment against the federal officials who allegedly violated his constitutional rights. Although no statutory cause of action for such misconduct was

the victims of a constitutional violation by a federal agent have a right to recover damages against the official in federal court" directly under the Constitution.[8] Under the doctrine of sovereign immunity, victims of official misconduct cannot sue the government itself for damages absent a congressional waiver of sovereign immunity.[9] Congress has waived such immunity with respect to most common-law torts, but not with respect to violations of constitutional rights such as those alleged in *Bivens*-type actions.[10] Presently, there are more than one thousand pending lawsuits alleging constitutional violations.[11] The increasing reliance on such suits is attributable largely to court decisions that recognize a broad variety of claims available under the *Bivens* doctrine,[12] and to the uncertainty enveloping the availability and scope of official immunity.[13]

The dissatisfaction with the system inaugurated by *Bivens* stems, in part, from the recognition that private damage actions against individual government officials do not afford victims of official misconduct "a financially responsible defendant."[14] Of

---

available, the Supreme Court found that a cause of action could be asserted directly under the Constitution.

[8]Carlson v. Green, 446 U.S. 14, 18 (1980) (describing the holding of *Bivens*).

[9]*See, e.g.*, United States v. Testan, 424 U.S. 392, 399 (1976); P. BATOR, P. MISHKIN, D. SHAPIRO & H. WECHSLER, HART & WECHSLER'S THE FEDERAL COURTS AND THE FEDERAL SYSTEM 1339–56 (2d ed. 1973); *cf.* Note, *Rethinking Sovereign Immunity After* Bivens, 57 N.Y.U. L. REV. 597 (1982). Notwithstanding the persuasive case for a judicially created system of governmental liability, *id.*, only legislation can effect the comprehensive changes that such a system would entail.

[10]28 U.S.C. §§ 1346(b), 2671–2680 (1976 & Supp. V 1981). Another major congressional waiver of sovereign immunity is the Tucker Act, 28 U.S.C. §§ 1346, 1491 (1976 & Supp. V 1981) (creating court of claims for suits against the government for money damages).

[11]*Tort Claims: Hearings on H.R. 24, H.R. 3060, and H.R. 3795 Before the Subcomm. on Administrative Law and Governmental Relations of the House Comm. on the Judiciary*, 97th Cong., 1st & 2d Sess. 23 (1982) (statement of Deputy Att'y Gen. Edward Schmults) [hereinafter cited as *1982 House Hearings*]; *see* Bell, *supra* note 3, at 2 n.5.

[12]Indeed, the Supreme Court has indicated that a federal damage remedy for violations of the Constitution is the general rule. Carlson v. Green, 446 U.S. 14, 18 (1980); *see also* Briggs v. Goodwin, No. 80-2269, slip op. at 17–20 (D.C. Cir. Jan. 11, 1983).

[13]Harlow v. Fitzgerald, 102 S. Ct. 2727 (1982), is the Supreme Court's most recent word on the subject. In this case, the plaintiff brought a civil suit against two presidential aides for damages based on his allegedly unlawful discharge from work at the Department of the Air Force. The defendants claimed that they were entitled to absolute immunity. The Supreme Court held, inter alia, that a presidential aide generally is entitled only to qualified immunity unless he can demonstrate that the functions of his office are sufficiently sensitive so as to require absolute immunity. In addition, he must show that he was performing the protected duty when taking the action from which liability allegedly stems. *See generally* Cass, *Damage Suits Against Public Officers*, 129 U. PA. L. REV. 1110 (1981); Schuck, *Suing Our Servants: The Court, Congress, and the Liability of Public Officials for Damages*, 1980 SUP. CT. REV. 281.

[14]*Federal Tort Claims Act: Hearings Before the Subcomm. on Administrative Law and Governmental Relations of the House Comm. on the Judiciary*, 95th Cong., 2d

equal concern to many is the perception that the threat of personal liability or, at a minimum, the risk of a law suit, has a chilling effect and therefore deters not only improper conduct by government officials, but proper conduct as well.[15] In addition to these concerns is the conviction that the government *ought* to be liable for the wrongs committed in its name, rather than impose the loss upon an individual official "who is more anvil than hammer."[16] Other criticisms of the present system focus on the expense to the government of retaining private counsel to defend individual government officials in *Bivens* suits,[17] and the possibility that the government will have interests at stake that will not be represented adequately when individual officials rather than the government are on trial. The United States Supreme Court has indicated in *Carlson v. Green* that Congress could displace the *Bivens* remedy by "provid[ing] an alternative remedy when it explicitly declared the remedy to be a *substitute* for recovery under the Constitution and viewed as equally effective."[18] A mere ipse dixit to this

---

Sess. 4 (1978) (statement of Att'y Gen. Griffin Bell) [hereinafter cited as *1978 House Hearings*].

[15]*See 1982 House Hearings*, *supra* note 11, at 27 (statement of Deputy Att'y Gen. Edward Schmults); *1978 House Hearings*, *supra* note 14, at 4 (statement of Att'y Gen. Griffin Bell).

[16]Schuck, *supra* note 13, at 347, *reprinted in Federal Tort Claims Act, pt. 2: Hearing on S. 1775 Before the Subcomm. on Agency Administration of the Senate Comm. on the Judiciary*, 97th Cong., 2d Sess. 35, 101 (1982) [hereinafter cited as *1982 Senate Hearing*].

[17]*See, e.g., 1979 House Hearings*, *supra* note 2, at 3 (statement of Deputy Att'y Gen. Benjamin Civiletti); *Amendments to the Federal Tort Claims Act: S. 2117, pt. 1: Joint Hearing on S. 2117 Before the Subcomm. on Citizens and Shareholders Rights and Remedies and the Subcomm. on Administrative Practice and Procedure of the Senate Comm. on the Judiciary*, 95th Cong., 2d Sess. 7 (1978) (statement of Att'y Gen. Griffin Bell) [hereinafter cited as *1978 Senate Hearing*]. It has been estimated that the Justice Department has spent over three million dollars for such private counsel. 129 CONG. REC. S2697 (daily ed. Mar. 11, 1983) (statement of Sen. Thurmond). *But see 1979 House Hearings*, *supra* note 2, at 12–13 (Dep't of Justice Memorandum from Ass't Att'y Gen. Patricia Wald to Ass't Att'y Gen. Kevin Rooney, May 7, 1979) (discussing cost estimate for the proposed amendments to FTCA).

Insofar as the substitution of the United States as the defendant in place of the wrongdoing government official would afford plaintiffs "a financially responsible defendant," it is far from obvious that adopting an enterprise liability system would save the government any money at all, except to the extent that legislation establishing such liability would permit the government to settle rather than to pay to litigate cases involving alleged official misconduct. *See 1982 House Hearings*, *supra* note 11, at 248 (statement of Royce Lamberth, Chief, Civil Div., Office of the U.S. Att'y for the Dist. of Columbia).

Thus, it appears that the major gain for the government would be to ensure its complete control over the defense of constitutional tort actions. *See id.* at 237–39 (statement of U.S. Att'y Stanley Harris); *1979 House Hearings*, *supra* note 2, at 3–4 (statement of Deputy Att'y Gen. Benjamin Civiletti).

[18]446 U.S. 14, 18–19 (1980).

effect by Congress presumably would not suffice to meet the standards suggested in *Carlson*. A statutory civil sanction system would have to guarantee that persons whose constitutional rights have been violated by officials are adequately compensated, and also would have to perform the deterrent and corrective functions now intended to be served by the *Bivens*-type damage remedy.[19]

Most observers believe, and we agree, that to most effectively serve the primary goals of compensation, deterrence, and fairness in dealing with alleged constitutional violations by federal officials, and to afford a solution to the problems perceived to flow from the current system of individual liability, Congress should replace liability of individual officials with governmental liability for constitutional wrongs done in the public's name.[20]

Legislative efforts to establish such a system so far have foundered on disagreements over the advisability of allowing the United States to assert the qualified immunity of its officials in constitutional tort suits, and over the appropriate substitute mechanisms of deterrence and accountability. Failure to reach agreement on these issues reflects the fact that ensuring compensation and deterrence, and fostering vigorous decisionmaking by federal officials, are often perceived to be incompatible goals.[21] Although bills to establish some form of governmental

---

[19]*See id.* at 20–23; *see also* Cutts v. Fowler, 692 F.2d 138, 140 (D.C. Cir. 1982) ("creation of a new statutory remedy does not necessarily displace judicially-created remedies for constitutional deprivations").

[20]Generally, direct government liability is considered preferable to indirect liability through insurance and indemnification programs for officials found liable to a private party. In contrast to the direct liability proposals before Congress, neither insurance nor indemnification would relieve the defendant official of the need to defend lawsuits. There are other problems as well. The adoption of an insurance program could involve the added costs of dealing with the insurer, costs that would not exist in a system of direct liability. Indemnity is troubled by potential conflicts of interest between the government and the defendant official.

Federal officials apparently have difficulty purchasing liability insurance on their own. *See 1978 House Hearings*, *supra* note 14, at 31 (statement of John McNerney, Nat'l President, Fed. Criminal Investigators Ass'n). There is no general statutory indemnification of officials, even where the liability clearly arose within the scope of the official's employment. *See Federal Tort Claims Act, pt. 1: Hearings on S. 1775 Before the Subcomm. on Agency Administration of the Senate Comm. on the Judiciary*, 97th Cong., 1st Sess. 4 (1981) (statement of Deputy Att'y Gen. Edward Schmults) [hereinafter cited as *1981 Senate Hearings*]; Schuck, *supra* note 13, at 373.

[21]*See 1978 Senate Hearing*, *supra* note 17, at 8 (statement of Att'y Gen. Griffin Bell). The current administration has taken the position that "[t]he problem here is not really Government employees acting in an unlawful fashion [but rather] the overhang of litigation and the burden on Government employees of having . . . to be worried about their own cases." *1981 Senate Hearings*, *supra* note 20, at 7–8 (statement of Deputy Att'y Gen. Edward Schmults). The Carter administration, by contrast, took the view that the problem of deterrence was sufficiently serious to warrant a new employee

liability at the federal level have been introduced regularly since *Bivens* was decided,[22] none has been adopted.[23] The bills introduced in the Ninety-eighth Congress illustrate the principal problems involved in the establishment of a system of federal liability for the constitutional wrongs of government officials.

II. Proposed Legislation

A. *H.R. 595*[24]

H.R. 595 would amend the FTCA to provide that the United States would be exclusively liable with respect to a claim arising under the Constitution of the United States for torts committed by government officials acting within the scope of their office or employment.[25] For any constitutional tort claim, the bill provides that the United States would be made liable for the greater of either actual damages or liquidated damages, which would be assessed in an amount that was the greater of either (a) $2,000 or (b) in the case of a continuing violation, $200 per day for each violation.[26] If it is established that the defendant official

---

disciplinary mechanism "to replace the sanction against employee misconduct presumably lost by immunizing employees from civil liability." *1979 House Hearings*, *supra* note 2, at 4 (statement of Deputy Att'y Gen. Benjamin Civiletti).

[22]*See supra* note 2.

[23]Of the 16 bills introduced prior to the 98th Congress, only three—H.R. 7034, 97th Cong., 2d Sess. (1982) (in lieu of H.R. 24, 97th Cong., 1st Sess. (1981)), S. 3314, 95th Cong., 2d Sess. (1978), and H.R. 9219, 95th Cong., 1st Sess. (1977)—were even forwarded to full committee by the responsible subcommittee. Not one of these bills ever was reported out of committee.

[24]98th Cong., 1st Sess. (1983).

[25]H.R. 595 § 202 (proposed 28 U.S.C. § 2699(b)). H.R. 595 and S. 775 provide a remedy against the United States not only with respect to constitutional torts, but also generally in suits based upon acts or omissions of United States officials committed within the scope of office or employment. *See* H.R. 595 § 101 (proposed 28 U.S.C. § 2679(b)); S. 775 § 5(a) (same). Currently, a plaintiff may, as a rule, sue both the official and the government for the same conduct. To do so, the plaintiff must allege a constitutional tort claim against the official under *Bivens*, and a nonconstitutional tort claim against the government under the FTCA. *See* Carlson v. Green, 446 U.S. at 20. Only in a limited set of cases has Congress explicitly made suit against the government under the FTCA an exclusive remedy. *See, e.g.*, 28 U.S.C. § 2679(b) (1976 & Supp. V 1981) (operation of motor vehicles by federal employees); 38 U.S.C. § 4116(a) (1976), 42 U.S.C. §§ 233(a), 2458a (1976 & Supp. V 1981) (malpractice by certain government health personnel). The statutory distinction between constitutional and nonconstitutional torts has been criticized as unsound and unnecessary. *See, e.g.*, *1979 House Hearings*, *supra* note 2, at 123 (statement of George Bermann, Professor of Law, Columbia Univ. Law School).

[26]H.R. 595 § 202 (proposed 28 U.S.C. § 2693). Proposals to replace the *Bivens* remedy typically have provided for liquidated damages to assure successful plaintiffs monetary

acted with malicious intent or reckless disregard for the plaintiff's constitutional rights, additional damages of up to $100,000 would be authorized.[27] A successful claimant also would be entitled to receive reasonable attorney fees and all other litigation costs reasonably incurred, including attorney fees up to twenty-five percent of the judgment, and costs attributable to processing an administrative claim for money damages based on the alleged constitutional tort.[28]

Procedurally under H.R. 595, a suit against an individual government official for an alleged constitutional tort would become a suit against the United States only "[u]pon certification by the Attorney General that the defendant employee was acting within the scope of his office or employment at the time of the incident out of which the suit arose . . . ."[29] The United States would be free to assert as a defense to a constitutional tort claim "the absolute or qualified immunity of the employee of the government whose act or omission gave rise to the claim, or his reasonable good-faith belief in the lawfulness of his conduct."[30]

Finally, H.R. 595 provides that where a constitutional tort action results in a judgment against the United States or an award, compromise, or settlement paid by the United States, "the Attorney General shall forward the matter to the head of

---

recovery where actual damages are small or nonexistent. Such liquidated damages are provided on the premise that proof of concrete injury "is often impossible in the case of nonviolent conduct involving violations of constitutional rights which are of an intangible nature." *1978 Senate Hearing*, *supra* note 17, at 8 (statement of Att'y Gen. Griffin Bell).

[27]H.R. 595 § 202 (proposed 28 U.S.C. § 2693). Notwithstanding the provision for such "additional" damages, H.R. 595 purports to preclude "punitive damages." *Id.* ("[t]he United States . . . shall not be liable . . . for punitive damages.").

[28]*Id.* (proposed 28 U.S.C. § 2698(a)). Under H.R. 595, an action could not be instituted against the United States unless the claim had first been presented to, and denied by, the appropriate federal agency. *Id.* (proposed 28 U.S.C. § 2694(a)). A similar condition applies to suits against the United States under the FTCA. *See* 28 U.S.C. § 2675 (1976).

[29]H.R. 595 § 202 (proposed 28 U.S.C. § 2698(d)(1)). Certification by the Attorney General would trigger removal to federal court of a suit brought in a state court. *Id.* (proposed 28 U.S.C. § 2698(d)(2)). The Attorney General's certification that the defendant official was acting within the scope of his office or employment at the time of the incident would be "binding and conclusive [only] on the United States and the defendant employee," *id.* (proposed 28 U.S.C. § 2698(d)(3)), and not on the party bringing the action.

If a plaintiff successfully challenges the Attorney General's certification upon removal, the district court would be required to remand the case to the state court, inasmuch as the Attorney General's certification was the predicate for removal. *Id.* If the Attorney General did not make such a certification, the defendant official would be allowed to request the federal district court—but apparently not the state court—in which the suit has been filed to certify that the defendant was acting within the scope of his office or employment at the time of the alleged injury. *Id.*

[30]*Id.* (proposed 28 U.S.C. § 2693).

the Federal agency which employed the employee at the time of the employee's alleged act or omission giving rise to the claim . . . for such further administrative investigation or disciplinary action as may be appropriate."[31] H.R. 595 is drafted to preserve the right to a jury trial in constitutional tort cases, a right now provided by the *Bivens* remedy.[32]

B. *S. 775*[33]

The Senate bill is similar to H.R. 595 except for two differences.[34] First, no provision is made under S. 775 for any award to a prevailing plaintiff of attorney fees or other litigation costs, including costs of processing administrative claims. And second, S. 775 does not provide for jury trial in all constitutional tort cases, but only in cases where demand for trial is made prior to or on the date of enactment of the bill.[35]

III. Issues

There is broad bipartisan support for the proposition that the United States, and not individual federal officials, should bear the ultimate financial responsibility for constitutional violations.[36] But differences have arisen over specific elements of

---

[31]*Id.* (proposed 28 U.S.C. § 2700).

[32]*Id.* (proposed 28 U.S.C. § 2695). Significantly, the power to award additional damages for malicious constitutional torts is vested under H.R. 595 in the court, not in the jury. *See supra* note 27.

[33]98th Cong., 1st Sess. (1983).

[34]Under the legislation introduced by Senator Grassley during the 97th Congress, the individual officer and the United States would be jointly liable for punitive damages of up to $50,000 for the employee's willful and malicious violation of constitutional rights. S. 1775, 97th Cong., 1st Sess. § 3(4) (1981) (proposed 28 U.S.C. § 2674(b)(3)) (amendment of Sen. Specter, adopted Sept. 23, 1982). The United States would be entitled to indemnity for any amount paid under this subsection. *Id.* Senator Grassley's current bill does not include this provision. The Senator did include a provision added to S. 1775 during subcommittee mark up to amend 50 U.S.C. § 1810 so that it would not apply to any civil cause of action against an official while acting within the scope of his office or employment. S. 775 § 10. Finally, under S. 775, the maximum amounts allowed for liquidated damages also are lower than those set by H.R. 595—the greater of $1,000 or, in the case of a continuing violation, $100 per day for each day up to a maximum of $15,000. S. 775 § 3(4) (proposed 28 U.S.C. § 2674(b)(2)).

[35]S. 775 § 11(b)(1).

[36]*See, e.g.*, Federal Officials' Liability for Constitutional Violations, 47 Fed. Reg. 58,207 (1982) (to be codified at 1 C.F.R. § 305.82-6) [hereinafter cited as *ACUS Recommendation*]. There also is substantial scholarly support for the proposition. *See* K. Davis, Administrative Law Text § 26.06, at 492–93 (3d ed. 1972); Bermann, *Inte-*

legislation that would create an alternative system of civil sanctions. The following issues are the focus of the current legislative controversy.

A. *Retention of Good-Faith Immunity*

The most significant and hotly disputed issue in the establishment of a system of exclusive governmental liability is whether the government should be allowed, as provided in H.R. 595 and S. 775, to assert as a defense an official's "qualified immunity . . . or his reasonable good-faith belief in the lawfulness of his conduct."[37] We submit that retention of good-faith immunity[38]

---

grating Governmental and Officer Tort Liability, 77 COLUM. L. REV. 1175, 1194–1203 (1977). But support for the proposition is not universal. *See 1982 House Hearings*, *supra* note 11, at 431, 435–40 (statement of Thomas Devine, Legal Director of the Gov't Accountability Project (GAP), Inst. for Pol'y Studies). Disputing the premise that the threat of personal liability "functionally paralyzes any conscientious federal employee," *id.* at 435, and that judgment-proof government officials thwart the goal of victim compensation, *id.* at 438–39, GAP proposed a modified version of the then-current House bill H.R. 7034 that would have supplemented, but not supplanted, the *Bivens*-type remedy. *Id.* at 451.

Professor Neuborne of New York University Law School has suggested that the FTCA simply be amended to permit *Bivens* defendants to implead the United States, establishing "a self-executing mechanism [that] would allocate the economic loss caused by a constitutional violation to the proper party." *1981 Senate Hearings*, *supra* note 20, at 29 (statement of Professor Burt Neuborne, Professor of Law, New York Univ. Law School, on behalf of American Civil Liberties Union (ACLU)).

These proposals for an alternative to an exclusive system of governmental liability echo the joint government and official liability recommended by the Senate Select Committee to Study Governmental Operations with Respect to Intelligence Activities, S. REP. NO. 755, *supra* note 2, at 336–38, and the Senate Committee on Government, S. REP. NO. 588, 93d Cong., 1st Sess. 2–4, *reprinted in* 1974 U.S. CODE CONG. & AD. NEWS 2789 (analyzing proposed amendment to H.R. 8245; H.R. 8245 was enacted into law as Act of Mar. 16, 1974, Pub. L. No. 93-253 § 2, 88 Stat. 50, 50 (codified at 28 U.S.C. § 2680(h) (1976 & Supp. V 1981)), to provide, among other things, a remedy against the United States for certain intentional nonconstitutional torts of its investigative and law enforcement officers). The Carter administration did not favor legislation that would allow a plaintiff to choose between suing the official and suing the government. *1978 Senate Hearing*, *supra* note 17, at 119 (letter from Ass't Att'y Gen. Patricia Wald to Sen. Abourezk (Mar. 13, 1978)).

[37]H.R. 595 § 202 (proposed 28 U.S.C. § 2693); S. 775 § 5(b) (proposed 28 U.S.C. § 2679(d)). The Reagan administration's proposal would allow the government to assert all defenses available to an individual employee, including the absolute or qualified immunity of the employee. S. 829, 98th Cong., 1st Sess. § 1303 (1983) (proposed 28 U.S.C. § 2674). Both H.R. 595 and S. 775 would allow the government to assert the absolute immunity of officials entitled to such immunity. Although many of the arguments for eliminating qualified immunity would seem applicable to absolute immunity, there is a strong consensus in Congress that the government should not be liable for conduct already protected by absolute immunity.

[38]The Supreme Court appears to define qualified immunity as the right to invoke the good-faith defense. *See* Harlow v. Fitzgerald, 102 S. Ct. 2727, 2737 (1982). For convenience, the issue of whether to allow the government to assert either the qualified-

in a system that purports to establish exclusive governmental liability is unfounded and ill conceived: it would undermine the objectives of compensation and accountability without relieving officials of the burdens now imposed by *Bivens*-type litigation.[39]

At the outset, it should be understood that nothing in established tort doctrine applicable to every analogous situation of private and public vicarious liability supports the proposition that the United States should be permitted to invoke the good-faith immunity of its officials in constitutional tort suits. At common law, the majority rule in the private employment context is that an employer sued for the tort of an employee under the doctrine of respondeat superior may not invoke the employee's immunities.[40] Moreover, the majority common-law rule in the public employment context at the state and local levels is that a government agency sued for the tort of a public official under the doctrine of respondeat superior may not invoke the official's immunities.[41] Indeed, in FTCA suits, virtually every federal court that has considered the issue has concluded that (a) the doctrine of respondeat superior does not permit the United States to invoke the immunities of its officials,[42] and

---

immunity or the good-faith defenses of its officials is discussed herein under the rubric of whether to allow the government to assert the defense of good-faith immunity. The convention is consistent with established usage. *See, e.g.*, Saldana v. Garza, 684 F.2d 1159, 1162–63 (5th Cir. 1982).

Several earlier bills, in contrast to H.R. 595 and S. 775, would have precluded the United States from asserting—except in cases involving a Member of Congress, a judge, a prosecutor, or a person performing analogous functions—the absolute or qualified immunity of the official whose conduct was at issue, or the official's reasonable good-faith belief in the lawfulness of his conduct. *See* H.R. 24, 97th Cong., 1st Sess. § 3 (1981); S. 695, 96th Cong., 1st Sess. § 3 (1979); H.R. 2659, 96th., 1st Sess. § (3) (1979); S. 3314, 95th Cong., 2d Sess. § 3(b) (1978). This approach was supported by the Carter administration. *See infra* note 48. Under S. 2117, 95th Cong., 1st Sess. § 3 (1977) and H.R. 9219, 95th Cong., 1st Sess. § 3 (1977), a federal officer's or employee's "good faith reliance on a court order or legislative authorization" would have constituted a complete defense in a constitutional tort suit against the United States, but the United States would nevertheless have been required to compensate the person whose constitutional rights had been violated.

[39]*But see ACUS Recommendation*, *supra* note 36.

[40]*See, e.g.*, Schubert v. August Schubert Wagon Co., 249 N.Y. 253, 164 N.E. 42 (1928); Hallmark Ins. Co. v. Crary Enters., 72 Wis. 2d 472, 475, 241 N.W.2d 171, 173 (1976); RESTATEMENT (SECOND) OF AGENCY § 217 Comment b (1958); 2 F. HARPER & F. JAMES, THE LAW OF TORTS § 26.17, at 1427 (1956); W. SEAVEY, HANDBOOK ON THE LAW OF AGENCY § 93, at 167 (1964).

[41]*See, e.g.*, James v. Prince George's County, 288 Md. 315, 331–32, 418 A.2d 1173, 1182 (1980); Muntan v. City of Monongahela, 45 Pa. Commw. 23, 26–28, 406 A.2d 811, 813–14 (1979); Maynard v. City of Madison, 101 Wis. 2d 273, 283–84, 304 N.W.2d 163, 169 (1981).

[42]*See* United States v. Trubow, 214 F.2d 192, 196 (9th Cir. 1954); Jackson v. United States, 196 F.2d 725, 726 (3d Cir. 1952); United States v. Hull, 195 F.2d 64, 68 (1st Cir.

(b) Congress did not intend for the United States to be able to invoke the immunities of its officials in suits brought under 28 U.S.C. § 2680(h).[43] Allowing the United States to plead the good-faith immunity of federal officials in constitutional tort suits would afford the government an advantage enjoyed neither by private and public employers at common law, nor by the United States in nonconstitutional tort suits.

The principal argument behind retaining good-faith immunity is that where the official has acted in good faith, not only has no genuine wrongdoing occurred, but a judgment against the government in such cases would lead to an unwarranted imputation of wrongdoing on the official's part.[44] This argument does not withstand close scrutiny. Good-faith immunity is made available to the individual government official not because the official's good faith somehow negates the fact of constitutional injury or because a determination of liability would unfairly stigmatize him. Instead, "underlying the qualified immunity which public officials enjoy for actions taken in good faith is the

---

1952); Downs v. United States, 382 F. Supp. 713, 749–51 (M.D. Tenn. 1974), *rev'd on other grounds*, 522 F.2d 990 (6th Cir. 1975). The only decision to the contrary appears to be Brooks v. United States, 152 F. Supp. 535, 537 (S.D.N.Y. 1957).

[43]Four of the five district courts that have considered the issue have reached this conclusion. *See* Picariello v. Fenton, 491 F. Supp. 1020, 1040–42 (M.D. Pa. 1980); Towsend V. Carmel, 494 F. Supp. 30, 36–37 (D.D.C. 1980); Crain v. Krehbiel, 443 F. Supp. 202, 216–17 (N.D. Cal. 1978); Norton v. Turner, 427 F. Supp. 138, 146–52 (E.D. Va. 1977), *rev'd sub nom.* Norton v. United States, 581 F.2d 390, 393–96 (4th Cir. 1978); *see also* Recent Decisions, *Sovereign Immunity*, 47 GEO. WASH. L. REV. 651 (1979) (criticizing Fourth Circuit's *Norton* decision). The Fourth Circuit, declining to rely on the common-law doctrine of respondeat superior, has held that Congress intended to permit the United States to plead the qualified immunity of its officials in suits under 28 U.S.C. § 2680(h) (1976 & Supp. V 1981). Norton v. United States, 581 F.2d 390, 393–96 (4th Cir. 1978), *rev'g* 427 F. Supp. 138, 146–52 (E.D. Va. 1977); *cf.* Downs v. United States, 522 F.2d 990, 998 (6th Cir. 1975) (holding that immunity of police officer from damage suit is not available to government under FTCA where basis of officer's immunity is policy of encouraging uninhibited exercise of discretion).

[44]*See 1982 House Hearings*, *supra* note 11, at 334 (statement of Donald Baldwin, Executive Director, Nat'l Law Enforcement Council); *id.* at 343–44 (statement of Ira Lechner, Legislative Counsel, Nat'l Ass'n of Police Orgs.); *id.* at 138 (statement of Joseph Morris, Gen. Counsel, Office of Personnel Management).

Viewed in this light, good-faith immunity is close to a denial not only of culpability but also that any *injury* has occurred. Thus, Deputy Attorney General Edward Schmults has stated the administration's view that good-faith immunity

> really goes to the merits of the plaintiff's claim by testing the action of an employee against the standard of reasonableness and good faith . . . . And certainly no employee wants to be found "guilty," if you will, of unconstitutional acts and suffer the resulting stigma. So that, even where the United States and not the employee would be the defendant, our view is that employees would be discouraged from acting in uncertain areas where they might subject the Government to financial liability.

1981 Senate Hearings, *supra* note 20, at 9.

fear that exposure to personal liability would otherwise deter them from acting at all."[45] Although a government official's good faith may be relevant to whether a party whose constitutional rights have been violated should be entitled to compensation from the official, the official's good faith cannot be relevant to whether the victim is entitled to compensation from the government.[46] Allowing the government to assert good-faith immunity also would limit the operation of any disciplinary mechanism that would be triggered by a determination of liability or settlement by the government.[47] At least in the view of the current administration, disciplinary proceedings would be activated only when the official had intentionally violated the plaintiff's constitutional rights.[48]

---

[45]Carlson v. Green, 446 U.S. at 21 n.7 (*citing* Butz v. Economou, 438 U.S. 478, 497 (1978)); *see* Owen v. City of Independence, 445 U.S. 622, 638 (1980) (city may not assert good faith of its officers as defense to a 42 U.S.C. § 1983 action); Scheuer v. Rhodes, 416 U.S. 232, 240 (1974). As FBI Director William Webster has stated, good-faith immunity "is for the individual protection of the agent." *1978 House Hearings*, *supra* note 14, at 97.

[46]*See* Norton v. United States, 581 F.2d at 398 (Butzner, J., dissenting); *1979 House Hearings*, *supra* note 2, at 126–27 (testimony of George Bermann, Professor of Law, Columbia Univ. Law School). As Deputy Attorney General Benjamin Civiletti testified:

> the good faith defense doesn't mean that the conduct is not regrettable, terribly negligent, that everyone doesn't feel awful remorse about it, and there hasn't been real harm or substantial harm, if not physically, to money values or at least to the dignity of the individual and the privacy of the individual.
>
> I will give you a quick example. Investigators, three of them, go to a home of suspects looking for racketeering—numbers, or whatever. When they go in the home and they are queried about it later, each one says he thought the other had a search warrant. None of them had a search warrant.
>
> They find a man and woman 75 years old in there. They take the woman and strip the woman down and do cavity searches and whatever—in the wrong home.
>
> That is not a gift when you are talking about compensation for that kind of wrong under the fourth amendment.

1979 House Hearings, *supra* note 2, at 19.

[47]*See* H.R. 595 § 202 (proposed 28 U.S.C. § 2700); S. 775 § 5(c) (proposed 28 U.S.C. § 2679 (f)).

[48]To be sure, even if the government were permitted to assert good-faith immunity, the agency that employed the official whose conduct was challenged theoretically would be free to initiate its own internal investigation and disciplinary proceedings under existing law at any time, irrespective of the pendency of the lawsuit. But the current administration has strongly implied that a determination of liability or money settlement by the United States would be a necessary predicate for disciplinary action against the defendant official. *See 1981 Senate Hearings*, *supra* note 20, at 9 (testimony of Deputy Att'y Gen. Edward Schmults) (arguing against waiver of good-faith immunity on ground that disciplinary proceedings would be triggered by determination of liability, and that where the official has acted in good faith "there should be no disciplinary proceedings").

As initially drafted (in the form of H.R. 24), H.R. 7034, 97th Cong., 2d Sess. (1982), was criticized by some for eliminating good-faith immunity in constitutional tort suits. *See 1982 House Hearings*, *supra* note 11, at 341 (statement of J. Brian Hyland, President, Ass'n of Fed. Investigators); *id.* at 343–45 (statement of Ira Lechner, Nat'l Ass'n of Police Orgs.). The original good-faith waiver in H.R. 24 was specifically endorsed

As for the claim that a determination of liability absent bad faith would unfairly stigmatize the defendant official, it must be remembered that in such a case the official did commit a constitutional injury, even though his conduct may have been unintentional—even well-meant. Solicitude for the wrongdoer's reputation should not be allowed to override the victim's right to a remedy, especially when there is no possibility of recovering damages from the official.[49] Moreover, the official's good faith should not be allowed to remove the basis for disciplining him on grounds other than intent, such as negligence or incompetence. In short, there is simply no reason why an official should not be held accountable for his misfeasance.[50]

Congress, of course, may specifically provide by legislation that no finding of liability on the government's part in a constitutional tort case shall give rise to any adverse inference as to

---

by, among others, the Federal Executive and Professional Association (FEPA), *see id.* at 328 (statement of Richard Pelz, President, FEPA), and waiver of immunity under earlier legislative proposals was supported by the FBI, *see 1979 House Hearings*, *supra* note 2, at 80 (statement of William Webster, Director, FBI) and by the Carter administration, *see 1978 House Hearings*, *supra* note 14, at 14 (statement of Irving Jaffe, Deputy Ass't Att'y Gen., Civil Div., Dep't of Justice), as well as by the American Bar Association, *1979 House Hearings*, *supra* note 2, at 152 (statement of B. James George, Jr., Former Chairperson, Criminal Justice Section, ABA).

[49]In actions for declaratory and injunctive relief, the good faith vel non of government officials is not an issue in determining whether such relief should be awarded. *See generally* Note, *supra* note 9, at 604, 607–17. If a constitutional violation has occurred, a court must enjoin its continuance whether or not the officials involved had a good-faith belief in the legality of their actions. *See, e.g.*, O'Connor v. Donaldson, 422 U.S. 563, 577 n.12 (1975); Knell v. Bensinger, 522 F.2d 720 (7th Cir. 1975); National Treasury Employees Union v. Nixon, 492 F.2d 587, 609 (D.C. Cir. 1974); Sostre v. McGinnis, 442 F.2d 178 (2d Cir. 1971); *see also* McCormick v. Attala County Bd. of Educ., 541 F.2d 1094 (5th Cir. 1976); Hutchinson v. Lake Oswego School Dist. No. 7, 519 F.2d 961 (9th Cir. 1975); Hupart v. Board of Higher Educ., 420 F. Supp. 1087, 1107–08 (S.D.N.Y. 1976); Demkowicz v. Endry, 411 F. Supp. 1184 (S.D. Ohio 1975).

[50]Public Citizen (a nonprofit, public interest organization) criticized H.R. 7034, 97th Cong., 2d Sess. (1982), for its original *limited* retention (in H.R. 24) of good-faith immunity with respect to the conduct of Members of Congress, judges, prosecutors, or others performing analogous functions. *1982 House Hearings*, *supra* note 11, at 357–58 (statement of Alan Morrison, Director of Litig., Public Citizen). Presumably, new legislation proceeding along those lines would now provide for retention of immunity with respect to the conduct of the President as well. *See* Nixon v. Fitzgerald, 102 S. Ct. 2690, 2705 (1982). Nevertheless, there would appear to be some merit to Mr. Morrison's contention that "all the reasons that lead to the elimination of the good-faith defense in all other situations apply here." *1982 House Hearings*, *supra* note 11, at 357 (statement of Alan Morrison). If the concern is to spare such officials the ordeal of administrative proceedings in the event that liability is established, future legislation might be drafted to bar any administrative investigation or proceeding against such officials upon a showing that an otherwise available immunity could have barred recovery in a given case in which liability was found. The power of Congress to abrogate the judicially established absolute or qualified immunity of any federal official, even without transferring liability to the United States, seems clear. *See* Nixon v. Fitzgerald, 102 S. Ct. at 2719 n.27.

the official's good faith. Indeed, it would appear sound for Congress to provide that once the United States has been substituted for an official in a constitutional tort suit, the official shall be free to treat that suit—for instance, in answering questions of potential creditors, employers, or others—as though he never had been named as a defendant. Congress also could provide specifically that the defendant official is entitled to assert a good-faith defense in any administrative investigation or disciplinary proceeding under the statute.[51]

Waiver of good-faith immunity by the government is necessary to reduce the official's involvement in litigation. Such a waiver also would operate to limit discovery and facilitate the disposition of claims.[52] Reducing constitutional tort suits to two issues—did a violation of the plaintiff's constitutional rights occur, and was that injury caused by a government official acting within the scope of his office or employment—could simplify trials and encourage early settlement.[53]

Proponents of the government's retention of an official's good-faith immunity often note that the recent Supreme Court decision of *Harlow v. Fitzgerald* was intended to simplify the

---

[51]*See* H.R. 595 § 202 (proposed 28 U.S.C. § 2700). As Deputy Assistant Attorney General Irving Jaffe noted in connection with legislation supported by the Carter administration, "in any disciplinary proceeding, the good faith of the employees will be very much an issue, and if [a defendant official] can establish that he had a good-faith belief that what he was doing was lawful and proper, then disciplinary proceedings on that basis would not proceed." *1978 House Hearings*, *supra* note 14, at 14.

[52]*1982 Senate Hearing*, *supra* note 16, at 125. As the Justice Department observed in commenting on the cost impact of earlier legislation providing for waiver of good-faith immunity:

> Major discovery savings would . . . accrue to the agencies and departments employing those individuals who have been sued. Currently, the burden on an agency can be immense. The FBI, for example, has had to examine hundreds of thousands of documents in connection with suits against individuals. In addition to the savings associated with the discovery process, simplification of the issues involved in a constitutional tort action will reduce the amount of attorney time necessary for each case.

1979 House Hearings, *supra* note 2, at 13 n.1 (Dep't of Justice Memorandum of May 7, 1979).

[53]The strength of the official's good-faith immunity has been an important element in settlement discussions. *1982 House Hearings*, *supra* note 11, at 251 (testimony of Royce Lamberth, Chief, Civil Div., Office of the U.S. Att'y for the Dist. of Columbia). The assertion of good-faith immunity may be expected frequently to postpone settlement pending appellate review of a district court's ruling on summary judgment for either party with respect to the so-called "objective" factors upon which good-faith immunity must now rely. *See, e.g.*, McSurely v. McClellan, 697 F.2d 309, 315–16, 321 n.20 (D.C. Cir. 1982) (post-*Harlow* opinion); *see also* Forsyth v. Kleindienst, No. 82-1812, slip op. at 3, 4 & n.2 (3d Cir. Jan. 20, 1983) (opinion sur grant of stay) (acknowledging that appealing denial of summary judgment on qualified-immunity issue is "not free from difficulty" after *Harlow*).

trial of the good-faith issue.[54] Although *Harlow* may have somewhat simplified litigating the good-faith issue with respect to the question of unclear applicable law, other subjective elements certainly remain to be asserted or denied in establishing the official's state of mind, and thus the difficulty surrounding proof of good faith will continue to complicate suits alleging official misconduct. This forecast is borne out by post-*Harlow* cases.[55]

Another argument advanced in favor of retaining good-faith immunity for the government in constitutional tort suits is that waiver of immunity would enable plaintiffs to "prevail more easily, thereby increasing economic costs to the [g]overnment."[56] As to this argument, it should be noted that any increased cost to the government by virtue of its expanded liability will be offset to some extent by the savings that would result from both the simplification of the litigated issues and the incentive for early settlement. And, certainly, the "economic costs to the government" may be moderated by setting liquidated damages in a reasonable amount.[57] But there is a more compelling reason for allowing recovery: "[W]here someone has a genuine grievance—where someone has a genuine claim that he has been deprived of a constitutional right—we ought to afford him avenues of relief without regard to cost. We cannot put a price on constitutional liberties."[58]

If legislation removes the liability of officials for constitutional torts, the policy choice is whether the government or the victim

---

[54]102 S. Ct. 2727, 2736–39 (1982). Defining "[q]ualified or 'good faith' immunity," the Court held that "government officials performing discretionary functions generally are shielded from liability for civil damages insofar as their conduct does not violate clearly established statutory or constitutional rights of which a reasonable person would have known." *Id.* at 2738. This "[r]eliance on the objective reasonableness of an official's conduct, as measured by reference to clearly established law," *id.* at 2739, supersedes reliance on both objective factors and subjective factors (*i.e.*, "permissible intentions") in determining whether an official acted in good faith. *Id.* at 2737.

[55]*See, e.g.*, Briggs v. Goodwin, No. 80-2269 (D.C. Cir. Jan. 11, 1983); McSurely v. McClellan, 697 F.2d 309 (D.C. Cir. 1982); Dale v. Bartels, 552 F. Supp. 1253 (S.D.N.Y. 1982); *cf.* Sampson v. King, 693 F.2d 566 (5th Cir. 1982); Saldana v. Garza, 684 F.2d 1159 (5th Cir. 1982); United States v. Irving, 684 F.2d 494 (7th Cir. 1982); Standridge v. City of Seaside, 545 F. Supp. 1195 (N.D. Cal. 1982). Indeed, the Reagan administration's most recent position on the subject appears to concede that considerable litigation will occur. The administration curiously argues that such litigation is beneficial. *See* 129 CONG. REC. S3154 (daily ed. Mar. 16, 1983).

[56]*1982 House Hearings*, *supra* note 11, at 138 (statement of Joseph Morris, Gen. Counsel, Office of Personnel Management).

[57]*See supra* note 34.

[58]*1982 House Hearings*, *supra* note 11, at 141 (testimony of Joseph Morris, Gen. Counsel, Office of Personnel Management) (arguing that the possibility of increased damage awards by virtue of substitution of the government as defendant in constitutional tort suits is a risk worth taking).

should bear the loss. We can discern no substantial public interest that would be served by requiring individual citizens who are harmed by even well-intentioned officials to bear the cost of constitutional injury. The loss is a cost of government activity and should be borne by the public.

B. *Substitution of Effective Mechanisms of Deterrence*

The Supreme Court has specifically emphasized "the doubt . . . cast on the validity of the assumption that there exist adequate mechanisms for disciplining federal employees" who violate constitutional rights.[59] The Court has implied that the threat of disciplinary action under existing mechanisms is, at most, a supplement to, and not a substitute for, the deterrence imposed by the risk of personal liability.[60]

A basic question, then, for legislation replacing official liability with governmental liability is how to ensure that the substitute deterrent mechanism will serve the goal of measured deterrence the *Bivens* remedy was designed to promote. The approach taken in H.R. 595 and S. 775 is one of bare reliance on existing mechanisms.[61] In this respect, the bills are in marked contrast to earlier proposals whose somewhat elaborate provisions for administrative action upon a determination of liability,

---

[59]Carlson v. Green, 446 U.S. at 21 & n.8, *citing 1978 House Hearings*, *supra* note 14, at 6 (statement of Att'y Gen. Griffin Bell).

[60]*Id.* at 21. The efficacy of existing mechanisms has been questioned especially in the context of the law enforcement and the intelligence agency abuses revealed during the last decade. *See 1979 House Hearings*, *supra* note 2, at 45 & n.32 (statement of Karen Christensen, Legislative Counsel, ACLU); *see also* U.S. DEP'T OF JUSTICE, REPORT CONCERNING INVESTIGATION AND PROSECUTORIAL DECISIONS WITH RESPECT TO CENTRAL INTELLIGENCE AGENCY MAIL OPENING ACTIVITIES IN THE UNITED STATES (1977); Association of the Bar of the City of New York, Committee on Civil Rights, *Intelligence Agency Abuses: The Need for a Temporary Special Prosecutor*, 31 REC. A.B. CITY N.Y. 601 (1976); *Prosecution for Domestic Spying Rejected in 1977*, N.Y. Times, Mar. 15, 1982, at A12, col. 2 (describing contents of 1976 report by Justice Department acknowledging apparent violation of "fundamental constitutional rights" by federal agencies and officials, but recommending against criminal prosecution).

One witness noted in hearings on H.R. 2659, 96th Cong., 1st Sess. (1979), the egregious lack of deterrent mechanisms for certain federal officials in connection with the FBI's administrative inquiry into illegal investigative activities directed against the Weather Underground. Of 61 special agents and seven supervisors suspected of illegal entries, illegal wiretaps, and numerous illegal mail openings, disciplinary action was taken against only two of the agents and four of the supervisors. *1979 House Hearings*, *supra* note 2, at 45 & n.32 (statement of Karen Christensen, ACLU).

[61]*See* H.R. 595 § 202 (proposed 28 U.S.C. § 2700); S. 775 § 5(c) (proposed 28 U.S.C. § 2679 (f)); *cf. ACUS Recommendation*, *supra* note 36.

and for subsequent judicial review, may have accounted for their failure to have been adopted.[62]

Plainly some middle ground must be found. An administrative mechanism that depends on an agency's inclination to repudiate the conduct of its officials cannot be expected to foster public confidence that officials either will be made to account for their actions, or will be effectively dissuaded from future misconduct.[63] On the one hand, an agency may be reluctant to punish officials who have committed wrongful acts in the belief that "they were doing their duty to their country and to their agency."[64] At the same time, however, internal proceedings may afford the agency an opportunity to make the defendant official serve as a scapegoat for the wrongdoing of his superiors, or in retaliation for "whistleblowing." Indeed, administrative proceedings arising from constitutional torts committed by defendant officials frequently involve claims implicating agency policy or policymakers and thereby compromise the agency's impartiality.

For this reason, some have urged that disciplinary proceedings should "not be conducted by the head of the agency in which the employee or employees work," but instead that an independent body should be designated for that purpose.[65] Reliance on an independent body would remove the discharge of disciplinary responsibilities from those with the most incentive to misuse it. Presumably, this would meet a few of the concerns aired by members of the Carter administration and public interest groups, both of which had argued for victim participation in independent disciplinary proceedings as "essential" to any

---

[62]*See* H.R. 2659, 96th Cong., 1st Sess. § 8 (1979); S. 695, 96th Cong., 1st Sess. § 12(a) (1979); S. 3314, 95th Cong., 2d Sess. § 13 (1978). H.R. 2659 and S. 695 were proposals supported by the Carter administration.

[63]The concerns expressed go beyond the law enforcement and intelligence context. Senator Howard Metzenbaum (D-Ohio) stated, to Attorney General Bell's expression of assent, that:

> [I]n the case of employees [accused of wrongdoing], you have to have some kind of a review board that is totally unresponsive to the agency that is itself involved in the alleged wrongdoing. Or else the person who claims to have been wronged will never feel that they got justice; and probably the American people, when they learn about it, will not think they got justice. It will always be suspect—whether it is the CIA investigating its own people, or Cointel, or whether it is HEW or any other agency.

1978 Senate Hearing, *supra* note 17, at 13.

[64]*1979 House Hearings*, *supra* note 2, at 127 (remarks of Rep. Robert McClory (R-Ill.)).

[65]*1982 House Hearings*, *supra* note 11, at 330–31 (statement of Richard Pelz, FEPA).

"meaningful substitute" for the existing mechanism of actual and punitive damages.[66]

One approach along these lines would be to establish, for independent investigation and, in appropriate cases, for prosecution of disciplinary proceedings, an Office of Disciplinary Counsel modeled on the existing Office of Special Counsel under the Civil Service Reform Act of 1978.[67] Another approach would be to utilize the Merit Systems Protection Board (MSPB) as a disciplinary body.[68]

The objective of independent control should not be to replace or needlessly duplicate the existing administrative disciplinary system. The agency that employs the official whose conduct has been challenged should continue to be responsible in the first instance for investigating and, where appropriate, for disciplining the official or for implementing other corrective steps even prior to the outcome of the constitutional tort suit against the government. Because existing disciplinary mechanisms are directed toward conduct related to job performance and are designed to promote administrative efficiency, it also would be desirable for Congress, quite apart from the need for indepen-

---

[66]Under current law an investigation may, of course, result from a victim's complaint, *e.g.*, 5 U.S.C. § 1206 (Supp. V 1981), but there is no automatic initiation of an investigation upon a private complaint. *See 1982 House Hearings*, *supra* note 11, at 361 (statement of Alan Morrison, Director of Litig., Public Citizen); *1978 Senate Hearing*, *supra* note 17, at 6 (statement of Att'y Gen. Griffin Bell) (advocating "procedures in which the injured person can participate in a meaningful way"). In its testimony on H.R. 24, 97th Cong., 1st Sess. (1981), Public Citizen called for the creation of:

> (1)[T]he right of the victim to initiate an investigation which cannot be terminated without adequate reasons;
> (2)the right of the victim to participate in an appropriate manner in the investigation and subsequent disciplinary proceedings, if any; and
> (3)the right of the victim to agency, and ultimately judicial, review if no punishment is meted out, or if the punishment is wholly inadequate for the violation.

1982 House Hearings, *supra* note 11, at 361 (statement of Alan Morrison). These recommendations were incorporated in legislation supported by the Carter administration. *See* Bell, *supra* note 3, at 12–13.

[67]Pub. L. No. 95-454 § 202, 92 Stat. 1111, 1121 (codified at 5 U.S.C. §§ 1201–1209 (Supp. V 1981)). Establishing an Office of Disciplinary Counsel was proposed by the Fund for Constitutional Government. *See 1979 House Hearings*, *supra* note 2, at 74. The Special Counsel is already authorized to investigate abuses of authority in the context of prohibited personnel practices. 5 U.S.C. § 1206(a)(4) (Supp. V 1981).

[68]*1982 House Hearings*, *supra* note 11, at 311 (testimony of Richard Pelz, FEPA), *citing* 5 U.S.C. § 1207 (Supp. V 1981) (authorizing MSPB to impose disciplinary sanctions). The MSPB is authorized, inter alia, to hear and adjudicate allegations of prohibited personnel practices, 5 U.S.C. § 1205(a) (Supp. V 1981), and to enforce compliance with its orders by any federal agency or employee. *Id.* § 1205(a)(2). Compliance may be enforced by orders that salary payments be withheld pending compliance. *Id.* § 1205(d)(2).

dent disciplinary controls, to provide explicit authorization for applying existing disciplinary mechanisms to official misconduct that causes constitutional injuries.[69]

C. *Jury Trial and Additional Damages*

In rejecting the argument that the FTCA provides as effective a remedy as that of *Bivens*-type actions, the Supreme Court has specifically stressed the availability of a jury trial and punitive damages in *Bivens* suits.[70] Opposition to jury trials in *Bivens*-type cases stems largely from the belief that juries are biased in such situations in favor of plaintiffs. This notion has not impressed the Court.[71] In view of the small number of judgments in favor of plaintiffs in *Bivens* suits,[72] concern with this issue does not appear to be serious. Admittedly, juries may be more ready to award damages against the "deep pocket" government than against individual *Bivens* defendants. But any legitimate concern about the risk of excessive jury awards should be met by setting a ceiling on the amount of damages, especially those that may be awarded for a constitutional tort committed with malice, and not by denying the plaintiff either such damages altogether or the the right to a jury trial.

Needless to say, punitive damages that otherwise would be available against an official are more properly cast as additional or exemplary when the government becomes the defendant. The object of assessing such damages against the government is not to punish the sovereign. Instead, the object is to express sharp

---

[69]*See ACUS Recommendation*, *supra* note 36.

[70]Carlson v. Green, 446 U.S. at 22–23. A jury trial is currently provided for under the FTCA only in civil actions against the United States for the recovery of internal revenue taxes erroneously or illegally assessed or collected, penalties collected without authority, or any other sum alleged to have been excessive or in any manner wrongfully collected under the internal revenue laws. 28 U.S.C. § 2402 (1976).

The Supreme Court noted that punitive damages may be the only significant remedy available where constitutional rights are maliciously violated and the victim cannot prove compensable injury. 446 U.S. at 22 n.9. This problem would not arise under legislation such as H.R. 595 and S. 775, which provide for liquidated as well as actual damages.

[71]Carlson v. Green, 446 U.S. at 22–23.

[72]According to Deputy Attorney General Schmults, "several thousand" constitutional tort actions had been filed as of October 1981, but only nine had resulted in money judgments against federal employees. *1982 House Hearings*, *supra* note 11, at 23; *see* Bell, *supra* note 3, at 2 n.5 (listing seven cases in which money judgments had been entered for plaintiffs in *Bivens* actions). The number of cases settled by the parties is unknown.

social disapproval of constitutional torts committed with malice, to spur the government to exercise administrative controls, and to minimize the possibility that such conduct will recur.[73]

It also should be noted that allowing additional damage awards against the United States in appropriate cases may be constitutionally *required* on the theory that Congress may not strip the federal judiciary of the power to fashion remedies held necessary to ensure the protection of constitutional rights.[74] At least in the context of civil rights litigation, the Court has suggested that some availability of punitive damages is necessary.[75] Whether the right to jury trial in constitutional tort cases may be eliminated would seem to depend, at least in part, on the credibility of the requisite congressional declaration that the new nonjury trial remedy would be "equally effective" as the existing jury trial remedy.[76]

---

[73]*See 1982 House Hearings*, *supra* note 11, at 328–29 (statement of Richard Pelz, FEPA) ("[T]he Attorney General in settling a suit [should be allowed] to agree to punitive damages, which would constitute a statement on his part that the action of the Federal employee or employees transcended the bound[s] of acceptability and excuse."). Although subjecting the government to punitive damages liability for the bad faith of officials without permitting the government to assert a qualified-immunity defense for the good faith of officials seems asymmetrical, this result is consistent with broader principles of tort law. Good intent does not ordinarily immunize the tortfeasor while bad intent may increase his liability.

[74]*See 1981 Senate Hearings*, *supra* note 20, at 25–26 (statement of Burt Neuborne, Professor of Law, New York Univ. Law School); *cf.* Jacobs v. United States, 290 U.S. 13, 16–17 (1933) (constitutional requirement of just compensation held to require inclusion of interest in damage awards; "[T]he right to just compensation could not be taken away by statute or be qualified by the omission of a provision for interest where such an allowance was appropriate in order to make the compensation adequate."). *See generally* Sager, *The Supreme Court, 1980 Term--Foreword: Constitutional Limitations on Congress' Authority to Regulate the Jurisdiction of the Federal Courts*, 95 HARV. L. REV. 17 (1981); Tribe, *Jurisdictional Gerrymandering: Zoning Disfavored Rights Out of the Federal Courts*, 16 HARV. C.R.-C.L. L. REV. 129 (1981).

[75]In holding on policy grounds that municipalities ought not to be liable for punitive damages under 42 U.S.C. § 1983 (1976 & Supp. V 1981), the Court relied on the availability of punitive damages against the offending official. City of Newport v. Fact Concerts, Inc., 101 S. Ct. 2748, 2761 (1981). "In our view," the Court stated, "this provides sufficient protection against the prospect that a public official may commit recurrent constitutional violations by reason of his office." *Id.* But if Congress adopts a system of exclusive governmental liability, the availability of punitive damages against the offending official could not be relied upon to "directly advance[] the public's interest in preventing repeated constitutional deprivations." *Id.* The legislative alternative that would create partial exclusive government liability, leaving officials themselves liable for punitive damage, is unsatisfactory. This would virtually eliminate any benefits to be gained by removing officials from *Bivens*-type suits because punitive damages could be expected to be, and are in fact, pleaded as a matter of course.

[76]Carlson v. Green, 446 U.S. at 19. To be sure, H.R. 24, 97th Cong., 1st Sess. § 3 (1981), declares that the exclusive remedy against the United States "shall be deemed an equally effective substitute for any recovery against any employee of the United States for tort claims arising under the Constitution." Although the Supreme Court in

Allowing malice, the normal prerequisite for punitive damages, to be placed in issue at trial may entail some sacrifice of the litigation economies achieved by waiver of good-faith immunity. But such economies would not be lost completely because good faith would be eliminated as an issue at trial, thus simplifying the liability determination. If additional damages are available against the government, it might be desirable to litigate cases in two stages: the first stage would determine whether there is any liability and the second, to determine the issue of malice, would be necessary only if liability is established.[77] Such bifurcation could avoid problems that the prospect of punitive damages might otherwise generate. Whatever the merits of the proposal, there does not appear to be strong congressional support for any provision authorizing damages that could be construed as intended to punish or penalize the government.

D. *Other Issues*

1. *Attorney Fees.*

As noted above,[78] H.R. 595, but not S. 775, provides that a prevailing plaintiff would be entitled to reasonable attorney fees and costs.[79] The debate over this issue is familiar. Noting that attorney fees are not available to successful plaintiffs in nonconstitutional tort suits, some argue that the possibility of attorney fees only would encourage litigation and inspire "artful pleading" by lawyers to set forth constitutional rather than nonconstitutional tort claims whenever possible.[80] But the same considerations that support awarding attorney fees in cases where civil rights have been violated would seem to apply here

---

Carlson stated that a *Bivens*-type action *may* be defeated by such a declaration, 446 U.S. at 19–20, the validity of the declaration itself surely would be subject to judicial review.

[77]FED. R. CIV. P. 42. Although bifurcated trials in cases not involving third-party claims usually involve only liability and damages phases—distinguishing between the merits of the claim and the remedy—courts under some circumstances have bifurcated trials into two merits phases. *See* North Cent. Airlines v. Continental Oil Co., 574 F.2d 582, 587 (D.C. Cir. 1978); Berman v. Gerber Prods., 454 F. Supp. 1310, 1316 (W.D. Mich. 1978).

[78]*See supra* text accompanying notes 28 & 35.

[79]H.R. 595 § 202 (proposed 28 U.S.C. § 2698(a)).

[80]*1982 House Hearings*, *supra* note 11, at 341 (statement of J. Brian Hyland, President, Ass'n of Fed. Investigators); *see id.* at 284 (statement of Sec. and Exch. Comm'n); *id.* at 258, 261 (statement of William Taft IV, Gen. Counsel, Dep't of Defense).

as well.[81] In short, the attorney-fees issue has its own momentum and will be determined by forces beyond the relatively narrow context of constitutional torts.

2. *Former Officials.*

H.R. 595 and S. 775 appear to cover claims for acts or omissions committed by former government officials while in government service. But neither bill provides an administrative mechanism for bringing the former officials to account for their wrongful conduct. The need for such an administrative mechanism was recognized in a Carter administration bill.[82] Under the Carter bill, former officials could elect either to be sued individually after leaving government service, or to have the government substituted in their place. In choosing the latter, the former official would have agreed to submit to a disciplinary proceeding that could have resulted in a fine equal to as much as one-twelfth of the annual salary earned at the time the act or omission occurred.[83] Regardless of the merits of the approach embodied in the Carter administration bill, it seems anomalous to place beyond accountability officials who had chosen to leave government service before a determination of liability. Of course, offering some wrongdoers an incentive to leave office might not always be a totally unsatisfactory result.

3. *"Scope of Office" versus "Color of Law."*

An important issue presented by the Attorney General certification requirement under H.R. 595 and S. 775 is whether it is advisable to use a "scope of office or employment" requirement or a broader criterion of "color of law," thereby following the criterion of liability applicable to state and local officials under 42 U.S.C. § 1983.[84] The broader color of law criterion would

---

[81]*See* 42 U.S.C. § 1988 (1976 & Supp. V 1981); S. REP. No. 1011, 94th Cong., 2d Sess. 2, *reprinted in* 1976 U.S. CODE CONG. & AD. NEWS 5908, 5910:

> [C]ivil rights laws depend heavily upon private enforcement, and fee awards have proved an essential remedy if private citizens are to have a meaningful opportunity to vindicate the important Congressional policies which these laws contain.
>
> . . . If private citizens are to be able to assert their civil rights, and if those who violate the Nation's fundamental laws are not to proceed with impunity, then citizens must have the opportunity to recover what it costs them to vindicate these rights in court.

See also id. (noting that Congress, since 1870, has authorized such fee shifting under more than 50 laws).

[82]H.R. 2659, 96th Cong., 1st Sess. (1979).

[83]*Id.*; *see 1979 House Hearings*, *supra* note 2, at 81 (statement of William Webster, Director, FBI) (praising former employee provision as "fair and worthwhile").

[84]This was the approach embodied by the Carter administration bill, H.R. 2659; *see 1979 House Hearings*, *supra* note 2, at 8 (statement of Deputy Att'y Gen. Benjamin

increase the prospect of liability and thereby seems better as a matter of public policy. Such an approach

> is more consistent with the values of full compensation and loss-spreading for injuries somehow connected with governmental action. It would more effectively relieve the official of the prospect of personal litigation and liability, which is one of the chief purposes behind [the proposed legislation]. Finally, it would often avoid the necessity of premature threshold determinations on the question whether the employee was acting within the scope of his authority or solely under color of office. Both in terms of its relevance and its feasibility, this determination is best left to the disciplinary phase of the affair.[85]

At present, the political support for applying a color of law criterion in a federal context appears to be negligible. Before exposing the United States to liability that is vastly greater than that under the *Bivens* doctrine, prudence dictates a more limited experiment; liability of the United States should be established only when an official has committed a constitutional tort within the scope of his office or employment.[86]

4. *Constitutional versus Nonconstitutional Wrongs.*

H.R. 595 and S. 775 both distinguish between constitutional and nonconstitutional torts.[87] The distinction is significant insofar as the existing *Bivens* remedy for constitutional torts provides plaintiffs with rights not afforded plaintiffs in FTCA suits against the government, such as the right to jury trial and the right to punitive damages.[88] The availability of such "extra" rights has been understood to render the *Bivens* remedy "more effective" than the FTCA remedy.[89]

The problem posed by the distinction between constitutional and nonconstitutional torts is that in practice, factual differences

---

Civiletti), *1978 Senate Hearing*, *supra* note 17, at 39–40 (S. 2117), as well as by earlier legislative proposals. *See* S. REP. NO. 588, *supra* note 36, at 34; S. REP. NO. 755, *supra* note 2, at 337–38.

[85] *1979 House Hearings*, *supra* note 2, at 124 (remarks of George Bermann, Professor of Law, Columbia Univ. Law School).

[86] Under H.R. 595 and S. 775, an individual government official would remain personally liable for constitutional misconduct committed outside of the scope of his office or employment, but under color of law. The bills do not, and could not, purport to provide an "equally effective" substitute for the existing system of treating such misconduct. *Bivens*, 403 U.S. at 397.

[87] *See supra* note 2.

[88] *See* Carlson v. Green, 446 U.S. at 22; Casto, *Government Liability for Constitutional Torts: Proposals to Amend the Federal Tort Claims Act*, 49 TENN. L. REV. 201, 240–41 (1982) (differences between FTCA and *Bivens* suits charted).

[89] Carlson v. Green, 446 U.S. at 20–23.

between the two types of tort are sometimes difficult to discern. A single set of facts may give rise to a cause of action under both a constitutional and a nonconstitutional tort theory. In *Carlson v. Green*, for example, the Supreme Court acknowledged that the conduct at issue gave rise to claims under the Eighth Amendment and under the FTCA.[90] If legislation making the United States liable for constitutional torts preserves for plaintiffs who allege such torts the extra rights available to *Bivens* plaintiffs, while precluding FTCA plaintiffs from exercising such rights, much litigation may be devoted to the appropriate theory of liability in any given case.[91]

One way to avoid such litigation would be to abolish the distinction between constitutional and nonconstitutional torts—a distinction that has been criticized as unsound and unnecessary.[92] As a practical matter, this could be accomplished either by extending the extra *Bivens* rights to nonconstitutional suits, or by eliminating such extra rights altogether from constitutional tort suits against the government.[93]

In our view, the distinction between constitutional and nonconstitutional torts should be preserved. Although the facts of some cases may support claims of both constitutional and nonconstitutional injury, this will not be so in every case. The protections afforded by the Constitution and the common law are not coextensive. If they were, the Supreme Court has recognized, the *Bivens* remedy itself would never have been devised, for one who has suffered a constitutional injury always could have recovered in tort.[94] Conversely, the Court has made clear that not every common-law wrong rises to the level of constitutional injury.[95] Since the rights of plaintiffs claiming con-

---

[90]*Id.* at 20.

[91]*See* Public Citizen Litigation Group, Memorandum on Handling of Tort Claims Act Cases Under H.R. 2659 Where Event Gives Rise to Constitutional and Common-law Torts 3–5 (Oct. 19, 1979) (on file in office of HARV. J. ON LEGIS.).

[92]*See 1979 House Hearings*, *supra* note 2, at 123–24 (statement of George Bermann, Professor of Law, Columbia Univ. Law School).

[93]The elimination of such extra rights would not mean, however, the elimination of independent administrative action where constitutional injury had been demonstrated. Courts would still have to make the distinction between constitutional and nonconstitutional tort liability for the former triggers special disciplinary sanctions.

[94]*See Bivens*, 403 U.S. at 392 ("[T]he Fourth Amendment operates as a limitation upon the exercise of federal power regardless of whether the State in whose jurisdiction that power is exercised would prohibit or penalize the identical act if engaged in by a private citizen.").

[95]*See, e.g.*, Paul v. Davis, 424 U.S. 693 (1976); *cf.* Ingraham v. Wright, 430 U.S. 651, 677–82 (1977) (due process does not mandate hearing where common-law damages are available for deprivation of liberty).

stitutional injuries differ significantly from the rights of plaintiffs claiming nonconstitutional injuries, and since the distinction between the two types of claims is itself valid, the distinction should be retained.

To be sure, when the facts of a case only arguably support claims of both constitutional and nonconstitutional injuries, it sometimes will be difficult to decide whether claims of both types of injuries, or of only one, should be allowed. If Congress affords the extra rights now available to *Bivens* plaintiffs to those who allege constitutional injuries in actions against the United States, without extending such extra rights to plaintiffs alleging nonconstitutional wrongs under the FTCA, one can expect attempts to take advantage of these rights by "skillful counsel [who] plead the existence of a constitutional tort when their case in reality sounds in a traditional, common law cause of action."[96] That counsel can be expected to do the most for their clients, however, does not justify curtailing substantive rights and remedies. There is no reason of course, for Congress not to make clear in the legislative history its intention that courts be alert to the possibility of such skillful pleading, and to decide a plaintiff's entitlement to the special advantages of pleading constitutional torts "based upon the true gravamen of the tort as alleged and proven,"[97] and not simply upon the label used by counsel to characterize the wrong.

IV. Conclusion

Congress should amend the Federal Tort Claims Act to provide that the United States shall be substituted as the exclusive party defendant in all civil damage actions for violations of rights secured by the Constitution of the United States committed by federal officials while acting within the scope of their office or employment. In amending the FTCA, Congress must ensure that its statutory remedy for violations of constitutional rights by federal officials is no less effective than the remedy fashioned by the Supreme Court in *Bivens* and its progeny.

To meet that standard, a legislative remedy should provide, among other things, that:

---

[96]U.S. Dep't of Justice, Section-By-Section Analysis of S. 695, at 8 (1979) (on file in office of Harv. J. on Legis.).

[97]*Id.*

(1) In damage actions alleging constitutional violations, the United States may not assert as a defense any qualified immunity available to the official whose conduct gave rise to the claim, or his reasonable good-faith belief in the lawfulness of his conduct. Good-faith immunity of officials has been judicially created for policy reasons to ensure that officials exposed to personal liability shall not be discouraged from actively discharging their responsibilities. The policy basis for qualified immunity disappears when the government is substituted as the defendant. Allowing the United States to assert the official's qualified immunity in such circumstances would unjustifiably deny compensation to many victims of constitutional torts. It also would afford the government an advantage enjoyed neither by private nor public employers at common law, nor by the United States itself in nonconstitutional tort suits under the FTCA.

(2) An alternative system of administrative sanctions should be available to perform the deterrent and corrective functions now intended to be served by the damage remedy available in lawsuits against individual officials. At a minimum, Congress should provide every federal agency with explicit statutory authority to employ existing administrative controls and sanctions on officials who have violated the constitutional rights of any person. Congress should consider seriously the need for an independent administrative disciplinary mechanism as well as other measures believed necessary to strengthen administrative methods of dealing with official misconduct.

Legislation incorporating these two elements would go far toward satisfying the twin goals of ensuring adequate compensation for those whose constitutional rights have been violated, and deterring future violations of such rights by holding federal officials accountable for their misconduct. It is in large measure because the *Bivens* remedy has not satisfactorily served these goals that a legislative alternative to that remedy is necessary.

A decision by Congress to afford plaintiffs in constitutional tort suits in which the United States is the defendant the extra rights now available to *Bivens* plaintiffs—the right to jury trial and the right to punitive or additional damages—undoubtedly would be considered by the courts in determining whether the legislative remedy fashioned by Congress is an adequate substitute for those currently provided by *Bivens*. Whether attorney fees should be provided successful plaintiffs under the new

legislation, whether former officials, like sitting officials, should be made subject to administrative sanctions in exchange for being released from *Bivens* liability, whether the new legislation should reach action taken not only within the scope of an employee's office or employment, but under color of law, and whether the new legislation should preserve the distinction between constitutional and nonconstitutional torts—are issues that Congress must face in fashioning legislation that will replace the *Bivens* remedy. The Ninety-eighth Congress should meet the challenge of making the compromises necessary to enact such long overdue legislation.

ARTICLE

THE APPLICABILITY OF THE EQUAL TIME DOCTRINE AND THE REASONABLE ACCESS RULE TO ELECTIONS IN THE NEW MEDIA ERA

ROBERT S. KOPPEL*

The equal time doctrine and the reasonable access rule are the most prominent federal statutory provisions regulating the appearance of political candidates on broadcast television. These provisions require broadcasters who allow one candidate a use of their station to provide all other candidates with an equal opportunity, and require broadcasters to allow federal candidates to purchase broadcast time on their stations. The courts have held similar regulation of the print media unconstitutional, but have upheld the equal time and reasonable access requirements as necessary to ensure fair and diverse presentation of political candidates on the limited number of broadcast media outlets. Now, however, the emergence of new electronic media, such as cable television and electronic publishing, has rendered uncertain whether an extension of the current regulatory scheme is either desirable or constitutionally permissable.

In this Article, Mr. Koppel proposes a framework for analyzing the problem of regulating communications by political candidates in the new media era. He first describes the goals that such regulation must strive to achieve, and explores the rationales that the courts have used in applying the equal time and reasonable access requirements to broadcasting. His framework generates and analyzes regulatory policy options for each new medium individually as well as for the communications media as a whole. Finally, he proposes a three-step sequence of policy options that gradually reduces the overall level of media regulation, culminating in complete deregulation.

The equal time doctrine and the reasonable access rule have become important features of the American political landscape because they represent significant governmental regulation of content in the broadcast media. The equal time doctrine, section 315(a) of the Communications Act of 1934, provides that any television or radio broadcaster who permits a candidate for any public office to use his station must provide all other candidates

---

*A.B., Cornell University, 1979; J.D., Harvard Law School, 1982. Mr. Koppel is an attorney practicing communications law in Washington, D.C. This Article was written with the Harvard University Program on Information Resources Policy.

for the same office an equal opportunity to use the station.[1] The reasonable access rule, section 312(a)(7) of the Communications Act of 1934, requires that broadcasters permit federal candidates to purchase reasonable amounts of time for broadcasting on their facilities.[2]

The advent and rapid expansion of various new electronic communications media—such as cable television, direct broadcast satellite (DBS), multipoint distribution service (MDS), teletext, videotex, videocassettes, videodiscs, subscription television (STV), high definition television (HDTV), and low power television (LPTV)—have raised the question of whether the equal time doctrine and reasonable access rule should be applied to each of these new electronic media. The proliferation of these new delivery systems also has raised questions regarding the continued desirability of applying these statutes to broadcast television. In addition, the general trend toward deregulation and the heightened concern over governmental interference with the media have focused attention on the current regulatory scheme. In September 1981, the Federal Communications Commission (FCC) recommended that Congress abolish the equal time doctrine and the reasonable access rule.[3]

This Article establishes an analytical framework for determining the applicability of the equal time doctrine and the reasonable access rule to each of the new electronic media. This framework is also used to determine whether the equal time doctrine and the reasonable access rule should be retained for broadcast television. The remainder of the Article is devoted to discussing specific policy options for the future regulation of

---

[1]Enacted in the Communications Act of 1934, Pub. L. No. 73-416, § 315, 48 Stat. 1064, 1088 (current version at 47 U.S.C. § 315(a) (1976)). The equal time doctrine provides:

> If any licensee shall permit any person who is a legally qualified candidate for any public office to use a broadcasting station, he shall afford equal opportunities to all other such candidates for that office in the use of such broadcasting station . . . . No obligation is imposed under this subsection upon any licensee to allow the use of its station by any such candidate.

[2]Added to the Communications Act of 1934 by the Federal Election Campaign Act of 1971, Pub. L. No. 92-225, § 104, 86 Stat. 3, 7 (1972) (current version at 47 U.S.C. § 312(a)(7) (1976)). It provides:

> (a) The Commission may revoke any station license or construction permit . . .
>
> . . . .
>
> (7) for willful or repeated failure to allow reasonable access to or to permit purchase of reasonable amounts of time for the use of a broadcasting station by a legally qualified candidate for Federal elective office on behalf of his candidacy.

[3]*See* BROADCASTING, Sept. 21, 1981, at 23.

electoral communications.[4] While there is no requirement of special solicitude for political candidates in establishing these policies, a long tradition of special rules for political candidates suggests that governmental regulation of electoral communications should not be lumped too readily with the broader question of general content regulation in the new media.

Section I sets out the underlying goals of media regulation. Any regulatory scheme must be designed to maximize these goals, which include minimum governmental interference, diversity, access, and fairness. Section II describes the three models currently used to achieve these goals: the print model, the common carrier model, and the broadcast model, which includes the equal time doctrine and the reasonable access rule. These three models, and combinations thereof, constitute the universe of policy options available for the regulation of electoral communications. Section III outlines the rationales or justifications for regulating broadcasting and defines the basic constitutional constraints involved in regulating electoral communications on any medium.

Section IV proposes two basic approaches to regulating electoral communications. The micro approach selects a regulatory scheme for each electronic medium separately, paying little attention to electronic communications as a whole. Current media regulation seems to have developed under this approach. The macro approach attempts to regulate the communications media comprehensively, taking account of how the individual electronic media could work together in a unified system of electoral communications. Section IV also analyzes various micro schemes of regulation as applied to cable television and to several of the new media. Section V analyzes several macro schemes as they would apply comprehensively to all media.

Finally, Section VI recommends a three-step macro policy option for the future regulation of electoral communications in the new media. The proposed policy options are not limited by current statutory provisions. Although this Article considers the structure and justification of current laws, its major focus is on establishing a new legal structure for the future.

---

[4]This Article uses the term "electoral communications" to denote the speech regulated by §§ 315(a) and 312(a)(7). These statutes apply only to political candidates during political campaigns.

I. GOALS

The starting point of any policy analysis must be to posit the goals or ends that the policy should achieve.[5] From the First Amendment,[6] courts and commentators have drawn four overlapping and frequently conflicting goals of regulating electoral communications. These four goals underlie the basic constitutional rationales for the regulation of electoral communications.

A. *Minimum governmental interference and maximum journalistic discretion.* Although the First Amendment has never been held to constitute an absolute prohibition on regulation,[7] it is clearly the foundation of a long and powerful tradition of minimum governmental interference with the press. Maximum journalistic discretion through private control of the communications media is merely the mirror image of minimum governmental interference.

B. *Diversity.* One of the purposes of free speech and a free press is to assure "the widest possible dissemination of information from diverse and antagonistic sources."[8] This fulfills "the public need for information and education with respect to the significant issues of the times"[9] and thereby leads to an informed and participating electorate. Dissemination of information from diverse sources also helps to assure that there will be a "market-place of ideas,"[10] which will help prevent control of public opinion and of political elections by a few powerful media operators.[11]

C. *Access.* Courts and commentators discuss two types of access, but do not distinguish them. On the one hand, there is the right of access to the communications media to *send* messages. On the other hand, there is the right of access to *receive* messages, in other words, "the right of the public to receive suitable access to social, political, esthetic, moral, and other ideas."[12] Access to receive messages transmitted over the com-

---

[5]These four goals of regulation are discussed in numerous First Amendment cases and by various commentators. *See infra* notes 15–40 and accompanying text.

[6]U.S. CONST. amend. I ("Congress shall make no law . . . abridging the freedom of speech, or of the press").

[7]*See, e.g.*, Red Lion Broadcasting Co. v. FCC, 395 U.S. 367 (1969).

[8]Associated Press v. United States, 326 U.S. 1, 20 (1945).

[9]Thornhill v. Alabama, 310 U.S. 88, 102 (1940).

[10]*See* Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

[11]Barrow, *The Equal Opportunities and Fairness Doctrines in Broadcasting: Should They Be Retained?*, 1 COMM/ENT 65, 66 (1977).

[12]Red Lion Broadcasting Co. v. FCC, 395 U.S. at 390.

munications media is essentially diversity. Access to send messages, what shall be meant herein as access, is a means of achieving diversity, but also is a goal in itself.[13] Access is more meaningful where the channels of communication are widely available and reasonably affordable.

D. *Fairness.* Fairness means that where one candidate is allowed to speak or one idea is expressed, opposing candidates or ideas should be heard, thereby providing some balance and at least a limited marketplace of ideas. Fairness is a more limited goal than diversity or access because fairness dictates that a candidate be heard only if other candidates appear on the medium.[14] Diversity and access, on the other hand, suggest that a candidate should be heard regardless of whether other candidates are using the media.

These four goals often conflict, and must be balanced against one another. The fundamental question whenever the government seeks to impose content regulation, such as the equal time doctrine, is whether the right of the public to receive information from diverse sources outweighs the right of the editors of the communications media to exercise unrestrained journalistic discretion. This balancing and weighing of the various goals are played out in the contrasting models of regulation.

II. Models of Regulation

The government presently uses three models for regulating electoral communications: the print model, the common carrier model, and the broadcast model. Each of these regulatory models is closely related to the characteristics of the media to which it has been applied. These three models, along with hybrid schemes, represent the universe of policy options available to policymakers, who must choose for each new electronic medium the model best suited to achieving the goals of regulation.

---

[13]For example, access can be seen as a goal in itself in developing a respect for order, and in encouraging dissident candidates to work within the system. If dissident candidates are guaranteed access to the communications media, they will be less likely to "conclude that unrest and violence alone will suffice to capture public attention." Barron, *Access to the Press--A New First Amendment Right*, 80 Harv. L. Rev. 1641, 1650 (1967); *see* Whitney v. California, 274 U.S. 357, 375 (1927) (Brandeis, J., concurring) ("the path of safety lies in the opportunity to discuss freely supposed grievances and proposed remedies").

[14]The equal time doctrine serves primarily a fairness goal.

A. *The print model.* The print model, in its purest form, countenances neither content nor structural regulation;[15] "[f]or print the First Amendment stands behind the idea that no government, either federal or local, has authority to interfere with the right of a publisher to print whatever information that person cares to put on paper."[16] Electoral communications are completely unregulated in the print media; the goal of maximum journalistic discretion prevails over all other goals.

In 1974, the Supreme Court held in the landmark case of *Miami Herald Publishing Co. v. Tornillo*[17] that the First Amendment was violated by a Florida statute granting political candidates a right to reply to newspaper editorials criticizing them.[18] The Court struck down the statute, declaring:

> The choice of materials to go into a newspaper, . . . and treatment of public issues and public officials—whether fair or unfair—constitute the exercise of editorial control and judgment. It has yet to be demonstrated how governmental regulation of this crucial process can be exercised consistent with the First Amendment guarantees of a free press . . . .[19]

The Court did not once refer to an earlier opinion upholding similar regulations for broadcast television.[20] It was unwilling to accord special access rights to the print media for political candidates, or "to balance the need of the public to know with the editorial control and judgment of the editors."[21] The Court may have believed that print, unlike broadcasting, is inherently accessible to all without need for governmental intervention.

B. *The common carrier model (structural regulation).* The common carrier model is a type of structural regulation; unlike the two content regulation models, it is not specifically related

---

[15]The print media are not completely free from regulation. Antitrust, libel, and copyright laws, as well as general business regulations, do apply. *See* Schmidt, *Pluralistic Programming and Regulation of Mass Communications Media*, in COMMUNICATIONS FOR TOMORROW: POLICY PERSPECTIVES FOR THE 1980s 191–92 (G. Robinson ed. 1978). In addition, unprotected communications such as obscenity are prohibited.

[16]W. Read, The First Amendment Meets the Second Revolution 2 (Jan. 1981) (Program on Information Resources Pol'y, Harv. Univ. Center for Information Pol'y Research, paper no. P-81-1).

[17]418 U.S. 241 (1974).

[18]The Florida Supreme Court, in upholding the statute, had determined that it furthered the "broad societal interest in the free flow of information to the public." *Id.* at 245.

[19]*Id.* at 258.

[20]Red Lion Broadcasting v. FCC, 395 U.S. 367 (1969).

[21]Barrow, *supra* note 11, at 73.

to electoral communications.[22] Structural regulation controls access to transmission facilities, not the information itself. Because it achieves access, diversity, and minimum governmental interference with content, structural regulation poses a reduced threat to First Amendment values.

The distinguishing feature of a common carrier is the separation of the conduit from the content; the operator of the transmission facilities does not provide any of the information transmitted. Under this "separations" policy, the common carrier holds out his transmission facilities to the public on a first-come, first-served basis, and "members of the public who choose to employ such facilities may communicate or transmit intelligence of their own design and choosing."[23]

The common carrier model, used to regulate natural monopolies such as local telephone lines, guarantees that the monopolist's services will be available to all customers at a reasonable price. The Eighth Circuit has noted that the "basic rationale for regulation of common carriers [is] to insure fair and equal access to the carrier's service."[24] Telephone lines for campaigning and for fundraising are equally available to all candidates willing to pay the tariffs. No candidates are barred because of their views, and the telephone company may not edit their messages. Thus, this model achieves access while minimizing governmental interference with content.

Access channels are a form of structural regulation that may be especially important in generating policy options for the new media.[25] Under an access channel scheme, the government would require media owners to reserve certain channels for public access either free of charge or for reasonable rates. During political campaigns, an entire channel might be devoted to access for political candidates. Governmental regulation would

---

[22]Although the reasonable access rule may appear to be a structural regulation, it is better classified as a content regulation because the access is contingent on the content of the message, which must be a political candidate's message "on behalf of his candidacy." 47 U.S.C. § 312(a)(7) (1976).

[23]FCC v. Midwest Video Corp., 440 U.S. 689, 701 (1979) (citing Industrial Radiolocation Serv., 5 F.C.C.2d 197, 202 (1966)).

[24]Midwest Video Corp. v. FCC, 571 F.2d 1025, 1036 (8th Cir. 1978) ("Neither the basic rationale for regulation of common carriers (to insure fair and equal access to the carrier's service) nor that for regulation of broadcast transmissions (to preclude bedlam on broadcast frequencies), is applicable to cable systems *per se*."), *aff'd*, 440 U.S. 689 (1979).

[25]*See infra* text accompanying note 88.

be limited to requiring the access channels and perhaps to regulating their rates.

C. *The broadcast model.* The broadcast model for the regulation of electoral communications consists of the equal time doctrine and the reasonable access rule. When a broadcaster allows one candidate a nonexempt "use" of his station, the equal time doctrine requires not just equal time, but equal treatment in all respects for other candidates.[26] If a candidate has paid for his time, the broadcaster need only offer an opportunity to purchase paid time to all opposing candidates; he has no duty to offer free time unless he has already allowed one candidate to use his station free of charge. The equal time doctrine thus serves primarily to promote fairness.

The reasonable access rule for federal candidates is a congressional attempt to move beyond the fairness goal of the equal time doctrine to achieve access as well. A broadcaster now must allow a federal candidate to buy time on his station, and, by virtue of the equal time doctrine, must provide all opposing candidates the same opportunity. As always, this access may be limited by the candidates' inability to pay for broadcast time.

In *Red Lion Broadcasting Co. v. FCC*,[27] the Supreme Court upheld the FCC's statutory authority to adopt personal attack and political editorial rules.[28] It also upheld, in dictum, the constitutionality of the equal time doctrine, declaring:

> It is the right of the viewers and listeners, not the right of the broadcasters, which is paramount . . . .
>
> . . . .
>
> . . . Otherwise, station owners and a few networks would have unfettered power to make time available only to the highest bidders, to communicate only their own views on public issues, people, and candidates, and to permit on the

---

[26]47 U.S.C. § 315(a) (1976). Certain appearances by candidates are exempt from the equal time doctrine, including appearances on newscasts, news interview programs, news documentaries, and on-the-spot coverage of news events. *Id.* § 315(a)(1)–(4). Debates and press conferences may also qualify as on-the-spot news coverage when they are outside the control of the broadcaster. Aspen Inst., 55 F.C.C.2d 697 (1975), *aff'd sub nom.* Chisholm v. FCC, 538 F.2d 349 (D.C. Cir.), *cert. denied*, 429 U.S. 890 (1976).

[27]395 U.S. 367 (1969). Justice White wrote for an 8–0 majority.

[28]*Id.* at 391. The personal attack rules require that if an individual or group is personally attacked during the presentation of a controversial issue of public importance, the broadcaster must notify the attacked party and offer that party a reasonable opportunity to respond. 47 C.F.R. § 73.1920 (1981). The political editorial rules require that if a broadcaster in an editorial endorses or opposes a candidate, the disfavored candidate(s) must be notified and offered a reasonable opportunity to respond. *Id.* § 73.1930.

> air only those with whom they agreed. There is no sanctuary in the First Amendment for unlimited private censorship operating in a medium not open to all.[29]

In *Red Lion*, the Supreme Court seemed unconcerned with the goal of minimum governmental intervention, focusing instead on fairness and access.

In contrast, the Court in *Columbia Broadcasting System v. Democratic National Committee* (*CBS v. DNC*)[30] held that broadcasters could not be compelled to sell advertising time to individuals or groups, including political parties, who wish to comment on public issues. Treating broadcasters as part of the press for the first time,[31] the Court expressed concern that there would be "a further erosion of the journalistic discretion of broadcasters in the coverage of public issues," and noted that "[f]or better or worse, editing is what editors are for; and editing is selection and choice of material."[32] The Court left open the possibility that in the future Congress or the FCC might "devise some kind of limited right of access that is both practicable and desirable."[33]

In *CBS v. FCC* (*Carter-Mondale*),[34] however, the Supreme Court upheld the reasonable access rule, finding that the need to assure access outweighed the need to preserve journalistic discretion. The Court emphasized that:

> The First Amendment interests of candidates and voters, as well as broadcasters, are implicated by § 312(a)(7) . . . . Section 312(a)(7) thus makes a significant contribution to freedom of expression by enhancing the ability of candidates to present, and the public to receive, information necessary for the effective operation of the democratic process.[35]

The reasonable access rule is a significant regulation of content; the broadcaster retains no control over programming content for broadcast time turned over to a candidate under section

---

[29]395 U.S. at 390–92.

[30]412 U.S. 94 (1973) (Burger, C.J.).

[31]*See* Bollinger, *Freedom of the Press and Public Access: Toward a Theory of Partial Regulation of the Mass Media*, 75 MICH. L. REV. 1, 26 n.81 (1981).

[32]CBS v. DNC, 412 U.S. at 124.

[33]*Id.* at 131.

[34]CBS v. FCC, 453 U.S. 367 (1981). This case arose out of the attempt by the Carter-Mondale Presidential Committee to purchase time for a 30-minute program featuring President Carter, to be aired in December 1979 in conjunction with Carter's formal announcement of his candidacy for reelection. All three networks refused to make the requested time available.

[35]*Id.* at 382.

312(a)(7). Whereas in *CBS v. DNC* broadcasters had for the first time been referred to as a part of the "press," in *Carter-Mondale* broadcasters had once again become merely broadcasters or licensees.[36]

Carter-Mondale demonstrates the special treatment that the Supreme Court is willing to give political candidates seeking access to broadcast television. The Court noted that congressional legislation reflects "the importance attached to the use of the public airwaves by political candidates."[37] While many of the content regulations imposed by the FCC arise out of the broad public interest standard of the Communications Act of 1934,[38] both the equal time doctrine and the reasonable access rule reflect specific policies regarding political campaigns. This is reflected in sections 315(a) and 312(a)(7), explicit statutory provisions that grant political candidates special rights of access to broadcast television for personal appearances. A comparison of *Carter-Mondale* with *CBS v. DNC* also suggests that the Supreme Court is more likely to uphold access rules for political candidates than for noncandidates such as political parties.

In addition to content regulations, certain structural regulations have been applied to broadcasting.[39] The two primary forms of structural regulation are restrictions on cross-ownership and on multiple ownership.[40] For electoral communications, the cross-ownership rules are most interesting. These rules prevent persons with interests in broadcast stations from owning any cable system or any newspaper in the same community, with some exceptions. One purpose of these rules is to promote diversity by preventing control of all the media outlets in a single community from being concentrated in the hands of a few media owners. These rules also assure greater access and balance,

---

[36]*See* Bollinger, *supra* note 31, at 26 n.81.

[37]*Carter-Mondale*, 453 U.S. at 386.

[38]*See* 47 U.S.C. §§ 303(g), 307(a), 309(a) (1976).

[39]Structural and content regulation need not be mutually exclusive; they can be used together to achieve the posited goals. Many of the policy options proposed below rely in part on structural regulation.

[40]The cross-ownership rules for broadcast and cable may be found at 47 C.F.R. § 76.501 (1981). The other broadcast multiple ownership and cross-ownership rules may be found at: *id.* § 73.35 (AM radio), *id.* § 73.240 (FM radio), and *id.* § 73.636 (television). The antitrust laws also serve as a form of structural regulation applicable to all communications media. They encourage diversity by promoting competition and discouraging monopolies, and it is possible that they may be invoked directly to gain access. *See* Neustadt, *Regulating Abundance: Policy Issues Raised by Electronic Publishing*, TELECOMMUNICATIONS, Sept. 1981, at 54.

because no single entity can arbitrarily deny access to a candidate, and a candidate excoriated by one media outlet may be praised by another.

The success of the broadcast model has been questioned, which is important to remember in analyzing various regulatory schemes. The most frequent criticism is that it has a chilling effect because the equal opportunities requirement discourages stations from airing political programming for fear that they will then have to provide equal time to all the minor party candidates free of charge.

III. Rationales for Regulation of Broadcasting

Three rationales have been used to justify imposing content regulation on broadcasting. These rationales also provide the framework for discussing whether various types of regulation would be constitutional if applied to the new media.

A. *Scarcity.* Governmental regulation of broadcasting was instituted because of the technological scarcity of radio spectrum space. Without government allocation of scarce frequencies, "the medium would be of little use because of the cacophony of competing voices, none of which could be clearly and predictably heard."[41] As the Supreme Court declared in *National Broadcasting Co. v. United States*:

> Freedom of utterance is abridged to many who wish to use the limited facilities of radio. Unlike other modes of expression, radio inherently is not available to all. That is its unique characteristic, and that is why, unlike other modes of expression, it is subject to governmental regulation.[42]

In *Red Lion*, the Court relied primarily on the scarcity rationale to uphold the fairness and equal time doctrines, observing that "[w]here there are substantially more individuals who want to broadcast than there are frequencies to allocate, it is idle to posit an unabridgeable First Amendment right to broadcast com-

---

[41] *Red Lion*, 395 U.S. at 376.

[42] 319 U.S. 190, 226 (1943). Frankfurter's reasoning in this case can be criticized on the ground that, while scarcity may justify orderly allocation of frequencies, it does not justify content regulation. Today, scarcity of frequencies, to the extent there is scarcity, may also be a function of economic as well as technological constraints. For example, bandwidth compression could increase the number of available broadcast frequencies, but it has been resisted because both broadcasters and viewers would be faced with the added costs of new equipment.

parable to the right of every individual to speak, write, or publish."[43] Governmental intervention is considered an aid to free speech because it ensures that both diversity and access are not severely hampered by the scarcity of broadcast frequencies.[44]

On the other hand, the Supreme Court has refused to accept a technological scarcity rationale for regulating the print media. One reason for this refusal is that newspaper "is technologically open to all, [while broadcasting] is . . . not."[45] The Court has also refused to accept *economic* scarcity as a rationale for content regulation of print.[46] Thus, in single-newspaper towns, a newspaper's monopoly of print outlets could result in a complete lack of access for some political candidates.

Even with an abundance of media outlets from new technologies, continued regulation of broadcast television may be justified if a few national broadcast networks continue to dominate viewer preferences in the future.[47] If traditional broadcast television were a unique medium, the scarcity rationale could justify continuing to regulate its availability for electoral communications. The scarcity, however, would be of "viewer preferences,"[48] not of spectrum frequencies. Moreover, using viewer preference scarcity to justify regulating electoral communications would demonstrate a strong degree of paternalism.[49] If political programming is available but viewers choose not to watch it, it may not be appropriate for the government to try to force them to watch it.

B. *Public domain and public trustee.* The public domain and public trustee rationales for the regulation of broadcasting rest upon the idea that, like navigable waterways, the radio spectrum is a public resource whose ownership is vested in the American

---

[43]395 U.S. at 388.

[44]*See* Bollinger, *supra* note 31, at 6.

[45]*Electronic Journalism and First Amendment Problems, Recommendations of Communications Law Committee Section on Science and Technology, American Bar Association*, 29 FED. COM. B.J. 1, 10 n.11 (1976) [hereinafter cited as *ABA Recommendations*].

[46]Miami Herald Publishing Co. v. Tornillo, 418 U.S. 241. Bollinger has criticized the Supreme Court, suggesting that, for access regulation, whether the cause of concentration is technological or economic is "far less relevant from a first amendment standpoint than the fact of concentration itself." Bollinger, *supra* note 31, at 11.

[47]Such dominance might continue, despite a broader range of programming options, if viewers retain their taste for popular mass programming. *See* Schmidt, *supra* note 15, at 215.

[48]*Id.*

[49]*Id.* at 215–16.

people.[50] Although the government could have retained full control of the radio spectrum or decreed that each frequency be shared by many,[51] it has chosen instead to allocate portions of the radio spectrum for exclusive use by private companies.[52] This specific allocation to private broadcasting is partly "because of the contribution that [broadcasting] can make to an informed electorate."[53]

Private broadcasters thus are public trustees of their allotted frequencies. The ability to broadcast is a privilege granted by the government only to those who will broadcast in the public interest.[54] Broadcasters, unlike newspaper publishers, must be licensed by the federal government. The entire licensing and license-renewal process is designed to assure that broadcasters serve the public interest.[55] An important element of this public service is to aid in developing an informed electorate, which is at least partly achieved by assuring political candidates some opportunity to present themselves and their platforms to the electorate.[56]

A corollary to the public domain rationale is what might be called a quid pro quo theory: in exchange for the free use of a valuable public resource, the broadcaster is expected to use some of his supranormal profits to produce or air programming that serves the public interest.[57] Since newspapers, by contrast, do not use a public resource, and do not enjoy a protected market position, they are not considered public trustees required to serve the public interest.[58]

---

[50]*See, e.g.*, CBS v. DNC, 412 U.S. at 173–74; Schmidt, *supra* note 15. Although this Article has separated the closely connected public trustee and scarcity rationales, the courts and other commentators rarely do. *See, e.g.*, *Red Lion*, 395 U.S. at 394 (broadcasters are "licensees given the privilege of using scarce radio frequencies as proxies for the entire community").

[51]*Red Lion*, 395 U.S. at 390–91.

[52]*See* W. Read, *supra* note 16, at 22 n.21.

[53]*ABA Recommendations*, *supra* note 45, at 5.

[54]*See* 47 U.S.C. §§ 307(a), 309(a) (1976) (public interest standard of Communications Act of 1934).

[55]*See ABA Recommendations*, *supra* note 45, at 11.

[56]*See id.* at 5.

[57]W. Read, *supra* note 16, at 5. The quid pro quo theory operates with special force when the FCC bars a new entry into a broadcast market, even when such an entry is technologically feasible, in order to protect the profitability of the existing stations. This theory also assumes that public interest programming cannot also be profit-maximizing programming, but the popularity of local news programming may refute this assumption.

[58]Newspapers do in fact receive certain government benefits, most notably subsidized mail rates. These benefits could be seen as derived in part from newspapers' public information function.

C. *Impact and invasiveness.* Television's impact and invasiveness are sometimes used to justify content regulation. This impact and invasiveness has at least three closely related elements. First, television is invasive because its messages can be absorbed by the viewer or listener with little or no effort.[59] Broadcast messages, unlike written messages, cannot be avoided merely by "averting the eyes."[60] Second, broadcasting is considered invasive because it comes into the home so easily. Finally, the visual impact of the television screen itself may raise concern. The special impact of television is not only on the individual, but on society in general. Ninety-seven percent of United States households have televisions; these televisions are turned on an average of seven hours per day, and have become the primary news source for two-thirds of adult Americans.[61]

The invasiveness of broadcast television was used to justify the banning of cigarette commercials on television in *Banzhaf v. FCC*:

> Written messages are not communicated unless they are read, and reading requires an affirmative act. Broadcast messages, in contrast, are "in the air." . . . [A]n ordinary habitual television watcher can *avoid* these commercials only by frequently leaving the room, changing the channel, or doing some other such affirmative act. It is difficult to calculate the subliminal impact of this pervasive propaganda, which may be heard even if not listened to, but it may reasonably be thought greater than the impact of the written word.[62]

The Supreme Court also has relied upon the invasiveness of broadcast radio to justify the FCC's strict regulation of indecency on the broadcast media. In *FCC v. Pacifica Foundation*,[63] the Court upheld the FCC's authority to impose sanctions on licensees who engage in obscene, indecent, or profane broadcasting. The Court noted that "of all forms of communication, it is broadcasting that has received the most limited First Amendment protection."[64] One reason for that limited protec-

---

[59]*See* Bollinger, *supra* note 31, at 14; Schmidt, *supra* note 15, at 203–04.
[60]Cohen v. California, 403 U.S. 15, 21 (1971).
[61]*Schmidt*, *supra* note 15, at 219.
[62]405 F.2d 1082, 1100–01 (D.C. Cir. 1968), *cert. denied*, 396 U.S. 842 (1969).
[63]438 U.S. 726 (1978).
[64]*Id.* at 748.

tion is that "the broadcast media have established a uniquely pervasive presence in the lives of all Americans."[65]

Television's impact may be particularly important in political campaigns, where the candidate can figuratively step right into the viewer's home to deliver his message personally. Because of this face-to-face visual impact, television may be the most important and effective means of reaching the electorate, particularly in federal election campaigns.

Despite the impact and invasiveness of television, these rationales have never been used to justify the regulation of electoral communications. It is nevertheless possible that courts will rely on the impact rationale to justify future content regulation of this powerful medium after the scarcity rationale has become inapplicable.

There have been serious criticisms, however, of using an impact rationale as justification for content regulation. First, many commentators have questioned the assumption that television has a special impact on its viewers.[66] Any special impact on opinion and attitude formation may be merely the result of the dominance of the three networks; programming pluralism could alleviate this problem.[67] Second, regulating broadcasting specifically because of its wide audiences and great impact seems, as one commentator has put it, "inconsistent with the underlying purpose of the first amendment, which presumably is to protect effective as well as ineffective speech."[68] Another commentator has noted that "the more powerful a medium, the better it can serve to check governmental abuses and the greater the risks of government control."[69] These criticisms, however, do not seem to apply with full force if content regulation is limited to giving special treatment to political candidates. Unlike obscenity and indecency regulations, the equal time doctrine and reasonable access rule do not in themselves restrict the broadcaster's right to say what he pleases. Therefore, impact and invasiveness may become important rationales for future regulation of both broadcast television and the new media.

---

[65]*Id.*

[66]*See, e.g.*, Bollinger, *supra* note 31, at 15; Schmidt, *supra* note 15, at 219; W. Read, *supra* note 16, at 7.

[67]*See* Schmidt, *supra* note 15, at 220.

[68]Bollinger, *supra* note 31, at 15; *see infra* text accompanying notes 132–37.

[69]Schmidt, *supra* note 15, at 220.

IV. Policy Approaches for Regulating Electoral Communications in the New Media

All of the policy options proposed herein for regulating electoral communications in the new media fit into one of two underlying approaches to regulation: the micro approach and the macro approach. The micro approach analyzes each medium in isolation, as if it were a unique form of communication. The macro approach regulates the communications media comprehensively, seeking to take advantage of each medium's unique characteristics. This approach seeks to make the various media work together to maximize achievement of the four posited goals, rather than trying to achieve these goals through individual, uncoordinated regulation of each medium. Since these two approaches underlie all the proposed schemes of regulation, policymakers should decide at the outset whether to adopt the micro approach or the macro approach.[70]

Under the micro approach, policymakers analyze each new medium separately to identify which of its characteristics are relevant to electoral communications. Based on these characteristics, they then choose the regulatory model that is most likely to achieve the posited goals when it is applied to the new medium. This method of policy selection can be difficult, however, because it requires policymakers to set priorities among the goals.

Another means of selecting policy under the micro approach is simply to adopt the regulatory scheme currently applied to the existing medium that most resembles the new medium.[71] For example, if the new medium most resembles traditional broadcast television, then the broadcast model would be applied. This functional approach can be difficult to apply because many of the new electronic media exhibit regulatory characteristics of more than one existing medium. On the other hand, this method not only avoids the difficulty of setting priorities among the posited goals, but it is also more likely to lead to constitutionally acceptable regulation. Current regulations that are constitutional

---

[70]As one commentator has noted, "[t]he ultimate legal issue . . . is whether each form of communication should be judicially considered in isolation or as part of a collection of competing forms of communication." W. Read, *supra* note 16, at ii.

[71]However, selecting policy in this manner rests on the assumption that the media are appropriately regulated under the current scheme.

as applied to existing media may well be constitutional if applied to an analogous new medium.[72]

Another method of micro policy selection is to create a hybrid regulatory scheme. If a new medium shares characteristics with two or more existing media, appropriate parts of existing regulatory models could be combined to regulate the new medium. For example, if the new medium partly resembles broadcast television and partly resembles telephone service, a mixture of broadcast and common carrier regulations might be designed and applied to the new medium.

A. *Micro Policy Options for Regulating Electoral Communications on Cable Television*

There are a variety of policy options available for regulating electoral communications on cable television.[73] Under present law, the equal time doctrine applies to cablecast programming. The Federal Election Campaign Act of 1971[74] amended the 1934 Act to make section 315 specifically applicable to cable television by including "community antenna television system" within the definition of a broadcasting station.[75] On the other hand, it is completely uncertain whether the reasonable access rule of section 312(a)(7) now applies to cable television. The FCC has not included a reasonable access rule in its cable rules, nor has it ever tried to enforce such an obligation.[76]

---

[72]For example, if technological scarcity justifies content regulation of broadcast television, it may justify similar regulation of any other technologically scarce medium.

[73]This Article gives special emphasis to cable television because cable is clearly the most widespread and probably the most important new medium. Cable is also unique because it does not rely substantially on over-the-air transmission. *See* FCC Staff, Policies for Regulation of Direct Broadcast Satellites 19 n.32 (Oct. 2, 1980).

[74]Pub. L. No. 92-225, § 104, 86 Stat. 3 (1972). This Act dealt largely with the charges to be made for political broadcasts.

[75]47 U.S.C. § 315(c)(1) (1976). The FCC, in its 1981 report to Senator Goldwater (R-Ariz.), states that the equal time rules clearly apply to cable television, but that this application may have "occurred without benefit of any extended consideration . . . of the issue." FCC, Cable Television Bureau, Cable Television and the Political Broadcasting Laws: The 1980 Election Experience and Proposals for Change 17 (Rep. to Sen. Goldwater) (Jan. 1981) [hereinafter cited as Goldwater Report]. It should be noted that the FCC has indicated that it would not apply equal time rules to public access channels "as long as [these public access] channels . . . have inherent in their functioning access of a type which makes possible equal opportunities for political candidates." *Id.* at 27 (quoting FCC, Memorandum Opinion and Order in Docket 20508 (No. 80-608 Oct. 21, 1980)). Regulations concerning electoral communications can be found in FCC, *Primer on Political Broadcasting and Cablecasting*, 43 Rad. Reg. 2d (P & F) ¶ 1353 (1978).

[76]In its report, the FCC concludes that "without further clarifying legislation or Commission rulemaking [the reasonable access rule] could only be enforced against

This Article uses three variables to generate policy options for regulating electoral communications on cable television: the equal time doctrine, the reasonable access rule, and access channels.[77] The different options serve different goals, and the eventual selection of one of these policy options will depend on one's priority among the goals. These concrete proposals are intended to stimulate debate about the goals of a policy, the regulatory model selected, and the rationales for regulation.

1. *The broadcast model: equal time doctrine, reasonable access rule, no access channels.* If the full broadcast model were applied to cable television, it would achieve fairness and access at the expense of journalistic discretion. Governmental interference is likely to be greater than it would be in an access channel scheme,[78] because implementing the equal time doctrine and the reasonable access rule would require the promulgation of extensive rules and regulations.

For this policy option, it is crucial to determine whether the equal time requirement should apply to each channel separately or to the system as a whole. The potential chilling effect of the equal time doctrine might be alleviated if the equal time requirements were applied on a system-wide basis, because cablecasters with high-capacity systems and unused channel time could satisfy their equal opportunity obligations on any of their channels. To achieve true fairness, however, equal time requirements may need to be applied to each channel individually, because an appearance by a political candidate on a popular programming channel is not likely to be "equal" to an appearance at the same time of day on a little-watched cable channel.

A major problem with applying the broadcast model to cable is that the regulations may be difficult to enforce. While the local cable operator is legally responsible for the content, he has only limited control over the programming on his system.[79] The regulatory power of the FCC and of state and local governments extends only to the cable operator, not to program sup-

---

cable television systems operators with great difficulty, if at all." See the extended discussion in the Goldwater Report, *supra* note 75, at 22–25.

[77]In their various combinations, these three variables generate eight policy options. However, this Article will only consider the most important ones, as well as an additional option, the pure common carrier approach.

[78]*See infra* text accompanying notes 88–89 for further discussion of access channel schemes.

[79]*See* Goldwater Report, *supra* note 75, at i, 34.

pliers. The FCC, in a recent report, suggested some possible answers to these enforcement problems:

> The high channel capacity of cable television makes inconsistent a simultaneous pursuit of both responsibility [fairness] and diversity. Moreover, control as a practical . . . matter may not be reasonably feasible. This may either suggest that responsibilit[y] . . . devolve on the channel programmer or that . . . any requirement of balance or fairness be abandoned in favor of limited but absolute access rights . . . . Alternatively, all program content controls might be limited to channels over which the operator has actual as opposed to merely legal control.[80]

The equal opportunities rules have been effectively enforced over local broadcast television stations, however, despite the fact that much of their programming comes from the major networks, which are not subject to the equal time doctrine. These networks comply with equal time requirements to protect their local affiliates, and cable program suppliers could be expected to comply with similar regulations for the same reasons.

A potentially greater problem than enforcement is the constitutionality of the broadcast model as applied to cable television. Cable seriously challenges the rationale of technological scarcity as a justification for content regulation. First, cable television does not occupy the radio spectrum; instead, its signals are transmitted through cables strung along public utility poles or laid underground.[81] Governmental regulation to prevent signal interference is thus unnecessary.[82] Second, cable television can offer many more channels than traditional broadcast television.[83]

---

[80]*Id.* at 48–49.

[81]Although cable systems actually do transmit their signals using the radio spectrum, the same frequencies can be reused in the physical space adjacent to the cable because the radiation of the waves is almost completely within the cable. FCC Staff, *supra* note 73, at 19 n.32.

[82]Goldwater Report, *supra* note 75, at 6.

[83]While the systems being installed today may have up to 108 channels, almost 70% of the systems in place have 12 channels or fewer. Goldwater Report, *supra* note 75, at 4 (citing S. MAHONEY, N. DEMARTINO, AND R. STENGEL, KEEPING PACE WITH THE NEW TELEVISION 87 (1980)). Due to the FCC's "must carry" rules, which require cable systems to carry all local television broadcast stations, these channels are used primarily for the retransmission of broadcast television signals. 47 C.F.R. §§ 76.51–.65 (1981). Moreover, cable penetration is unlikely ever to reach 100%. Cable is currently subscribed to by only 30% to 35% of American households, Weekly Television Dig., Jan. 31, 1983, at 8. One study indicates the possibility that no more than 55% of homes in the United States will ever actually subscribe to cable. NAT'L ASS'N OF BROADCASTERS, NEW TECHNOLOGIES AFFECTING RADIO AND TELEVISION BROADCASTING 1 (1981).

Courts and commentators agree that spectrum scarcity cannot justify content regulation of cable.[84] In *Home Box Office, Inc. v. FCC*, for example, the court explicitly stated that

> [t]he First Amendment theory espoused in *National Broadcasting Co.* and reaffirmed in *Red Lion Broadcasting Co.* cannot be directly applied to cable television since an essential precondition of that theory—physical interference and scarcity requiring an umpiring role for government—is absent.[85]

Moreover, the court refused to rely on economic scarcity as a rationale for regulation, declaring that

> even though there is some evidence that local distribution of cable signals is a natural economic monopoly, which may raise the spectre of private censorship by the system owner, . . . scarcity which is the result solely of economic conditions is apparently insufficient to justify even limited government intrusion into the First Amendment rights of the conventional press, and there is nothing in the record before us to suggest a constitutional distinction between cable television and newspapers on this point.[86]

The public domain and public trustee rationales are also questionable justifications for regulating the content of cable television. Most cable systems use spectrum space to retransmit broadcast signals which then are distributed to the public by cable.[87] Cable systems also use municipal rights of way when they string their cables along public utility poles or underground. These facts would support application of the public domain and public trustee rationales. These rationales are weakened, however, by technological abundance, for it is hardly a privilege to use an abundant public resource. The greater the availability of electronic media outlets, the less the value of each outlet. The result is that the value of the use of the public resource diminishes, and any supranormal profits decline, thereby removing one justification for imposing unprofitable public interest programming requirements on cablecasters. The grant of a valuable

---

[84]*See, e.g.*, Goldwater Report, *supra* note 75, at 42; Schmidt, *supra* note 15, at 214.

[85]567 F.2d 9, 44–45 (D.C. Cir.), *cert. denied*, 434 U.S. 829 (1977).

[86]567 F.2d at 46 (citations omitted); *cf. supra* note 46 and accompanying text (discussing the Court's refusal to apply the economic scarcity rationale to print in Miami Herald Publishing Co. v. Tornillo, 418 U.S. 241 (1974)). *But see* Schmidt, *supra* note 15, at 214.

[87]The Cable Television Relay Service (CTRS) is specifically authorized by the FCC for this retransmission of broadcast signals to cable systems. 47 C.F.R. §§ 78.1, 78.5 (1981).

public franchise by the local or state government, however, may justify requiring that cable operators program in the public interest, particularly if such a grant guarantees them a monopoly and thus the likelihood of earning supranormal profits.

If neither the scarcity nor the public trustee rationales seem appropriate, the impact and invasiveness rationales might justify regulating electoral communications on many of the new media and continuing similar regulation of broadcast television. To apply this rationale, policymakers must determine whether a medium has special impact or invasiveness, and, if so, why. If the impact of the visual image itself raises concern, cable and all the other new technologies, as well as traditional broadcast television, would require content regulation. If policymakers are concerned with messages that can be absorbed with little or no affirmative effort, they may seek to regulate passive media such as television, but leave unregulated the media resembling print, in which the reader must make an affirmative effort to use the medium. They may also vary the level of regulation according to the amount of control a viewer or listener can exert over his own exposure to the messages being transmitted.

2. *No equal time doctrine, no reasonable access rule, access channels.* Under this regulatory scheme, cable operators would be required only to reserve certain channels for public access, either free or paid. During political campaigns, for example, an entire channel might be devoted to candidates' use. These access channels, like common carriers, would be available on a first-come, first-served basis. Governmental interference in this "partial separations" model would be limited to requiring the existence of access channels and perhaps to regulating their rates.

An alternative proposal is to apply the broadcast model, but to waive the equal time and reasonable access requirements for cablecasters who provide access channels.[88] Cable operators with large-capacity systems could then free themselves of content regulation and of the equal time doctrine's chilling effect. At the same time, this proposal would avoid imposing the burden of an access channel requirement on small-capacity cable systems that have little or no unused channel space.

---

[88] *See, e.g.*, Goldwater Report, *supra* note 75, at 49; *ABA Recommendations*, *supra* note 45, at 3.

An access channel scheme clearly would promote access to the media by political candidates while assuring minimum governmental interference with content. Treating the access channels like common carriers would eliminate the cable operator's programming discretion on these channels, but the operator would retain full journalistic discretion on nonaccess channels. Nonexempt interviews and issues programs should flourish, since cable operators could air them undeterred by the requirement of granting equal opportunities to all candidates. This regulatory scheme thus avoids the alleged chilling effect of the equal time doctrine and guarantees at least some media access for all candidates.

Such an access channel scheme, however, faces two significant problems. First, it may not be effective in actually informing the American electorate about political candidates; it is unlikely that many people will watch access channels when they can be watching entertainment or sports programming instead. Second, it raises a fairness problem. To achieve fairness, all candidates should have truly equal opportunities to reach their audiences. However, a cable operator could give a favored candidate time adjacent to a popular program while relegating opposing candidates to little-watched access channels.[89]

3. *No equal time doctrine, reasonable access rule, no access channels.* Either the reasonable access rule or an access channel requirement would assure candidates at least some access to cable and prevent a cablecaster from completely denying access to disfavored candidates. A policy option that uses the reasonable access rule would require substantial governmental intervention in the form of complicated regulations regarding who is entitled to access on what channels. An access channel requirement, on the other hand, would minimize such intervention but would allow cablecasters to assign appearances by political candidates to little-watched access channels.

A policy option using only the reasonable access rule might ensure access and avoid the problem of relegating candidates to little-watched channels if "reasonable" were interpreted to include access to all of a cablecaster's channels, including the popular entertainment channels. If cablecasters were required to provide access to noncommercial channels, how-

[89]*ABA Recommendations*, *supra* note 45, at 56 (separate views of K. Cox, member of Communications Law Comm.).

ever, the special nature and attraction of these channels would be disrupted.

4. *Equal time doctrine, no reasonable access rule, access channels.* This policy option combines content regulation with structural regulation. The access channel requirement would guarantee that if the cablecaster chose to avoid political programming on his entertainment channels, political candidates still could buy time on access channels. Cablecasters would not be required to grant access to entertainment channels whenever a candidate demands it. Moreover, if the equal time requirement were applied to each channel individually, a cablecaster choosing to air political programming would not be able to schedule a favored candidate's appearance on a popular entertainment channel unless he were willing to offer other candidates equal opportunities on the same channel. While this option may somewhat chill journalistic discretion, it requires less governmental intervention than the traditional broadcast model.[90]

The access channel requirement would render the reasonable access rule unnecessary, since there is no need to involve the government in both content and structural regulation when structural regulation alone will achieve access. Dropping the reasonable access rule would also alleviate the problem of political commercials interfering with the commercial-free nature of cable television. This option thus achieves both fairness and access, although at the cost of some governmental intervention.

5. *The print model: no equal time doctrine, no reasonable access rule, no access channels.* The policy option of complete deregulation of cable television obviously would minimize governmental interference and maximize journalistic discretion. If effective, deregulation might also achieve access, diversity, and even fairness. The marketplace itself should operate to encourage access and fairness. Cable operators with excess channel capacity should be willing to sell time to political candidates at reasonable rates without governmental intervention. The cablecaster's own economic interest should also lead him to ensure a large degree of fairness, at least to the major candidates. If a cablecaster seeks a broad audience on a given channel, he will

[90]In contrast, the option of applying *only* the equal time doctrine to cable television without an access channel requirement would promote fairness, but would not ensure access. It would also chill the journalistic discretion of cablecasters who wanted to air political programming but had inadequate channel capacity on which to give fringe candidates equal time.

be less likely to give any candidate an unfair advantage because this would tend to alienate part of that audience. This tendency is no guarantee of fairness, however, and local cablecasters with strong political leanings and sympathetic audiences would have less incentive to give all candidates equal opportunities.[91]

The most significant advantage of this policy option is that it may be the only constitutionally acceptable regulatory scheme. First, as noted above, any regulation of electoral communications on cable television through an equal time or reasonable access rule may be unconstitutional if the rationales justifying content regulation of broadcasting are inapplicable.[92] Second, structural regulation in the form of access channels may also be unconstitutional. In 1976, the FCC promulgated so-called capacity and access rules, requiring cable systems with more than 3500 subscribers to make certain channels available for access by third parties.[93] The Supreme Court in *FCC v. Midwest Video Corp.*, overturned these rules on statutory grounds, declaring that section 3(h) of the Communications Act of 1934 "preclude[s FCC] discretion to compel broadcasters to act as common carriers, even with respect to a portion of their total services . . . . [T]hat same constraint applies to the regulation of cable television systems."[94] The Court concluded that the access regulations could not be "reasonably ancillary" to the FCC's regulation of broadcast television.

Although it did not reach the constitutional issue, the Court did note that the question of whether such rules might violate the First Amendment rights of cable operators was not frivolous.[95] The Court of Appeals had, however, discussed the constitutional issue in dictum. Expressing concern that the FCC did not consider the cablecaster's journalistic discretion in promulgating its rules, the court noted that access channel requirements

---

[91]The analysis of this paragraph applies to broadcast television as well.

[92]The primary regulatory rationale of scarcity may not apply to cable television. *See supra* notes 82–87 and accompanying text.

[93]*In re* Amendment of the Comm'n's Rules & Regulations Concerning the Cable Television Channel Capacity and Access Channel Requirements, 59 F.C.C.2d 294, 297 (1976); *see also* Midwest Video Corp. v. FCC, 571 F.2d 1025, 1034 (8th Cir. 1978), *aff'd*, 440 U.S. 689 (1979).

[94]440 U.S. 689, 705 n.15 (1979), *aff'g* Midwest Video Corp. v. FCC, 571 F.2d 1025 (8th Cir. 1978). The Court declared that Congress had deliberately "restricted the [FCC's] ability to advance . . . public access at the expense of the [broadcaster's] journalistic freedom . . . ." *Id.* at 707. Section 3(h) of the Communications Act of 1934, codified at 47 U.S.C. § 153(h)(3)(h) (1976), provides that "a person engaged in . . . broadcasting shall not . . . be deemed a common carrier."

[95]440 U.S. at 709 n.19.

take programming completely out of the cable operator's hands.[96] The court described cable television as a "private electronic publication," and, referring to *Miami Herald*, concluded: "[W]e have seen and heard nothing in this case to indicate a constitutional distinction between cable systems and newspapers in the context of the government's power to compel public access."[97]

Thus, even if Congress itself decides to assure access by requiring access channels, courts may strike down such a requirement as an unconstitutional interference with maximum journalistic discretion.

6. *The pure common carrier model.* If the common carrier model were applied to cable television, there would be no equal time or reasonable access requirements, but cable channels would be leased on a first-come, first-served basis. This pure common carrier approach would separate control of the conduit (the transmission) from control of the content. In other words, the "cable . . . systems [would] provide channel facilities but have neither control nor ownership interest in the programs and information distributed through the system."[98]

The advantage of this separations policy would be that cable operators would not be able to control the content on their cable system's channels.[99] Leasing channels on a first-come, first-served basis would assure, for example, that a fifty-four channel system really provided fifty-four points of access and fifty-four chances for diversity. Common carrier regulation would have the virtue of eliminating "the pressure to regulate the content of broadband messages because distribution systems are local monopolies."[100] As it does with other natural monopolies, however, the government might regulate rates and certain other aspects of distribution.

Like other purely structural regulations, the common carrier model has no specific applicability to electoral communications, and thus differs little from any of the policy options that use access channels.

---

[96]Midwest Video Corp. v. FCC, 571 F.2d 1025, 1054 (8th Cir. 1978).

[97]*Id.* at 1056.

[98]Schmidt, *supra* note 15, at 223.

[99]On the other hand, denying the owner of the transmission facility the opportunity to provide programming might well discourage investment in cable systems.

[100]Schmidt, *supra* note 15, at 223.

B. *Micro Policy Options for Regulating Electoral Communications on Electronic Media Other than Cable*

The new electronic technologies[101] can be divided into two categories: transmission media, and services available on transmission media.[102] Regulatory schemes currently are imposed according to the transmission medium, not the service. Because many electronic services, such as teletext and videotex, will be available on a variety of transmission media, they present a special regulatory problem: the same service might be regulated differently depending on how it is transmitted. This problem is endemic to the micro policy approach, which treats each medium in isolation.

1. *Direct Broadcast Satellite.* In DBS technology, satellites deliver programming directly to the viewer's home. In June 1982, the FCC unanimously approved DBS service. In so doing, the FCC intentionally imposed almost no regulations, but rather left individual system operators free to choose whether to act as common carriers, broadcasters, or anything else.[103]

DBS seems closest to broadcast television in characteristics relevant to electoral communications. First, there is technological scarcity: DBS operators will have available to them only a limited number of frequencies and orbital positions.[104] The FCC has estimated that only six to ten DBS channels will be available nationwide.[105] Second, DBS makes use of a public resource, the

---

[101]Many of the so-called "new" technologies are in fact old technologies that are just now being exploited. A number of them are only variations of traditional broadcast television. Since few of these new technologies have been fully established, important regulatory characteristics such as channel capacity, availability, and market penetration can only be estimated.

[102]*See* Report and Order, 47 Fed. Reg. 31,555 (1982) (to be codified at 47 C.F.R. §§ 2.106, 94.65, 100); Fagan, *Direct Broadcast Satellites and the FCC: A Case Study in the Regulation of New Technology*, 29 FED. B. NEWS & J. 378, 382 (1982).

[103]Report and Order, *supra* note 102. In September 1982, the FCC authorized the first construction permit for interim DBS service. *See* Taylor, *DBS Service Still Long Way Off Despite FCC Permits*, TELEVISION/RADIO AGE, Oct. 18, 1982, at 37. On November 2, 1982, the FCC authorized construction permits for seven more companies. *See* Fagan, *supra* note 102, at 382. While DBS service may start as early as 1984, one estimate predicts that it may be 1995 or later before a significant nationwide audience is reached. Taylor, *supra*, at 38.

[104]These frequencies and orbital slots cannot be allocated among domestic operators until they are first allocated among 32 western hemisphere nations at the 1983 Regional Administrative Radio Conference. Therefore, the FCC's 1982 DBS authorizations have been labelled "interim." Taylor, *supra* note 103, at 38–39.

[105]FCC Staff, *supra* note 73, at 105.

airwaves.[106] Third, DBS' video picture has the impact and invasiveness of broadcast television, although the viewer may have greater control over it because he can choose not to subscribe to DBS service.

The very similarity between DBS and broadcast television might justify using DBS as an experimental forum to test the effect of deregulation on electoral communications. Unlike traditional broadcast television, however, DBS will provide a somewhat limited forum for electoral communications. DBS service will be almost universally available, but total subscribers may never exceed three to nine million homes because of the high cost of service.[107] Since DBS will be a pay service for the forseeable future, it will probably be devoted largely to sports and entertainment programming.[108] In addition, DBS will be a viable forum only for national campaigns, since its satellites will provide signals to large areas covering many states.

Any of the three regulatory models could be applied to DBS, but an access channel requirement would probably be inappropriate because DBS systems require a heavy and somewhat risky investment, and operators will control only a few channels. The reasonable access rule might disrupt the noncommercial nature of certain DBS channels.[109]

2. *Multipoint Distribution Service.* MDS is a video delivery system transmitted on microwave frequencies. The end user requires a special antenna to receive the signal and convert it to a conventional television frequency. MDS now provides only one channel of programming in any given area, and that programming is almost exclusively for-pay, prime-time entertainment programming.[110] In 1980, however, the FCC proposed a reallocation of spectrum frequencies that might make available up to thirty-three MDS channels per area.[111] MDS presently is

---

[106]*See* Fagan, *supra* note 102, at 381–82.

[107]*See* Taylor, *supra* note 103, at 41. Subscribers will need to lease or purchase a receiving antenna dish, as well as pay a monthly fee for programming service.

[108]However, Satellite Television Corporation has proposed to operate a cultural and public affairs channel in its four-channel system. *Direct Broadcast Satellites: Ownership and Access to the New Technology*, 33 FED. COM. L.J. 245, 246 n.6 (1981).

[109]On the other hand, if DBS offers a unique service such as HDTV, an equal time requirement for HDTV channels might be desirable to guarantee fairness on that scarce medium.

[110]As of April 1980, 54 MDS stations offered video programming. FCC Staff, *supra* note 73, at 22.

[111]Notice of Inquiry, Proposed Rulemaking and Order, 45 Fed. Reg. 29,323, 29,324 (1980) [hereinafter cited as MDS Proposed Rulemaking].

regulated as a common carrier,[112] but station operators may transmit their own programming up to fifty percent of the time if they establish a separate programming affiliate.[113]

MDS, like DBS, seems closest to broadcast television for regulatory purposes. First, there is technological scarcity. There are now only two channels allocated to MDS, and any new channels would have to be taken away from other services.[114] Second, MDS uses a public resource, the airwaves. Third, MDS's video picture has the same impact and invasiveness as traditional broadcast television, although, as with DBS, a viewer can choose not to subscribe to it.

MDS probably will not be an important forum for political programming, however, because of its focus on entertainment programming and its limited availability.[115] Although the equal time doctrine probably would have little effect on the operation of an MDS station, the reasonable access rule might disrupt programming schedules and impose obligations that might discourage initiation of service on MDS channels.

3. *New broadcast technologies: Subscription Television, Low Power Television, and VHF drop-ins.* These new broadcast technologies all use the same spectrum band as traditional VHF and UHF television. Their regulatory characteristics are similar enough that they will be considered together.

Subscription Television (STV) is simply a standard broadcast television signal transmitted in a scrambled form so that the viewer must lease special receiving equipment to watch it on his home television set. STV provides for-pay, largely prime-time, and almost exclusively entertainment and sports programming.

STV adds no new technological channel capacity; it only increases available programming if an STV operator uses an otherwise idle television channel.[116] It may even decrease the free television programming available if a regular television station switches to STV operation.[117] STV is currently regulated

---

[112]47 C.F.R. § 21.903(b) (1981).

[113]*Id.*

[114]*See* MDS Proposed Rulemaking, *supra* note 111, at 29,326.

[115]MDS requires a clear line of sight from the transmitter to the receiving antenna. NAT'L ASS'N OF BROADCASTERS, *supra* note 83, at 7.

[116]FCC Staff, *supra* note 73, at 20.

[117]This is particularly the case because the commission has recently deleted the requirement that STV stations broadcast a certain amount of conventional programming each week. Third Report and Order, In the Matter of Subscription Television Service,

the same as traditional broadcast television.[118] Like similar limited-capacity new technologies with for-pay entertainment programming, however, STV is unlikely ever to be an important political forum.

Low Power Television (LPTV) stations can be added where the market is too small to support a full-power station, or where a full-power station would interfere with existing stations.[119] LPTV stations would be able to originate and to carry their own programming, and may support themselves by advertisements, subscriptions, or taxes.[120] The FCC has proposed to impose political broadcasting rules on LPTV stations concomitant with each station's origination capabilities.[121] Since LPTV service might add one new broadcast station to many areas, LPTV stations could form a new national or regional network. These stations could provide an important new forum for political programming.

VHF drop-ins are full-power VHF television stations that use more advanced directional signals to avoid unacceptable interference with nearby stations. VHF drop-ins also may add one new broadcast station to certain markets. The FCC has proposed to regulate such stations in the same way it regulates traditional broadcast stations.[122]

All the rationales justifying content regulation of traditional broadcast television would apply to these new broadcast technologies because they all require scarce spectrum space to transmit their signals, all use a public resource, and all have a visual impact equivalent to that of traditional television. As with DBS, however, their very similarity to broadcast television might justify using one or more of these new technologies to experiment with new forms of regulation, including deregulation.

Scarcity analysis depends on whether or not each medium is unique. Each of the new media depends on a scarce portion of

---

47 Fed. Reg. 30,069 (1982). As of February 29, 1980, 17 STV applications were for conversion of existing stations to STV service, and 90 STV applications were for initiation of STV service on previously unused channels. NAT'L ASS'N OF BROADCASTERS, *supra* note 83, at 20 n.33 (citing *Addendum: The Status of STV as of February 29, 1980*, Pay TV Newsletter, Mar. 7, 1980, at 166).

[118]47 C.F.R. § 73.643(b) (1981).

[119]Ferris, *Direct Broadcast Satellites: A Piece of the Video Puzzle*, 33 FED. COM. L.J. 169, 172 (1981).

[120]*See Report and Order in Inquiry into the Future Role of Low Power Television Broadcasting and Television Translators in the National Telecommunications Systems*, 51 Rad. Reg. 2d (P & F) ¶ 480 (Apr. 26, 1982).

[121]*Id.* at 519.

[122]*See* NAT'L ASS'N OF BROADCASTERS, *supra* note 83, at 16.

the radio spectrum, but if they are fungible, the new media in combination with broadcast television might greatly alleviate any perceived scarcity. These three new broadcast technologies combined, however, are unlikely to add more than two or three channels to any given market area. Thus, they probably do not significantly reduce scarcity in any given market, and therefore would not justify modifying current regulations based on scarcity.

The for-pay portion of STV and any LPTV station supported by subscription fees are unlikely to be important forums for electoral communications because they will be devoted almost exclusively to entertainment programs and will not be available to all viewers. On the other hand, LPTV stations that are advertiser supported and provide "free" programming may provide an important new political forum, especially for local candidates.

4. *Videocassettes and videodiscs.* Videocassettes and videodiscs are prerecorded tapes and discs that can be played with special equipment on an unused television channel. Viewers can record programs themselves or can play prerecorded cassettes and discs.

Videocassettes and videodiscs are similar to the print media because they offer nearly unlimited diversity and complete user control. Cassettes and discs do not suffer from technological scarcity, nor do they use a public resource. Only in visual impact do they differ from the print media.

Videocassettes and videodiscs should be regulated under the print model. Any greater content or structural regulation would be almost impossible to implement and almost certainly unconstitutional. These cassettes and discs also will not be an important forum for electoral communications because they offer mainly prerecorded entertainment programming.

5. *High Definition Television.* HDTV is a developing technology that would provide superior color and detail by using an 1100-line system instead of the current 525-line system. HDTV may be available on cable, videocassettes, DBS, and possibly traditional broadcast television stations. The unique visual impact of HDTV may make it important for electoral communications.

The scarcity rationale may justify content regulation of HDTV. Suppose HDTV were available only on cable, and because of technological constraints akin to bandwidth "hogging,"

it were available only on two cable channels. Suppose also that large-screen projector television sets are commonplace. The question then is whether HDTV, with its sharp picture quality, superior skin tones, and distinct advantage on the projector screen, is just another electronic medium, or whether it is a unique and especially powerful medium for political candidates. If HDTV is unique, then it is also scarce. Thus, scarcity could justify regulation of HDTV to assure at least fairness, and perhaps also access, for political candidates.[123] Without regulation, the owner of a cable system could give favored candidates a distinct advantage by giving them exclusive access to the two HDTV channels.

6. *Teletext and videotex.* Teletext and videotex, two forms of "electronic publishing," transmit text and graphics directly to the user's home video display screen. Teletext systems have the capacity to send data in only one direction, from the central transmitter to the end user's computer or television screen,[124] and can transmit only a limited amount of data.[125]

Videotex is a more sophisticated technology; its two-way capacity allows communication between the end user and the central computer, allowing the user to search more precisely for information. Videotex systems can provide access to an unlimited number of data bases and pages of information.[126] Videotex is now available almost exclusively over telephone lines, but it will soon be available on cable systems which have two-way capacity.[127]

Virtually any information that can be communicated in textual form can be put into a teletext or videotex data base. Editorial and factual information, including information about political candidates, could be included in the data base. Teletext and videotex thus offer a new and unique forum for electoral communications. However, because the user can select the information he desires, and because electronic publishing lacks

---

[123]The scarcity of HDTV is at least partly technological.

[124]Neustadt, Skall & Hammer, *The Regulation of Electronic Publishing*, 33 FED. COM. L.J. 331, 332 (1981).

[125]Broadcast teletext, which uses the vertical blanking interval of a regular broadcast signal, can only carry about two hundred pages of information. *Id.* at 333, 336. Teletext also can be transmitted by cable and MDS, and perhaps by DBS, LPTV, and the FM radio subcarrier signal. *Id.* at 337 n.20. Teletext was authorized by the FCC on March 31, 1983. *See Teletext Authorized by FCC*, N.Y. Times, Apr. 1, 1983, at D1, col. 1.

[126]Neustadt, Skall & Hammer, *supra* note 124, at 337, 340.

[127]*Id.* at 339.

the visual and auditory impact of the other electronic media, teletext and videotex may be of little importance for electoral communications.

The traditional rationales for regulation seem inappropriate for electronic publishing. Electronic publishing exhibits some technological scarcity in the limited number of pages available on some services and in the fact that some services can be transmitted only over certain kinds of transmission media.[128] If teletext and videotex are in the same product market, however, and if that market also includes the print media, then there may be no scarcity at all. Scarcity of transmission media, on the other hand, would be a serious regulatory concern only if videotex were considered unique because of its two-way capacity.

Impact and invasiveness do not seem viable rationales for regulating electronic publishing. Electronic publishing's impact is muted because nearly all the information must be read to be understood, and invasiveness is reduced because the user can control the flow of information into his home.

The traditional broadcast model seems inappropriate for electronic publishing. The potential chilling effect of the equal time doctrine might discourage electronic publishers from providing political information for fear that they would have to relinquish scarce page capacity to competing candidates.[129] The reasonable access rule might impose unfair burdens on electronic publishing services with limited page capacities, and even discourage their development.[130]

Either the common carrier model or an access requirement would alleviate the potential problem of a transmission media bottleneck and help to assure greater access and diversity. The print model, on the other hand, would avoid most of the special problems of regulating the content of electronic publishing. Complete deregulation would relieve the uneven burdens of the access requirement and the potential chilling effect of the equal time doctrine. The print model would also recognize the futility of adopting a paternalistic approach to electoral communica-

---

[128]For example, videotex is presently available only over telephone lines. Structural regulation of scarce transmission media may alleviate this potential bottleneck. *Id.*

[129]*See id.* at 350–52. On the other hand, videotex services, with their great capacity, would not be greatly burdened by an equal time or reasonable access rule. *Id.* at 351.

[130]*Id.* at 350–53. One way to alleviate this problem is to apply the reasonable access requirement to the regular broadcast signal and the teletext signal together. *Id.* at 368–69.

tions, and deregulation would avoid the First Amendment dangers of regulating the electronically transmitted content of newspapers and other print media.[131]

V. The Macro Approach

The macro approach to the regulation of electoral communications involves formulating a comprehensive scheme for regulating the communications media as a whole. Although it regulates comprehensively, the macro approach seeks to take advantage of each medium's unique characteristics in order to maximize achievement of the goals of minimum governmental intervention, diversity, access, and fairness.

Professor Lee Bollinger has explored a macro approach for two media: print and broadcast television. His "partial regulatory" scheme would leave print unregulated,[132] but would allow legislative and administrative experimentation with various types of access regulation of broadcasting.[133] Although Bollinger concedes for the purposes of his analysis that any differences between the print and broadcast media are "too insignificant to justify momentous distinctions in treatment under the first amendment,"[134] he nonetheless advocates different regulatory approaches for the two media in order to promote the competing regulatory goals of access and fairness, on one hand, and minimum governmental intervention, on the other hand.[135]

Bollinger argues that the Supreme Court adopted this partial regulatory approach in *Miami Herald* and *Red Lion*, the cases he uses as separate foundations on which to build his macro scheme. His approach strives to retain the advantages of an unregulated print medium, including preservation of a link with constitutional traditions,[136] assurance that at least one medium will remain independent of government,[137] and provision of "in-

---

[131]W. Read, *supra* note 16, at 14.

[132]Bollinger, *supra* note 31, at 32.

[133]*Id.*

[134]*Id.* at 16. The concept that the print and broadcast media are not fundamentally different is a hotly debated issue and by no means universally accepted.

[135]*Id.* at 36. Other commentators agree that the media are not fundamentally different, but conclude from this that they should be regulated the same way. Moreover, these commentators generally advocate complete deregulation. *See, e.g.*, W. Read, *supra* note 16, at 14.

[136]Bollinger, *supra* note 31, at 32–33.

[137]*Id.* at 32.

formation that might not be disseminated by the regulated sector."[138] The regulated broadcast medium, on the other hand, would provide balanced coverage of important issues and access to candidates and points of view that might be neglected in the unregulated media.

The policy options set forth in this Section take a comprehensive approach to regulating electoral communications on broadcast television, print, cable, and the other new media. This Section examines various macro policy options generated by combinations of the equal time doctrine, the reasonable access rule, and an access channel requirement,[139] as well as the option of complete deregulation.

A. *Unregulated print, equal time doctrine without reasonable access rule for broadcast television, access channels on cable, other media unregulated.* This first macro policy option[140] builds on Bollinger's scheme, by using cable television to provide access. It provides the advantages of an unregulated print medium. Retaining the equal time doctrine assures a large degree of fairness, which is particularly important if broadcast television remains the dominant medium.

Access channels would assure that political candidates have at least some access to the home television screen. Cable television, with its tremendous channel capacity in the hands of a single owner-operator, seems to be a more appropriate medium on which to place an access requirement than broadcast television. Even if the equal time doctrine has a chilling effect on broadcast television, political candidates will always have access available on cable. If candidates can get access on cable, the reasonable access rule becomes unnecessary and even paternalistic.[141] Presumably, if viewers are interested in political programming, then media operators will provide it; the

---

[138]*Id.*

[139]These three variables can yield eight policy options, but this Section discusses only the most important ones. Unregulated print is held as a constant in generating these options because the Supreme Court has consistently ruled that any content or access regulation of print is unconstitutional. Even more policy options could be generated by regulating radio differently from television. The FCC, recognizing the different characteristics of radio, has recently removed many regulations from radio, but not the political broadcasting rules. *Deregulation of Radio*, 84 F.C.C.2d 968, *recons. denied*, 87 F.C.C.2d 797 (1981).

[140]This option is very similar to Cable Micro Policy Option 4. *See supra* note 90 and accompanying text.

[141]*See* Schmidt, *supra* note 15, at 215–16.

government should not force such programming onto broadcast television.[142]

The viability of this approach depends on a vast majority of Americans subscribing to cable. It also assumes that cable and broadcast television are interchangeable for electoral communications and that imposing an access requirement on cable television is constitutional. This first macro policy option is obviously one for the future.

B. *Deregulation of broadcast television, access channels for cable, deregulation of all other media.* This macro policy option would deregulate electoral communications on all media except for an access channel requirement on cable television. The absence of the chilling effect of the equal time doctrine should encourage broadcasters and other media operators to offer more and better political programming. While fairness would no longer be assured, relative balance in political programming may be assured in a medium that tries to attract mass audiences because broadcasters would be careful to avoid unfairness to candidates at the risk of alienating part of their audience.[143] At the same time, political candidates could always get airtime on the access channels. The success of this macro policy option, however, depends on widely available cable systems with high channel capacity.

C. *Equal time doctrine (and perhaps reasonable access rule) for broadcast television, complete deregulation of all other media.* This macro policy option retains current regulation of broadcast television and deregulates all the other media. Such a policy might be appropriate if traditional broadcast television remains dominant, because it would assure some balance and fairness on a medium that political candidates may consider unique and indispensable to their campaigns.

This policy assumes that cable and the new media, given adequate channel capacity, will provide plenty of access for political candidates without governmental intervention. If this is so, an access channel requirement on cable and the reasonable access rule for broadcast television would be unnecessary. Moreover, if ownership of these media outlets is diverse, an overall balance of viewpoints may prevail. This experiment with deregulation of electoral communications should indicate how

---

[142]*Id.*

[143]This relative balance, if it exists at all, may only be inherent in national networks.

well such a scheme works; if it is successful, deregulation of all media might follow.

This policy option also alleviates a major implementation problem of many comprehensive regulatory schemes: the differential availability of and subscription to various new media in any given geographic area. For example, if ten years from now an area has only broadcast television, a national policy based on the absence of scarcity would be inappropriate; if only half of its residents subscribe to cable, an access channel requirement on cable television would not be viable.

D. *The print model: complete deregulation of all communications media.* This policy option would completely deregulate electoral communications on all communications media. Several considerations support this policy. For example, cable, DBS, MDS, and LPTV all add to the channel capacity of video delivery systems, thereby alleviating, if not eliminating, scarcity. The reduction of technological scarcity removes a prime justification for regulating electoral communications. A policy of complete deregulation also would have the virtue of sparing courts the difficult task of comparing and contrasting the relevant characteristics of the various media. Deregulation would reduce the danger of inconsistent regulatory distinctions among electronic services that can be transmitted over various media. In addition, deregulation will reduce the danger that "as newspapers . . . come to rely on satellites and other regulated communications technologies . . . they . . . will be drawn into the regulatory web."[144] Finally, any chilling effect of current equal time regulation would also be eliminated.

Complete deregulation would maximize journalistic discretion but would not assure fairness, access, or diversity. However, an important purpose of the First Amendment is to foster an unregulated "marketplace of ideas,"[145] with competition among media outlets ensuring access and diversity. If consumers use a variety of independently controlled media outlets, competition and diversity among these outlets should provide overall balance and fairness without regulation.

Some structural regulation may be necessary to assure access and diversity in a deregulated communications market. Cross- and multiple-ownership regulations may help maximize these

---

[144]W. Read, *supra* note 16, at 14.

[145]*See* Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

goals at both the local and national levels by assuring that media outlets are independently controlled. Otherwise, if a single entity controls the local broadcast station, cable system, and MDS system, that entity will be able to deny access to any given candidate or point of view, making the multiplication of available channels meaningless in terms of electoral communications.

Before adopting a policy of deregulation, policymakers must look closely to see whether scarcity has indeed been eliminated. As the new electronic technologies have increased channel capacity, the uses for that capacity have also grown.[146] In addition, many of these new electronic media will be devoted almost exclusively to entertainment programming, and thus may do little to provide additional media outlets for electoral communications.

Deregulation also depends on the new media being widely available and widely used, but economic factors may limit the number of new technologies in a given area. Many of these new technologies are pay or subscription services available only to those willing and able to pay the tariff, and some may require expensive receiving equipment. These new technologies, therefore, may not contribute to the true diversity of outlets necessary for successful deregulation.[147]

The elimination of scarcity also depends on who controls the new media. Newer cable systems may provide fifty-four or more new channels, but because they will all be controlled by a single entity, it may be inaccurate to say that they add that many new media outlets to any given market. The elimination of scarcity rests on the further assumption that all electronic media are fungible in terms of electoral communications.[148] A lack of fungibility may stem from the phenomenon known as "scarcity of viewer preferences."[149] If traditional broadcast television re-

---

[146]New channel capacity is by no means devoted exclusively to video delivery of public programming; new channels are now used for data transmission, closed-circuit television, and a variety of other business and educational purposes.

[147]For example, it is unlikely that a significant number of people will subscribe to both cable and DBS, because DBS will serve primarily as a substitute for cable television in those areas where cable service is not available.

[148]Antitrust analysis might help policymakers decide whether there is still scarcity in a given area. They might ask, for example, whether viewers and candidates consider the different media interchangeable, whether each of the media is equally available to viewers, and whether the rates per viewer minute are roughly equivalent. Services such as HDTV, for example, may not be fungible. *See supra* note 125 and accompanying text.

[149]*See* Schimdt, *supra* note 15, at 214.

mains dominant, it may be a unique and thus scarce medium to political candidates wanting to reach the widest possible audiences. Such circumstances may justify continued regulation of electoral communications to assure fairness and access. Paternalistic regulation of this type may seriously conflict with First Amendment values, however, and thus may be both unwise and unconstitutional.[150]

Despite the advantages of deregulation, some continued regulation of electoral communications may be desirable if scarcity persists, if the media are not fungible, or if the free market fails to produce optimal levels of political programming. The first three macro policy options, therefore, have proposed a large degree of deregulation, but retain some regulation as a safeguard.

VI. Conclusion and Recommendations: A Three-Step Macro Approach

The final task is to select the best policy for regulating electoral communications in the future. The micro approach focuses on each medium in isolation, neglecting the communications picture as a whole. It tends to lead to ad hoc, uncoordinated policymaking as new problems or new technologies arise.

The macro approach, on the other hand, is better suited to formulating future regulatory policies for both new and existing media. The comprehensive nature of the macro approach recognizes the large degree of fungibility among the electronic media, but is also flexible enough to make the best use of the unique characteristics of each medium. The macro approach can also take account of such realities as the apparent dominance of broadcast television, the likelihood that many of the new media will be devoted largely to entertainment programming, and the fact that many of the new media will not be widely available or widely used. A comprehensive regulatory policy also can respond more readily to changes in the makeup of mass communications.

The best macro policy approach is not a single policy option, but a three-step sequence of three different macro policy op-

---

[150] *Id.*

tions. Each successive policy option would be implemented when experience demonstrates that it is ready to be adopted.

Step 1. The first step would retain the equal time doctrine and reasonable access rule for broadcast television, and deregulate or leave unregulated all the other media.[151] Step 1 should be adopted immediately, and could be expected to remain in effect for the foreseeable future. It would require the deregulation of cable television, an important change from current policy.

This macro scheme is a pragmatic response to the unavoidable fact that television will remain the dominant medium in electoral communications for the foreseeable future. Even with the advent of so many new electronic media, there will remain many people who do not have available, cannot afford, or choose not to pay for these new sources of information and programming. Also, high channel capacity systems, if they are controlled by a single entity, may effectively add but a single outlet for electoral communications to a given area.

Although the equal time doctrine and the reasonable access rule unquestionably impair the journalistic discretion of broadcasters, they promote an important countervailing value: the maintenance of a fair and viable electoral system. This is particularly important in areas where only a few media outlets are available, and disfavored candidates face the possibility of being excluded from the airwaves. Such exclusion would shake assumptions regarding free elections and offend many Americans' sense of fairness, and might lead to an even further decline in voter participation and confidence in the electoral system.

Deregulating the other electronic media and leaving the print media unregulated will allow policymakers to experiment with deregulation, and to determine its effect on political programming. If there is public demand for more and better political programming, the unregulated media should provide it, particularly since they will not be deterred by equal time requirements. Regulators will be able to see whether political candidates are given adequate access to these other media, and whether deregulation can achieve relative fairness in the presentation of all candidates.

---

[151]This policy option has been discussed above as Macro Policy Option C. *See supra* Section V (C).

Structural regulations, especially the cross-ownership prohibition, might need to be retained in smaller markets. If one entity were allowed to control all the media outlets in a given area, that entity could deny access to any disfavored candidate and thereby subvert the goals of access, diversity, and fairness. This concern diminishes considerably in major metropolitan areas where there are abundant media outlets.

Step 2. The second step would drop the reasonable access rule from broadcast television, impose an access channel requirement on cable television, and leave the other media unregulated.[152] This step can be taken only when cable subscription exceeds sixty-five or seventy percent of American households. Furthermore, access regulation of cable television may be unconstitutional.[153]

If the vast majority of Americans should subscribe to cable, it would make sense to switch the access regulation from broadcast television to cable television. Cable television is better suited than broadcast television to providing access because its tremendous channel capacity is under the control of a single owner-operator. Imposing an access channel requirement on cable television would require, however, that most cable systems have a high channel capacity, which is not presently the case.

It may be unnecessary to impose an access channel requirement on cable if cable and other media operators demonstrate under Step 1 that they will make access available for political candidates on their own initiative. Cable operators would have a special incentive to provide such access so as to avoid access regulation during Step 2.

Step 3. The final step would completely deregulate all mass communications technologies.[154] There are no specific indicators of when it is appropriate to move to Step 3. Experience under Steps 1 and 2 must first demonstrate that adequate access to widely-watched media is available for all political candidates. Second, experience with deregulation must demonstrate that

---

[152]This policy option has been discussed above as Macro Policy Option B. *See supra* text accompanying note 145.

[153]If Step 2 were unconstitutional, it would have to be deleted, but it is largely a refinement of Step 1. Step 3 can be reached without Step 2, if necessary.

[154]This policy option has been discussed above as Macro Policy Option D. *See supra* text accompanying notes 144–50.

fungible unregulated media maintain at least a semblance of balance and fairness in their coverage of political candidates.

If the unregulated market can provide adequate access and relative fairness, governmental regulation becomes both unnecessary and undesirable. Complete deregulation of electoral communications will then maximize achievement of the four posited goals. Even if it cannot be successfully adopted immediately, complete deregulation should remain the ultimate aim of policymakers.

ARTICLE

COMCAR: A MARKETPLACE CABLE TELEVISION FRANCHISE STRUCTURE

MARK S. NADEL*

Three years ago, in an article in the Harvard Journal on Legislation *by Charles D. Ferris,** *then-chairman of the Federal Communications Commission, it was suggested that the Commission could use its power over common carriers in a more flexible way to encourage the development of new communications media. Since that time, technological advancement has made the need for regulatory adjustment in the media marketplace even more acutely apparent. In response to this need, this Article proposes a specific program for structuring cable television service, using a form of common carrier regulation.*

Mr. Nadel argues that a fresh approach to cable regulation must be taken if viewers are to receive the greatest array of services at the most competitive prices. He proposes a deregulated system, named COMCAR, in which cable operators function much like common carriers. COMCAR eliminates the operator's control over access to its system's channels and allows all programmers capable of paying a competitive market price to gain access to the system. While the proposal runs counter to the tenor of cable legislation now before Congress, the author argues that its deregulatory aims are best suited to achieving efficiency in the delivery of cable services.

When an industry such as cable television has characteristics of a natural monopoly,[1] the government may treat it in several

---

*Adjunct Assistant Professor, Pace University. B.A., Amherst College, 1978; J.D., Harvard Law School, 1981. The author would like to thank the many people who have commented upon and criticized earlier drafts of this Article, including Ben Compaine, Jeff Forbes, Budd Gibbs, Douglas Ginsburg, Ken Gordon, Paul Isaac, Alfred Kahn, Tony Oettinger, Carl Oppedahl, and Harlan Rosenzweig, but especially Eugene Nadel and Eli Noam, without whose extensive, constructive, and indispensible help this Article would have been abandoned many drafts ago.

**Ferris, *Common Carrier Regulation for the Future*, 17 HARV. J. ON LEGIS. 241 (1980).

[1]A natural monopoly exists when there are economies of scale so persistent that a single firm can serve a market at a lower unit cost than two or more firms. *See, e.g.*, P. AREEDA, ANTITRUST ANALYSIS 11–12 (3d ed. 1981); Posner, *Natural Monopoly and Its Regulation*, 21 STAN. L. REV. 548 (1969).

Studies finding significant economies of scale in cable service include E. Noam, Economies of Scale and Competitive Entry in Cable Television (Columbia Univ. Business School Working Paper No. 430A, Dec. 1982); B. Owen & P. Greenhalgh, Competitive Policy Considerations in Cable Television Franchising (Oct. 1982) (Economists, Inc., Washington, D.C.).

But see Greater Freemont, Inc. v. City of Freemont, 302 F. Supp. 652, 657 (N.D. Ohio 1968) (stating that "CATV is not a natural monopoly"); Community Communications Co. v. City of Boulder, 630 F.2d 704, 712 (10th Cir. 1980) (Markey, C.J., dissenting) (in which Chief Judge Markey claimed that "the city's sole defense is to pretend, disingenuously and contrary to the extensive, uncontradicted testimony and the specific

ways.[2] Public ownership is clearly one option,[3] but in a free market society private ownership is ordinarily preferred.[4] In either case, however, if the industry is producing an essential product or service without close substitute, the problem of misallocation of resources must be addressed.[5] Normally, this is handled by eschewing both laissez-faire treatment and public control in favor of government rate regulation.[6]

One commentator has argued that the administrative burdens involved in rate regulation could be avoided by allowing private sector firms to bid competitively for a franchise.[7] This arrangement, he alleged, would dissipate any monopoly profits without the need for rate regulation. Unfortunately, as critics soon pointed out, such a single auction cannot determine more than one variable efficiently. It fails to anticipate all of the complex

---

findings of the trial judge . . . that cable is a 'natural monopoly'"), *rev'd*, 445 U.S. 40 (1982).

Even if cable distribution does not enjoy economies of scale, the FCC has recognized that an operator must achieve 30% to 40% penetration to break even. This condition weighs against granting multiple cable franchises in a single locality. Inquiry Into the Economic Relationship Between Television Broadcasting and Cable Television, Report, 71 F.C.C.2d 632, 671 (1979). In addition, the "limited availability of space for pole attachments and underground ducts in many communities would preclude" competitive entry by cable systems. Miller & Beals, *Regulating Cable Television*, 57 WASH. L. REV. 85, 99 (1981); *see also infra* note 25.

[2]*See, e.g.*, M. FRIEDMAN, CAPITALISM AND FREEDOM 28 (1962) ("When technical conditions make a natural monopoly the natural outcome of competitive market forces, there are only three alternatives that seem available: private monopoly, public monopoly, or public regulation.").

[3]Public ownership of natural monopolies is most frequently employed in other countries, particularly in the telephone and television industries. *See* G. BROCK, THE TELECOMMUNICATIONS INDUSTRY 126–47 (1981). In the United States, public ownership has long been important in public utility industries. *See* R. SCHMALENSEE, THE CONTROL OF NATURAL MONOPOLIES 85–100 (1979). For a discussion of municipal ownership of cable television in the United States, see *infra* note 84.

[4]*See* 1 A. KAHN, THE ECONOMICS OF REGULATION: PRINCIPLES AND INSTITUTIONS 1–2 (1970); Posner, *supra* note 1, at 636–37 ("the major difficulty [with public ownership] is the absence of a profit incentive . . . . [A]s the Soviet Union has tacitly acknowledged, no one has yet discovered an adequate substitute for profits as the driving force of industrial efficiency in an advanced economy . . . . A public enterprise would seem especially susceptible to pressure from political interest groups . . . .").

[5]The problem of resource misallocation results when a monopoly produces fewer goods, at higher prices, than would be produced in a competitive market. For a discussion, see P. AREEDA, *supra* note 1, at 13–17.

[6]Regulation allows the government to prevent a single, fortunate, private monopolist from producing too few goods, charging a price that exceeds the one that would be charged in a competitive market, and thus earning an excess profit or monopoly rent. It also allows the public to have some influence over important investment and service decisions. Of course regulation imposes additional costs on the public because it requires extra labor and sometimes reduces beneficial competition. *See* Competitive Common Carrier Rulemaking, 77 F.C.C.2d 308, 358 (1980); 1 A. KAHN, *supra* note 4, passim.

[7]Demsetz, *Why Regulate Utilities?*, 11 J.L. & ECON. 55 (1968).

marketplace forces that, over time, create pressure for continuous adjustments of relative prices and outputs.[8] While it is possible that a flexible contract—adjustable over time to take account of market shifts—could be developed, such an approach merely involves another form of regulation.[9] Treatment of a monopolist thus requires an institutional choice among imperfect solutions.[10]

Because economic regulation of cable television service is costly,[11] and because the private sector appears willing and able to assume the risks of development,[12] this Article offers an alternative structure for cable television franchising. The proposed system incorporates a "double-bidding" mechanism:[13] applicants first bid for the system franchise, then programmers bid for channel leases on the system. This structure is presented under the name COMCAR.

Before presenting COMCAR, it is useful to review the present cable television industry; therefore, Section I provides a historical overview of the structure and regulation of the cable industry, and describes its economic problems and its First Amendment infirmities. Section 1 also demonstrates the inadequacies of current regulations and the need for change. Section II summarizes the basic elements of COMCAR and discusses the possible statutory and jurisdictional obstacles to its implementation. Section III sets out the process of establishing the terms of the

---

[8]*See* Telser, *On the Regulation of Industry: A Note*, 77 J. POL. ECON. 937 (1969); Williamson, *Franchise Bidding for Natural Monopolies: In General and with Respect to CATV*, 7 BELL J. ECON. 73 (1976).

[9]*See* Goldberg, *Regulation and Administered Contracts*, 7 BELL J. ECON. 426 (1976).

[10]*See* Demsetz, *Information and Efficiency: Another Viewpoint*, 12 J.L. & ECON. 1 (1969).

[11]*See* R. Peterson & A. Pierce, The Cost of Cable Television Regulatory and Franchise Requirements: A Preliminary Analysis 31–32 (Apr. 1982) (Ernst & Whinney, Tacoma, Wash.) (commissioned by the Nat'l Cable Television Ass'n) (estimating the regulatory costs in one sample system at more than five dollars per home per month); *see also* R. Posner, Cable Television: The Problem of Local Monopoly 27–33 (1970) (RAND Memo No. RM-6309-FF) (suggesting that the problems of regulation may be particularly acute with respect to cable television).

[12]The private sector has relied primarily on debt financing, including substantial bank loans to pay the tremendous construction costs of wiring a locality. M. HAMBURG, ALL ABOUT CABLE A411–15 (1981). This indebtedness to traditionally conservative lenders increases the pressure on cable owners to avoid new, unproven ideas that could disturb the stability of future earnings.

[13]A similar structure was proposed in Loeb & Magat, *A Decentralized Method for Utility Regulation*, 22 J.L. & ECON. 399 (1979). The authors argue that if the government gives a natural monopolist a subsidy equal to its uncaptured consumer surplus, the monopolist will produce and price at a competitive level. This subsidy would be financed either by selling franchises or by imposing a lump-sum tax on the monopolist.

franchise license and selecting the franchisee. In Section IV, the economics of the allocation of channels is analyzed. Section V then looks at a particular aspect of cable television—public services such as universal service, local access channels, and local studio facilities.

I. Structure of the Cable Industry

A. *Regulatory History*

Cable television originated in 1949 as community antenna television (CATV), a service designed to improve the reception of broadcast signals in mountainous rural areas that were far from any broadcast stations. A single powerful antenna was installed to collect and to amplify weak broadcast signals. The strengthened signals were then retransmitted to individual homes by coaxial cables.[14] The Federal Communications Commission (FCC) initially refused to exercise jurisdiction over cable, but by 1972 the Commission had enacted a wide range of broadcast-like regulations,[15] most of which were developed to limit a cable system's right to retransmit the broadcast television programming picked up by the system's receiving antenna.[16] The rules also required cable operators to provide new services, such as local access programming, forcing them to develop cable's major technical asset, its abundant channel capacity.[17] After 1973, however, the Commission began to lessen its control over cable; and, today deregulation has virtually eliminated fed-

---

[14]For a short description of the technology involved in cable television, see Sloan Comm'n on Cable Communications, On the Cable: The Television of Abundance 11–16 (1971) [hereinafter cited as Sloan Report]. For a short general history of the industry, see Besen & Crandall, *The Deregulation of Cable Television*, 44 Law & Contemp. Probs. 77 (1981).

[15]*See infra* Section II.B.

[16]Initially, this was accomplished by regulating the microwave carriers that transmitted distant broadcast signals to system antennas. *See, e.g.*, Carter Mountain Transmission Corp., 32 F.C.C. 459, *aff'd*, 321 F.2d 359 (D.C. Cir. 1962), *cert. denied*, 375 U.S. 951 (1963). Then, in the mid-1960's, the Commission began to impose direct restraints on cable carriage of broadcast signals. *See infra* note 63.

[17]*See, e.g.*, First Report and Order, 20 F.C.C.2d 201 (1969) (upheld in United States v. Midwest Video Corp., 406 U.S. 649 (1972)). Other access and origination requirements, which were imposed and later removed during the 1970's, were finally ruled outside of the Commission's jurisdiction over cable television in Midwest Video Corp. v. FCC, 440 U.S. 689 (1979). *See infra* Section II.B.

eral constraints.[18] Municipalities are now the major sources of cable regulation.[19]

Because cable distribution requires the use of public rights-of-way, it is doubtful that a cable television operator could wire a community without first securing a franchise license.[20] Presently, in the absence of preemption by federal[21] or state regulations,[22] a local governmental body is formed to conduct the franchising process.[23] Bids are solicited from all interested cable companies; and, because cable distribution exhibits character-

---

[18]The only remaining federal rules are found at 47 C.F.R. § 76 (1981). They are: (1) section 76.31, a limitation on franchise fees, *see infra* note 26; (2) sections 76.51 to 76.65, the rules on mandatory signal carriage ("must carries"), *see infra* notes 138 & 139; (3) section 76.67, which prohibits importation of sporting events that have been blacked out in the local area; (4) sections 76.92 to 76.99, which prohibit simultaneous importation of duplicative programming; (5) sections 76.205 to 76.311, which impose traditional broadcasting rules, such as the fairness doctrine and equal time requirements, on cablecasters; (6) section 76.501, which limits the cross-ownership of cable and other media; and, (7) sections 76.601 to 76.617, which articulate technical operating standards.

[19]Albert, *The Federal and Local Regulation of Cable Television*, 48 U. COLO. L. REV. 501, 508 (1977) ("Cable's very technology—including the laying of lines, conduits, and cables throughout a given community—in most cases triggers the involvement of the local government."); Pridgen & Engel, *Advertising and Marketing on Cable Television: Whither the Public Interest?*, 31 CATH. U.L. REV. 227, 258 (1982) ("In most cases, the municipality plays the leading role in regulating cable television.").

[20]Nevertheless, 32 cable operators in Pennsylvania claimed that they did not have formal franchise licenses. *See* D. Allen & D. Kennedy, Municipal Regulation of Cable Television in the Commonwealth of Pennsylvania 3-11 (Dec. 1982) (available from the Inst. of Pub. Admin., Penn. State Univ.)

If a municipality awards a franchise to one applicant, it is unclear whether it can keep others out, *see* Henderson, *Municipal Ownership of Cable Television: Some Issues and Problems*, 3 COMM/ENT 667, 668–70 (1981), particularly after Community Communications Co. v. City of Boulder, 455 U.S. 40 (1982); *see also infra* note 25.

[21]*See supra* note 18. A revised version of a bill introduced by Sen. Barry Goldwater (R-Ariz.), S. 66, 98th Cong., 1st Sess. (1983), was being considered by the Senate Subcommittee on Communications at the time this Article went to press. In its most recent form, the result of a compromise between the National League of Cities and the National Cable Television Association, the bill would expressly preempt much of a local government's authority to regulate cable service. That power would be vested solely in the FCC. *Compare, e.g., id.* § 607 (leaving municipalities with the power to set rates on basic cable service, subject to a finding by the FCC that cable service is part of a competitive market and should therefore be deregulated), *with* the arrangement reached in the compromise. *See Cities, Cable Operators Reach Agreement*, 41 CONG. Q. 637 (1983); *NCTA, League Deal Clears Deck for S. 66 Passage*, BROADCASTING, Mar. 14, 1983, at 118 (giving cable operators control over the rates charged for basic service).

[22]State statutes regulating cable vary greatly, from total state preemption, *see, e.g.*, CONN. GEN. STAT. § 16-331 (1981), to partial state preemption, *see, e.g.*, MASS. GEN. LAWS ANN. ch. 166A (West 1977), to minimum state regulation, as in Colorado, where power over cable television is exercised under a home-rule grant, COLO. CONST. art. XX, § 6. *See also* Briley, *State Involvement in CATV and Other Communications Services: A Current Review*, in 2 THE CABLE/BROADBAND COMMUNICATIONS BOOK 35 (M. Hollowell ed. 1980).

[23]Albert, *supra* note 19, at 509–13.

istics of a natural monopoly,[24] the locality usually awards only a single franchise for each area.[25] A city cannot, however, award its franchise to the applicant willing to bid the most dollars for it. An FCC regulation limits the fee a municipality may charge the successful applicant.[26] Instead, selection is based on a combination of factors that may include technical specifications of the proposed system, promised subscription rates, financial

---

[24]*See* E. Noam, *supra* note 1; B. Owen & P. Greenhalgh, *supra* note 1.

[25]Johnson & Blau, *Single Versus Multiple System Cable Television*, 18 J. BROADCASTING 323 (1974). Competitive franchising (where more than one franchisee serves a single locality) usually is caused by disputes about the scope of an initial award, and exists in only approximately eight locales out of more than 4,000. This situation is called "overbuild" in the industry. Only Allentown, Pennsylvania, and Phoenix, Arizona, have competitive franchises of appreciable size. *See* Dawson, *How Safe Is Cable's "Natural Monopoly"?*, CABLEVISION, June 1, 1981, at 333; *see also* FCC Staff, Report on Cable TV Cross Ownership Policies 97 n.28 (Nov. 27, 1981).

The right of a city to grant a single de jure exclusive license is unclear. A California court has held that the state constitution precludes such monopolies in the field of communications. *See* TM Cablevision v. Daon Corp., 6 Media L. Rep. (BNA) 2576 (San Diego, Cal., Super. Ct. 1981). Nevertheless, in a 1982 survey in Pennsylvania, 41.5% of the operators responding claimed that their franchise contracts precluded others from providing cable services. *See* D. Allen & D. Kennedy, *supra* note 20, at 3-15.

Even in the absence of any state legislation, antitrust laws may apply. The Supreme Court recently has held that municipalities are not exempt from the antitrust laws under the "state action" theory unless their creation of a monopoly results from the pursuit of a policy of regulation clearly adopted by the state. Community Communications Co. v. City of Boulder, 455 U.S. 40 (1982). However, the decision in Gold Cross Ambulance v. City of Kansas City, 538 F. Supp. 956, 965 (W.D. Mo. 1982) (exclusive franchise granted by city upheld because state has enacted "a comprehensive system for licensing and regulating ambulance companies"), suggests that the Supreme Court decision may not be applied too narrowly.

Whatever the legal rights of entry for cable operators, most feel that dual franchising benefits neither company. Thus, the cable industry has almost unanimously stayed true to the unspoken code opposing overbuilds. Companies also fear that encroaching on someone else's territory invites retaliation. *See* Dawson, *Overbuilds Subject of Study*, CABLEVISION, Apr. 4, 1983, at 22.

[26]To prevent localities from extracting huge license fees, thus hindering efforts by the cable industry to fund research and development, the FCC has limited the fees that could be charged by municipalities. 47 C.F.R. § 76.31 (1981).

The FCC explicitly limits franchise fees to three percent of gross revenues, but will permit a fee of up to five percent upon a showing by the municipality that such an amount not only "will not interfere with the effectuation of federal regulatory goals," but also "is appropriate in the light of the planned [local] regulatory program." *Id.* This means that the fee must not be unduly burdensome to the franchisee and that the local government must justify the amount in specific terms. *See* Note, *Cable Television: The Practical Implications of Local Regulation and Control*, 27 DRAKE L. REV. 391, 402 n.81 (1977).

Although repeal of this limit has been considered, *see* Clarification of Report and Order, 71 F.C.C.2d 569, 581–85 (1979), the compromise version of S. 66 would adopt the five percent limit. *Compare* the compromise provision, *see NCTA, League Deal*, *supra* note 21 (adopting a flat five percent fee), *with* the original provision for FCC determination of the franchise fee, S. 66, 98th Cong., 1st Sess. § 608 (1983) (authorizing the FCC to establish a "reasonable ceiling").

soundness of the applicant, public access offerings, and frequently politics.[27]

The principal difficulty with this procedure is that it requires political bodies to make highly subjective judgments based on ambiguous multiple standards.[28] There is tremendous pressure on applicants to expend their resources on political maneuvering[29] and to make unrealistic promises concerning rates[30] and technical features.[31] Resources also are normally allocated to public service offerings of the system, without any evaluation of whether the benefits derived from offering these services justify the costs.[32]

B. *Antitrust and First Amendment Problems in the Current Franchising System*

Applicable federal rules,[33] state or local statutes,[34] contractual provisions of the franchise agreement, representations made to potential subscribers, and competitive pressures from other media[35] all constrain the cable operator to a degree. Nonethe-

---

[27]*See generally* R. Posner, *supra* note 11 (discussing the wide variety of factors that can affect a franchise award under the present system). The particular political delicacies of the franchising process are recognized by all players, *see* 4 R. SMITH & R. GALLAGHER, THE EMERGENCE OF PAY CABLE TELEVISION 36–48 (1980), and "rent-a-citizen" and "rent-an-institution" strategies are freely admitted. Both strategies attempt to get either a local citizen or a local institution (often a civic group) involved as an equityholder in a franchise applicant's business. The theory is that such a local affiliation will incline the franchising authority to look favorably upon the applicant. *Id.* at 41–44.

[28]*See* R. Posner, *supra* note 11, at 6, 21–22 (pointing out that this situation requires the franchisor to make a difficult inference about what mix of services and prices viewers will prefer).

[29]This has led cable franchise applicants to spend millions of dollars to market their services to the public and to the franchise board. *See* Schmidt, *Millions Spent in Content for "Showcase" Denver Cable Contract*, N.Y. Times, Feb. 22, 1982, at A12, col. 2.

[30]*See* W.K. JONES, REGULATED INDUSTRIES, CASES AND MATERIALS 28–31 (2d ed. 1976); 4 R. SMITH & R. GALLAGHER, *supra* note 27, at 44–48 (illustrating the implausibility of rate proposals for Dallas); *see also infra* note 43.

[31]*See* 4 R. SMITH & R. GALLAGHER, *supra* note 27, at 39 (recognizing that cable operators undoubtedly often plan to "'[p]romise anything and everything necessary to win the franchise. [Then three years later] go back to the city and renegotiate the real promises you intended to offer all along.'") (quoting Dawson, *The Franchise Story*, CABLEVISION, May 19, 1980, at 83); Stoller, *The War Between Cable and the Cities*, CHANNELS OF COMMUNICATION, Apr.–May 1982, at 34. Teleprompter (now Group W Cable) is one franchisee that has been criticized for failing to live up to the "state of the art" promises that it had made to Manhattan. Schwartz, *Is Cable TV Doing Enough in Manhattan?*, N.Y. Times, Nov. 8, 1981, § 2, at 1, col 1.

[32]*See infra* Section V.

[33]*See supra* note 18.

[34]*See supra* note 22.

[35]Competing video media include movie theaters, traditional broadcast television, and new broadcast technologies (such as direct broadcast satellite service and subscription

less, the operator remains substantially free to exercise its market power while selecting from among competing program suppliers,[36] setting prices for pay channels,[37] and grouping program services into "tiers."[38] This raises a number of antitrust concerns.

As the sole buyer of programs for cable distribution in the area, the operator can distort the marketplace for programming. It can exercise censorial power over the kind of programming allowed on its channels[39] and can favor its own affiliated cable services over those of competitors.[40] The large multiple system

---

television), as well as videotape cassettes and videodiscs. *See* MAJORITY STAFF OF THE SUBCOMM. ON TELECOMMUNICATIONS, CONSUMER PROTECTION, AND FINANCE OF THE HOUSE COMM. ON ENERGY AND COMMERCE, 97TH CONG., 1ST SESS., TELECOMMUNICATIONS IN TRANSITION: THE STATUS OF COMPETITION IN THE TELECOMMUNICATIONS INDUSTRY 342–77, 383–87 (1981) [hereinafter cited as COMPETITION REPORT].

For some appraisals of intermedia competition, see Moozakis, *SMATV: Your Business or Theirs?*, CABLE TELEVISION BUSINESS, Nov. 15, 1982, at 34; Noam, *Towards an Integrated Communications Market: Overcoming the Local Monopoly of Cable Television*, 34 FED. COMM. L.J. 209, 233–41 (1982); Reidy, *An Economic Perspective: The Business of Communicating*, in *1983 Field Guide to the Electronic Media*, CHANNELS OF COMMUNICATIONS, Nov.–Dec. 1982, at 4.

[36]Presently, the cable operator may choose from more than four dozen cable satellite networks. *See, e.g.*, the charts contained in *Field Guide to the Electronic Media*, CHANNELS OF COMMUNICATION, Nov.–Dec. 1982, at 17, 20, 22, 40, 41.

[37]Its position as monopsonist buyer and monopolist seller of cable programming allows the operator to set all prices without violating the Sherman Act's prohibition against price fixing, 15 U.S.C. § 1 (1976), which would forbid the actual distributors of programming (networks) from discussing their pricing practices with each other. Still, it must compete with the media discussed in note 35 *supra*.

[38]Tiers are bundles of channels. Operators commonly offer one low-priced, heavily subscribed basic tier (which includes local commercial broadcast channels), and additional tiers, which are available to subscribers at additional charges (and usually carry specialty services such as all-movie channels). These additional tiers, because of cost, are often less heavily subscribed. *See The Buffalo Shuffle: TMC Gains, HBO Loses 27,000 in Channel Shift*, Multichannel News, Sept. 14, 1981, at 40, col. 1. Tiers are now created at the discretion of the cable operator. *See* S. BESEN & L. JOHNSON, AN ECONOMIC ANALYSIS OF MANDATORY LEASED CHANNEL ACCESS FOR CABLE TELEVISION 30–31, 78–79 (1982).

[39]The Motion Picture Association of America has argued to the Supreme Court that cable's natural monopoly position gives it "virtually unlimited anticompetitive powers." Amicus Brief at 5, FCC v. Midwest Video Corp., 440 U.S. 689 (1979).

A cable operator's right to exclude programmers may, however, be subject to antitrust attack. *See* Channel 100, Toledo, Inc. v. Comcast Cablevision Corp., No. 80-40,071 (E.D. Mich. May 5, 1980) (granting a preliminary injunction preventing a cable system operator from evicting a channel lessee); Mountain States Legal Found. v. City of Denver, No. 81-1738 (D. Colo. filed Nov. 1, 1982) (charging that a de facto exclusive license that permits the cable operator to exclude programmers violates the First Amendment).

[40]This danger arises when a cable operator is vertically integrated into program production. *See* CABINET COMMITTEE ON CABLE COMMUNICATIONS, CABLE: REPORT TO THE PRESIDENT ch. II, at 12, 20 (1974) [hereinafter cited as WHITEHEAD REPORT]; J. Ordover & R. Willig, Notes on Non-Price Anti-Competitive Practices by Dominant Firms (paper presented at the Ninth Ann. Telecommunications Pol'y Research Conference, Annapolis, Md., Apr. 1981).

operators (MSOs) also may implicitly agree to favor each other's programming networks, thereby excluding potential entrants.[41]

A number of committees and commissions have studied the anticompetitive potential of vertical integration between companies providing the distribution equipment (hardware) and those that provide programming (software); all have called for a general separation of the two operations.[42] Cable companies oppose this idea because of the profitability of vertical integration and because of a belief that such consolidation will lead to greater innovation in the industry as a whole.[43] Unless an alternative to separation can be developed to limit a system operator's control of access to its channels, integration may be attacked, despite its alleged benefits.[44]

One example of such a practice occurred when Group W Cable reportedly advised all of its cable systems to refrain from contracting with Cable News Network (CNN), since Westinghouse (Teleprompter's new parent) had plans for its own news network. *See* Noam, *supra* note 35, at 213. CNN has brought an antitrust suit, claiming that the ABC/Westinghouse Satellite News Channels are employing predatory practices in order to muscle CNN out of the market. *Turner Files Antitrust Suit Against SNC*, Multichannel News, Mar. 7, 1983, at 1, col. 1.

Monopoly power also can be exercised through tiering practices. By structuring the tiers of channels, a cable operator can put its own programming in the most favorable first tier (benefiting from cross-subsidies at times) while keeping its competitors in the least favorable tiers. *Cf. The Buffalo Shuffle*, *supra* note 38, at 40, col. 1. The FCC, however, has discounted these fears. *See* In re Teleprompter, 87 F.C.C.2d 531, 554–63 (1981).

[41]For example, Warner Amex Cable Co. could offer to provide channel space on its systems for Viacom's Cable Health Network affiliate if Viacom were to give access to a Warner program service, like Music TV. Programmers not affiliated with system owners would thus be placed at a competitive disadvantage.

[42]*See* SLOAN REPORT, *supra* note 14, at 142; STAFF OF SUBCOMM. ON COMMUNICATIONS OF THE HOUSE COMM. ON INTERSTATE AND FOREIGN COMMERCE, 94TH CONG., 2D SESS., CABLE TELEVISION: PROMISE VERSUS REGULATORY PERFORMANCE 90 (Subcomm. reprint 1976) [hereinafter cited as HOUSE REPORT]; WHITEHEAD REPORT, *supra* note 40, at ch. III, 1–4. For a comprehensive discussion of strict separation, see K. KALBA, SEPARATING CONTENT FROM CONDUIT (1977) (available from Kalba Bowen Assoc., Cambridge, Mass.). For a discussion of a similar issue in the film industry, see Conant, *The Impact of the* Paramount *Decrees*, in THE AMERICAN FILM INDUSTRY 346 (T. Balio ed. 1976) (concerning United States v. Paramount Pictures, 334 U.S. 131 (1948)).

[43]*See* K. KALBA, *supra* note 42, at 31; FCC Staff, *supra* note 25, at 134–38.

[44]Requiring that system operators lease channels to programmers at uniform prices is one way of limiting an operator's control of access without requiring strict separation. The concept of leasing has gained support from many industry observers. "Leased channels are the future of the cable industry," declared Bernard Wunder, Assistant Secretary of Commerce for Communications and Information, at a recent conference on leased channels in Washington, D.C. Gudgel, *Leased Channels Conference Held*, Nat'l Fed'n of Local Cable Programmers Newsletter, Nov.–Dec., 1982, at 1, col. 1; *see also* FCC Staff, *supra* note 25, at 127–40; K. KALBA, *supra* note 42; 4 R. SMITH & R. GALLAGHER, *supra* note 27, at 60–61; Huffman, *Pressure Grows for Law Requiring Leased Access*, Multichannel News, Mar. 8, 1982, at 1, col. 1.

In the telephone industry, AT & T was recently divested of its local operating companies. One effect of the divestiture was to separate AT & T's information-providing

Subscriber fees for cable services presently are set by the cable operator, although the franchise agreement may limit the operator's power substantially by establishing basic rates.[45] Competition from other video media also limits the prices charged by operators; however, if cable becomes the lowest-cost conduit for video services then it may well become the dominant distribution medium over the next decade.[46] In that case, inter-channel (and even intra-channel[47]) price competition will provide the only effective limitation on prices. If the present industry structure is retained, however, such competition is precluded.[48] Some form of rate regulation then would be necessary to constrain the operator's monopoly pricing power. This regulation, however, would impose significant burdens and expenses on both the operator and the municipality.[49]

Because the franchise involved here affects speech, in the form of cable programming, it raises First Amendment concerns as well as antitrust issues. The franchisee, with total control of access to its system, can regulate the speech of programmers by threatening to refuse access. When this franchise is exclusive, the operator's control of access necessarily determines whether or not a cable programmer will be able to speak at all through the cable medium.[50] The Seventh Circuit recently rec-

---

services from its local operating companies, which control access to local phone lines. In a speech at "The Lawyer and the New Video Marketplace" seminar on June 11, 1982, Assistant Attorney General William Baxter suggested that a similar fate may await cable industry giants. *A New Monopoly?*, CABLEVISION, June 28, 1982, at 79.

[45]"[R]ate regulation, as it is currently practiced, does not seriously contain cable operators' decisions . . . . [A]bout 92 percent of the requests by cable operators for rate increases were granted. Moreover, nearly 100% of the amounts requested were granted." S. BESEN & L. JOHNSON, *supra* note 38, at 49–50. Local regulation of pay television is presently preempted by the FCC. Clarification of Cable Television Rules, 46 F.C.C.2d 175, 186 (1974), *aff'd sub nom.* Brookhaven Cable TV v. Kelly, 573 F.2d 765 (2d Cir. 1978), *cert. denied*, 441 U.S. 904 (1979). As for basic rates, "[t]he view that the marketplace can control rates in some markets is gaining momentum." L. LEVINE, THE REGULATION OF CABLE TELEVISION SUBSCRIBER RATES BY STATE COMMISSIONS 94 (1978). Massachusetts already has deregulated some of its cable markets. *See* MASS. GEN. L. ANN. ch. 166A, § 15 (West Supp. 1982–1983); MASS. ADMIN. CODE tit. 207, §§ 6.51–6.54 (1980).

[46]*See* Moozakis, *supra* note 35; Reidy, *supra* note 35.

[47]Intra-channel price competition may be achieved by utilizing pay-per-view. This technology permits each program on a channel to be priced separately. *See* Wirth & Zenalty, *The Economics of Per-Program Pay Cable Television*, 22 J. BROADCASTING 143 (1978).

[48]*See supra* note 37.

[49]*See supra* note 11.

[50]*See* TM Cablevision of San Diego v. Daon Corp., 6 Media L. Rep. (BNA) 2576 (San Diego, Cal., Super. Ct. 1981) (discussed *supra* note 25).

ognized this possibility when reviewing the award of an exclusive cable franchise in Indianapolis.[51] Although the court explicitly claimed not to have decided the First Amendment question in the case, the opinion suggested that the cable franchising process bore a substantial relationship to situations in which public officials seek to impose reasonable time, place, and manner restrictions on public speech such as parades.[52] Relying on Supreme Court decisions covering regulation of speech in public fora, the court held that a franchising statute, to be constitutional, must have "specific criteria" that will guide the franchise decision.[53] While intimating that the grant of an exclusive cable franchise is not per se violative of the First Amendment, the court did remand the case to the district court for a determination of whether the Indianapolis franchising ordinance satisfied the constitutional requirements for regulation of a speaker in a public forum.[54]

In summary, the present system is deficient in at least three major respects: (1) the franchising procedure and FCC limitations constrain direct bidding, which leads applicants to make inefficient and at times illegal[55] indirect bids and promises that are difficult to keep or to enforce; (2) the operator's monopsony power enables it to hinder the efforts of many programmers to gain competitive access to a system, in violation of both antitrust laws and the First Amendment; and, (3) controlling the operator's monopoly power over subscribers requires burdensome and costly rate regulation. COMCAR proposes an equitable and efficient solution to these problems.

---

[51]Omega Satellite Prods. v. City of Indianapolis, 694 F.2d 119 (7th Cir. 1982).

[52]*See* Nadel, *A Unified Theory of the First Amendment: Divorcing the Medium from the Message*, 11 FORDHAM URB. L.J. 163 (1983).

[53]Omega Satellite Prods., 694 F.2d at 128 ("The City of Indianapolis could not [for example] deny a franchise to [the appellant] because [it] carried programs critical of the Republican Party.") (citing Niemetko v. Maryland, 340 U.S. 268 (1951), for the standard of permissible regulation).

[54]In another case, a group in Colorado has sued Denver, charging that the city's award of an exclusive franchise unconstitutionally allows government to determine who may speak using the medium of cable. Mountain States Legal Found. v. City of Denver, No. 83-1738 (D. Colo. filed Nov. 1, 1982). A cable operator also has claimed that the First Amendment prevents a locality from denying it a franchise renewal. Century Cable v. City of San Buenaventura, No. 82-5274 (C.D. Calif. filed Oct. 12, 1982).

[55]*See, e.g.*, Teleprompter Cable Sys. v. FCC, 543 F.2d 1379, 1381 (D.C. Cir. 1976) (discussing use of bribery to gain a franchise award); Barnett, *State, Federal and Local Regulation of Cable Television*, 47 NOTRE DAME LAW. 685, 691–92 (1972) (compiling a variety of instances where illegal methods were used in seeking franchises).

II. COMCAR: Basic Structure and Legal Ramifications

A. *Summary of the Proposed Structure*

Under COMCAR, cable operators act as common carriers.[56] They are limited to providing transmission services to "video publishers" in much the same way that newsstands limit themselves merely to distributing periodicals. As common carriers, they are forbidden from exercising any control over the prices or content of any of the networks or individual programs that they carry. Their sole functions are transmitting programs and billing subscribers. They are deprived of any power to restrict speech or competition.

It is the channel lessees—the video publishers (possibly including the cable operator's company)—who edit the content of programs and set both the subscription rates for their programming services and the advertising rates for sponsors of commercial programming channels that are offered free or at reduced rates to viewers. Marketplace competition alone determines the rates these media publishers face.

The cable operator is required to charge program suppliers a uniform lease rate.[57] It sets this access charge as a percentage of gross revenues that the program supplier collects from subscriber fees and advertising income. The only regulation necessary to ensure competitive prices is the requirement that any unused channels be leased to the highest bidder in a subsequent auction.

---

[56]Common carrier proposals for cable television were frequent in the early 1970's because, as economist Bruce Owen noted, they would allow "the carrier to take advantage of economies of scale in the transmission process, and at the same time provide an opportunity for considerable competition among message sources . . . ." Owen, *Public Policy and Emerging Technology in the Media*, 18 PUB. POL'Y 539, 546 (1970). The most recent efforts have come from Henry Geller, former head of the National Telecommunications and Information Admin. *See, e.g.*, Petition of Henry Geller and Ira Barron to Issue Notice of Proposed Rulemaking (Oct. 9, 1981) (available from the FCC). For a discussion of past efforts to regulate cable as a common carrier—and their failure—see *infra* Section II.B.

[57]Special rates and rules could be established for some groups such as for local broadcasters, *see, e.g.*, 47 C.F.R. §§ 76.51 to 76.65 (1981) ("must-carry" rules), which would grant these groups a zero-price-lease rate, or for candidates for public office, *see, e.g.*, 47 U.S.C. § 315(b) (1981), which was upheld as constitutional in Columbia Broadcasting Sys. v. FCC, 453 U.S. 367 (1981). Unfortunately, the establishment of a mechanism for granting preferences to particular groups might create unacceptable risks of politically motivated abuses in violation of the First Amendment.

All qualified applicants are required to compete for the franchise to ensure that cable operators do not use their dominant position to earn monopoly profits. The franchise license is awarded to the applicant who offers subscribers the highest refund on monthly cable bills. For example, a winning bid of twelve percent allows viewers and advertisers to deduct twelve percent from their monthly bills.[58] Such an auction forces applicants to estimate the so-called "monopoly rent" that they can earn from program suppliers and return this amount to the public. Both program suppliers and cable operators thus are prevented from earning excess profits; competitive bidding distributes such profits to the public.

Finally, short-term franchise licenses are granted, but with a strong renewal preference, so that compliance with the terms of the license would virtually guarantee the operator of renewal. This would ensure continuity of service and encourage the licensee to make long-term investments in the system.

B. *The Jurisdictional Status of Cable Television*

Federal common carrier regulation is presently precluded because of the peculiar way in which the FCC assumed jurisdiction over cable during the late 1950's and throughout the 1960's. In 1959, the FCC determined that it was powerless to regulate cable. The Commission held that cable was not a "common carrier" within the meaning of Title II of the Communications Act of 1934.[59] While the statutory definition of the term "common carrier" is unclear,[60] the Commission felt that because the

---

[58]The amount might, however, be taxed to pay for public service offerings. *See infra* Section V.

[59]Report and Order, 26 F.C.C. 403, 427–28 (1959). Title II is comprised of 47 U.S.C. §§ 201–224 (1976 & Supp. V 1981). It is the Title under which the FCC regulates interstate telephone and telegraph communications and has become a possible source of regulations for new media, such as multipoint distribution service (MDS), regulated by 47 C.F.R. § 21 (1981).

[60]Section 153(h) defines a "common carrier" as "any person engaged as a common carrier for hire, in interstate or foreign communication by wire . . . ." 47 U.S.C. § 153(h) (1976). The FCC does not consider cable to be a "common carrier for hire" as contemplated by Title II. But just what this provision means is unclear; the legislative history is unhelpful, in that it says only, "The definition does not include any person if not a common carrier in the ordinary sense of the term . . . ." H.R. REP. No. 1850, 73d Cong., 2d Sess. 46 (1934). Clearly, this statute did not contemplate cable technology,

system operator selected the signals it carried, it was not acting with the degree of passivity characteristic of a common carrier.[61] The FCC thus allowed cable's market use to define its regulatory status, rather than trying to prescribe a regulatory framework designed to serve the public interest.[62]

Despite this initial reluctance to regulate cable, the Commission began to impose rules on the medium in 1965, regulating it as a competitor of broadcast television.[63] In *Southwestern Cable Co. v. FCC*, the Supreme Court affirmed this form of FCC control, recognizing the FCC's power over cable as that "reasonably ancillary to the effective performance of the Commission's responsibilities for the regulation of television broadcasting."[64]

The *Southwestern Cable* case alone, however, did not articulate the limits of the Commission's jurisdiction over cable.[65] In fact, in 1966, the Court of Appeals for the District of Columbia Circuit had implied that the FCC could choose the most appropriate basis of regulation (common carrier, under Title II, or broadcast, under Title III) to cope with changes in the developing cable industry.[66] This suggests that the Commission could

---

which was not introduced until 1949. For a detailed, historical review of common carriers, concluding that they are firms with "monopoly control over services regarded as essential to the public welfare," see Competitive Common Carrier Rulemaking, 84 F.C.C.2d 445, 520–34 (1981).

[61]Report and Order, 26 F.C.C. at 427–28, *aff'g* Frontier Broadcasting Co. v. Collier, 24 F.C.C. 251 (1958). In the paradigm common-carrier situation, the program supplier would decide whether or not to use the transmission facilities (the cable); the cable would have to be available for hire by any person wanting to transmit and willing to pay the fee set by regulation.

[62]The Commission's public interest obligations are set forth in scattered sections of the Communications Act of 1934. *See, e.g.*, 47 U.S.C. §§ 151, 303(f), 303(h), 303(r), 307(b) (1976) (generally requiring distribution of radio and wire communications in a fair and efficient manner, throughout the nation, as "public convenience, interest, or necessity requires").

[63]First Report and Order, 38 F.C.C. 683 (1965). The assumption of jurisdiction came reluctantly, at the urging of broadcasters, who feared that cable's use of broadcast signals would damage commercial television's financial viability by diverting broadcast viewers to cable. The regulations that the FCC produced were largely an effort to restrict cable's free use of broadcast signals in the absence of a congressional decision to impose copyright liability on cable operators for the use of the broadcast signals. *See* Botein, *CATV Regulation: A Jumble of Jurisdictions*, 45 N.Y.U. L. REV. 816, 826–27 (1970).

[64]392 U.S. 157, 178 (1968).

[65]The agency's jurisdictional reach was seemingly extended by the Supreme Court's decision in United States v. Midwest Video Corp., 406 U.S. 649 (1972) (*Midwest Video I*); but, subsequently, that reach was restricted. *See infra* text accompanying notes 69–71.

[66]Philadelphia Broadcasting Co. v. FCC, 359 F.2d 282 (D.C. Cir. 1966).

The FCC also has recognized the merits of common carrier status for cable. In 1970, it requested comment on a proposal that more than half of the channels of large cable

impose a common-carrier scheme similar to COMCAR as long as it was able to find that the regulatory package furthered the goals of its statutory mandate.

During the 1970's, the Commission pursued a mixture of regulatory schemes. It imposed a series of rules with both broadcast[67] and common carrier[68] characteristics. In *FCC v. Midwest Video Corp.* (*Midwest Video II*),[69] however, the Supreme Court struck down the Commission's common-carrier-like access requirements for cable, ruling that section 153(h) of the Communications Act prohibited the FCC from treating broadcasters like common carriers.[70]

Thus, even though cable is not a typical broadcaster, in that it uses radio frequencies only tangentially to transmit its programming, *Midwest Video II* held that because the FCC had assumed jurisdiction over cable under Title III, cable could not be regulated as a Title II common carrier. Accordingly, as long as the FCC asserts jurisdiction over cable under Title III, the Communications Act, as construed by the Supreme Court, precludes the FCC from imposing a common carrier structure such as COMCAR without legislative action.[71]

---

systems be leased at reasonable, nondiscriminatory rates. Second Further Notice of Proposed Rulemaking, 24 F.C.C.2d 580 (1970). In 1972, the Commission rejected a full common-carrier model, however, because it felt the time premature for such a plan. Cable Television, Report and Order, 36 F.C.C.2d 143, 197 (1972) (commonly known as the Consensus Agreement); *see also* SLOAN REPORT, *supra* note 14, at 148; *supra* note 56.

In his dissent in FCC v. Midwest Video Corp., 440 U.S. 689, 713 (1979) (*Midwest Video II*), Justice Stevens quoted from Columbia Broadcasting Sys. v. Democratic Nat'l Comm., 412 U.S. 94 (1973), claiming that "'the point is that Congress has chosen to leave [the choice of the proper form of regulation for a medium] with the Commission, to which it has given the flexibility to experiment with new ideas as changing conditions require.'" 440 U.S. at 713; *see also* Note, *Access and Pay Cable Rates: Off-Limits to Regulators After* Midwest Video II?, 16 COLUM. J.L. & SOC. PROBS. 591, 600 n.76, 630–31 (1981) (suggesting the Commission might still be able to "characterize cable as a common carrier").

[67]47 C.F.R. §§ 76.205, 76.209 (1981).

[68]The FCC adopted a limited common carrier plan, *see* Rules and Regulations Concerning Cable Television Channel Capacity and Access Channel Requirements, Report and Order, 59 F.C.C.2d 294 (1976), that required systems to provide some special free-access channels, and one leased access channel (with a provision for additional leased access channels when all of its present leased channels were being used during 80% of the time during any consecutive three-hour period for six consecutive weeks). In response to the decision in *Midwest Video II*, the Commission repealed these requirements. 45 Fed. Reg. 76,178 (1980).

[69]440 U.S. 689 (1979).

[70]*Id.* at 705. This decision seemingly overruled the Court's decision in *Midwest Video I*, which had affirmed a set of FCC rules imposing similar obligations on cable operators. The Court, however, distinguished the earlier case, though in a somewhat conclusory fashion.

[71]*But see* Note, *supra* note 66, at 600–01, n.76 (pointing out that the Commission

Whether states may regulate cable as a common carrier is a question that has not yet been answered. In *Midwest Video II*, the Eighth Circuit suggested that the FCC's 1976 access requirements violated a cable operator's First and Fifth Amendment rights.[72] The Supreme Court did not reach these difficult constitutional questions when it affirmed *Midwest Video II*.[73] However, a recently filed New York case, which challenges the constitutionality of state access requirements for cable, may provide an opportunity for decision on a state's constitutional right to regulate cable as a common carrier.[74] Even if state common-carrier regulation were held constitutional, it might be subject to challenge as an interference with the federal regulatory scheme and therefore be preempted.[75]

Even if COMCAR's introduction did require a rewrite of the Communications Act, this would not be undesirable. The current cable rules represent a haphazard attempt to accommodate a new technology within the framework of an outdated regulatory scheme;[76] thus, such a rewrite would be an advantageous, albeit difficult, undertaking in its own right.

---

regulates some radio common carriers under both Title II and Title III); *but see also* Petition of Henry Geller, *supra* note 56, at 25–50 (arguing that the Commission has the power to impose common carrier-like regulations on cable television).

S. 66, 98th Cong., 1st Sess. (1983), if amended as suggested by the terms of the compromise agreement of the National Cable Television Association and the National League of Cities, would amend the Communications Act in a way that probably would make such common carrier regulation of cable service illegal. *See NCTA, League Deal*, *supra* note 21.

[72]Midwest Video Corp. v. FCC, 571 F.2d 1025, 1054–58 (8th Cir. 1977) (holding that in addition to a possible First Amendment problem, the FCC's access regulations might constitute a "taking" without just compensation, in violation of the Fifth Amendment).

[73]*See Midwest Video II*, 440 U.S. at 709 n.19 (noting, however, that the constitutional challenges were not "frivolous").

[74]Comax Telecommunications Corp. v. New York State Comm'n on Cable Television, No. 82-CV-746 (N.D.N.Y. filed July 19, 1982).

[75]*Cf.* Brookhaven Cable TV, Inc. v. Kelly, 573 F.2d 765 (1978), *cert. denied*, 441 U.S. 904 (1979) (FCC can preempt state regulation of pay-cable subscriber charges and leave the subject unregulated); New York State Comm'n on Cable Television v. FCC, 569 F.2d 58 (2d Cir. 1982) (affirming the FCC's preemption of New York's MATV regulation due to interference with federally regulated MDS transmissions).

[76]*See supra* text accompanying notes 59–62.

The FCC recognized that its cable rules were largely unnecessary once Congress enacted Section 111 of the Copyright Act of 1976, 17 U.S.C. § 111 (1976), which required cable operators to pay a compulsory license fee for the right to retransmit broadcast signals. The Commission thus repealed the major components of its remaining cable rules in 1980. Report and Order, 79 F.C.C.2d 663 (1980), *aff'd*, Malrite TV of New York v. FCC, 652 F.2d 1140 (2d Cir. 1981), *cert. denied sub nom.* National Ass'n of Broadcasters v. FCC, 454 U.S. 1143 (1982). It is ambiguous whether cable was ever truly regulated as a technology in its own right, rather than as a mere free user of broadcast television programming. Because cable is now within the copyright structure, future cable regulation must be evaluated in light of this very different circumstance.

The lack of justification for regulating cable as a Title III broadcaster has led cable operators to argue that they are video publishers and are entitled to the same First Amendment rights as newspaper publishers.[77] Such a claim, however, clouds the distinction between the rights of a publisher and those of an editor.[78] Under COMCAR, a cable operator is not recognized as equivalent to a newspaper editor, but rather as analogous to a conduit owner, like the owner of all the newsstands in a locality.[79] When a cable operator originates programming, it enjoys full First Amendment protection against government interference with its messages.[80] When it passively transmits, it is guaranteed the right to operate an economically viable system.[81] Nevertheless, as a mere transmitter it has no right under the First Amendment to exclude others from its wires.[82]

III. AWARDING THE FRANCHISE

A. *Establishing Technical Standards*

A locality that chooses to offer cable television service must first decide whether to allow free entry to all cable companies, or whether to award only a single franchise,[83] possibly to the municipality itself.[84] If cable is treated as a natural monopoly

---

[77]*See* Goldberg, Ross & Spector, *Cable Television, Government Regulation and the First Amendment*, 3 COMM/ENT 577 (1981) (originally a report to Sen. Bob Packwood (R-Or.) by the Nat'l Cable Television Ass'n).

Cable operators also oppose common carrier treatment for three other reasons. First, it would deprive them of the power to fix the consumer prices of their programming services. *See supra* note 37. Second, they fear that common carrier status would facilitate the establishment of rate regulation. *See* FCC Staff, *supra* note 25, at 127–40. Finally, they normally link common carrier treatment to a prohibition against vertical integration by operators into programming. Such a prohibition has been advocated by some for the reasons discussed in note 44 *supra*; still, it is a separate issue from the imposition of common carrier status. COMCAR would treat cable operators as common carriers, but permit them to integrate vertically.

[78]*See* Nadel, *supra* note 52.

[79]*See id.* at 216–23.

[80]*See id.*

[81]*See id.* at 221.

[82]*Cf.* Red Lion Broadcasting v. FCC, 395 U.S. 367, 391 (1969) ("the First Amendment confers no right on licensees to prevent others from broadcasting on 'their' frequencies").

[83]*See supra* note 25; Jones, *Origins of the Certificate of Public Convenience and Necessity: Developments in the States, 1870 - 1920*, 79 COLUM. L. REV. 426, 427 (1979) (contrasting exclusive and non-exclusive licenses).

[84]Municipal ownership of cable television systems has been advocated by many. *See, e.g.*, R. JACOBSON, MUNICIPAL CONTROL OF CABLE COMMUNICATIONS (1977); Com-

and an exclusive franchise is granted, the municipality presumably will set minimum specifications for the system to ensure satisfaction of public needs and financial stability. Similarly under COMCAR, the franchising process would serve to establish system standards as well as to select the franchisee and to determine the value of the franchise license.

A local franchise board would be appointed to solicit suggestions for standards and services from all concerned parties. Competitive pressures would lead each applicant to lobby the board to require those technical features which it could produce more efficiently than its competitors. The board then would select—from the list of features that prospective applicants and other interested parties had offered—those features that would best serve the public interest. All proposed options that were not demonstrated to be wasteful, unreasonable, unrealistic, or simply inefficient, would become part of the minimum specifications of the franchise license. This process would produce a franchise with specifications formulated through negotiation among all participants. If sufficient competitive pressure did not exist, cable consultants might be hired to recruit additional applicants or to evaluate the reasonableness of the proposals presented.[85]

The determination of a reasonable license period and the selection of an efficient renewal procedure are two aspects of the franchising process that have received particular attention from economists.[86] The dynamic nature of the industry suggests that licenses be granted for relatively short periods of time, allowing localities to adjust and to modify franchise standards fairly frequently. This allows franchisors to require and enforce "state of the art" technology provisions in the franchise agreement. The shorter the license period, however, the more frequently there arises the potential problem of transferring the system from existing franchisees to their successors.[87]

---

ment, *Toward Community Ownership of Cable Television*, 83 YALE L.J. 1708 (1974). Nevertheless, public ownership has its problems. According to cable analyst Paul Kagan, it now appears much less attractive, *see* CABLE TV REGULATION 1 (Paul Kagan Assoc. No. 137, Apr. 2, 1981), but a number of cities are still seriously considering the option. *See Public Owners*, BROADCASTING, July 12, 1982, at 66.

[85]Nevertheless, there are problems collecting information during the process of establishing franchise standards. *See* Goldberg, *Competitive Bidding and Production of Pre-Contract Information*, 8 BELL J. ECON. 250 (1977).

[86]*See* R. Posner, *supra* note 11, at 9–11, 22–24; R. SCHMALENSEE, *supra* note 3, at 50–53; Telser, *supra* note 8; Williamson, *supra* note 8.

[87]The tremendous reliance by the cable industry on debt financing, *see supra* note 12, however, requires that initial franchise terms be long enough to allow its lenders to feel

The transfer of goodwill presents an additional problem.[88] The cable operator has no incentive to develop long-term goodwill unless it can recover the benefits of its investment over a sufficiently long franchise term or can receive payment for that goodwill from its successor. Thus, relatively long-term licenses have been granted—in the absence of a practical method for compensating the operator for goodwill.[89]

Under COMCAR, cable television franchise licenses would be treated similarly to the way broadcast licenses are now handled.[90] A licensee would be evaluated at the end of every three-to-five-year franchise term. If it has met all the minimum standards of the franchise agreement, its performance would entitle it to a renewal.[91] But, if it failed to do so, it would be entitled to sell its "renewal preference" to another cable operator that could meet the minimums. The license then would effectively be perpetual (and at the original franchise fee), contingent only upon the satisfaction of the minimum standards of the agreement. This arrangement would provide the cable operator with normal incentives for making long-term investments in both tangible and intangible assets.[92]

---

secure that the operator's main asset will not be lost before the operator has been able to pay off its loans. *See also* Posner, *The Appropriate Scope of Regulation in the Cable Television Industry*, 3 BELL J. ECON. 98, 116 (1972); Williamson, *supra* note 8, at 84–87 (1976). Without such security, cable operation could well require government regulation as a form of insurance. *See* Jones, *supra* note 83, at 510.

[88]The value of the goodwill that one accumulates is recognized as property. *See* Levitt Corp. v. Levitt, 593 F.2d 463, 468 (2d Cir. 1979) (citing Hanover Star Milling Co. v. Metcalf, 240 U.S. 403, 412–13 (1916)).

[89]Williamson, *supra* note 8, at 90. In fact, until 1977, the FCC rules set 15 years as the suggested franchise length. Report and Order, 66 F.C.C.2d 380 (1977); 47 C.F.R. § 76.31(a)(2) (1981).

[90]47 U.S.C. § 301 (1981) specifically states that "no . . . license shall be construed to create any right, beyond the terms, conditions, and period of the license." Nonetheless, licenses granted to broadcasters are effectively granted in perpetuity. *See* Coase, *The Federal Communications Commission*, 2 J.L. & ECON. 1, 12–17 (1959) ("Even before the 1927 Act was passed, it was recognized that stations were transferred from one owner to another at prices which implied that the right to the license was being sold.").

In fact the FCC's present policy expressly includes a renewal preference as one of the many factors considered at renewal time, where "the strength of the expectancy depends on the merit of the past record." Cowles Broadcasting, 86 F.C.C.2d 993, 1012 (1981), *aff'd sub nom.* Central Fla. Enter. v. FCC, 683 F.2d 503 (D.C. Cir. 1982). It has been intimated by the Supreme Court that this degree of renewal preference is permissible. *See* FCC v. Nat'l Citizens Comm. for Broadcasting, 436 U.S. 775, 782 & n.5, 805–07 (1978).

[91]S. 66, 98th Cong. 1st Sess. § 609 (1983) is just such a provision. It would virtually guarantee renewal if a cable operator performs well—observing all of its contractual duties—and has a reasonable plan for future services.

[92]This renewal expectancy should at least partially alleviate a cable operator's fear that a municipality will appropriate its goodwill by refusing to renew its franchise and instead award it to a local group. In Hugo, Okla., Cablecom General, after serving the

B. *Selecting the Franchisee*

Once the local board reviewed all suggestions and comments, it would promulgate a franchise license that incorporated all desirable minimum standards. Applicants then would bid against each other for this license by offering subscriber discount rates. This procedure would standardize the terms of the bid and thus transform a complex, multi-dimensional auction into a more manageable, one-dimensional process.[93]

Franchise applicants would, however, have to be able to calculate their expected income streams before they could determine their maximum franchise bids. This is because under COMCAR their revenues come almost exclusively from channel lease charges. Basic monthly fees, which provide operators with the major portion of their revenues under current franchising arrangements, would be limited in COMCAR to nominal amounts, if any. COMCAR would, however, continue to permit an installation charge. Since lease charges would be derived from revenues earned by the channel lessees (from subscriber and advertiser fees), the size of these revenues must now be examined.

---

city for 20 years, was denied a renewal and the franchise instead was awarded to a group of local businessmen. *See* Albert, *supra* note 19, at 517–20.

See also Pridgen & Engel, *supra* note 19, where the authors note that "[t]he public franchise differs from ordinary contracts in that the grantor sovereign is not precluded from subsequently enacting legislation under its police power that regulates the exercise of the franchise so long as it does not interfere substantially with the main object of the franchise." *Id.* at 259 (citing Pearsall v. Great N. Ry., 161 U.S. 646 (1896); New Orleans Gas Co. v. Louisiana Light Co., 115 U.S. 650 (1885)).

As the Supreme Court has clearly stated:

> [I]t is settled that neither the "contract" clause nor the "due process" clause has the effect of overriding the power of the state to establish all regulations that are reasonably necessary to secure the health, safety, good order, comfort, or general welfare of the community; that this power can neither be abdicated nor bargained away and is inalienable even by express grant; and that all contract and property rights are held subject to its fair exercise.

Atlantic Coast Line R.R. v. Goldsboro, 232 U.S. 548, 558 (1914).

[93]The cost of using such a standardized bidding process, requiring only minimum technical and financial standards, is that cable operators offering proposals with superior technical and financial features (that are realistically attainable) would not receive an advantage unless they could convince the local cable board to require such superior qualities as part of the minimum standards. Still, this problem is tolerated in most standard government contract bidding: the weapons industry provides a classic example. *See* M. PECK & F. SCHERER, THE WEAPONS ACQUISITION PROCESS: AN ECONOMIC ANALYSIS (1962).

IV. COMCAR IN OPERATION

A. *Subscriber and Advertiser Rates*

Each programmer would have complete control over the contents of its leased channels, the subscription rates it charged to viewers, and the advertising rates it set for sponsors—just as print publishers control the prices and content of their various publications. As suppliers of video programming, they would face intra-medium as well as inter-media competition from other video publishers.[94] Thus, there would be no need to regulate the channel fees they charged viewers.

There would be no danger of antitrust violations occurring through monopolisitic tying arrangements[95] or the preemption of interchannel price competition[96] because the cable operator would lack the power to interfere in the pricing decisions of programmers. The monthly charge collected by the cable operator would be simply the sum of the subscription rates for all of the channels (or individual programs[97]) selected by the viewer-households.[98] Billing by tiers (groups of channels selected by the operator and priced as a single package) would be prohibited; but, to make billing practical, low-priced commercial networks would be allowed to group themselves into bundles or tiers.[99]

These prices still would not represent the true competitive market rates because they would include the cable operator's monopoly rent, the amount that it could inflate its lease rate above its actual distribution cost.[100] Eliminating this distortion would require the cable operator to deduct its own appraisal of this monopoly rent (as determined by its bid in the initial franchise auction) from all programming bills sent to consumers.

---

[94]*See supra* note 35.

[95]This is achieved through tiering. *See supra* note 38.

[96]*See supra* note 37.

[97]*See supra* note 47.

[98]Bills for programming could resemble telephone bills. *See* Rosenfeld, *Let the Seller Beware!*, TVC, July 15, 1982, at 68 (advocating an unbundled approach as an excellent marketing device).

[99]*See supra* note 38.

[100]According to the economic theory of monopolies, the operator would be able to charge lessees a rate that included a monopoly rent charge which would be exactly the value of the franchise license.

This deduction, expressed as the successful applicant's percentage bid for the franchise, would precisely equal the amount that subscription and advertiser rates were inflated above cost as a result of the monopoly power of the cable operator. Thus, if the relatively low cost of video distribution over cable were to enable the operator to undercut the prices of competing video distribution media by, for example, eight percent (after including its own *reasonable* profit), the operator would be able to earn a monopoly rent of eight percent of gross revenues. The franchise license thus would be worth this eight percent; and, a competitive franchise process would ensure that the successful franchise applicant bids eight percent for the license. Viewers and advertisers, in turn, would get an eight percent discount on their monthly bills. Both would be billed initially by programmers at an inflated price (108% of cost); but, they would receive an eight percent rebate from the operator. They would thus pay a price reflecting only 100% of actual distribution costs.

B. *Selecting Channel Lessees (Minimum Lease Rate)*

As a common carrier, the operator would be required to announce a two-part tariff publicly. The components of this would be a minimum charge (P_{min}) and a percentage lease rate. Channels would be available to all those who agreed to pay the operator P_{min}, thus ensuring that all of the most valued programmers gained access.[101] The cable operator would maximize its profit by setting P_{min} at exactly the market clearing price (where the number of lessees willing to pay the price exactly equalled the number of available channels—see Figure 1 below). COMCAR would ensure this by stipulating that if the minimum price were set above that "competitive" level,[102] leaving channels

---

[101]Estimated subscriber and advertising revenues would serve to value most programming, and public and private grants and donations would allow merit programming to be considered on an equal basis. Some might claim that this marketplace system would permit only the wealthiest programmers to gain access, but as Ronald Coase has pointed out:

> [R]esources do not go, in the American economic system, to those with the most money, but to those who are willing to pay the most for them. The result is that, in the struggle for particular resources, men who earn $5,000 per annum are every day outbidding those who earn $50,000 per annum.

Coase, *supra* note 90, at 19.

[102]The cable operator might do so if it wanted to exclude competitors of its affiliated program channels. *See* B. OWEN, J. BEEBE & W. MANNING, TELEVISION ECONOMICS 83–88 (1974).

vacant, these channels would have to be offered to programmers in a special auction administered by the ongoing governmental cable monitor. The law of supply and demand would guarantee that this auction terminated at exactly the market clearing price (P_{min}).

To provide opportunities for smaller entrepreneurs, the cable operator would be required to lease channels in the smallest practical time segments.[103] Programmers, however, would be permitted to lease entire channels to facilitate marketing efforts and to preserve valuable audience flows.[104] The duration of all leases could either be established in the franchise agreement or governed only by normal antitrust provisions prohibiting restraints of trade.[105] There would seem to be no reason to prohibit subleasing or joint leasing.

C. *Distribution of Channel Revenues (Percentage Lease Rate)*

Although the cable operator would charge the same minimum fee to all channel lessees, the majority of its revenues, under COMCAR, would come from the percentage lease rate, expressed as a percentage of the gross revenues of the programmer. If, for example, an operator were to set P_{min} at $30,000, and the percentage lease rate at thirty percent of gross revenues, the price to each programmer would be either $30,000 or thirty percent of gross revenues, whichever was greater. This lease rate would permit the operator to recover its costs in proportion to the benefits that lessees received from the use of the system, as approximated by the lessee's revenues from subscriber charges and advertising fees.

Under COMCAR, channel revenues collected from viewers and advertisers are distributed according to Figure 1 below. As the graph shows:

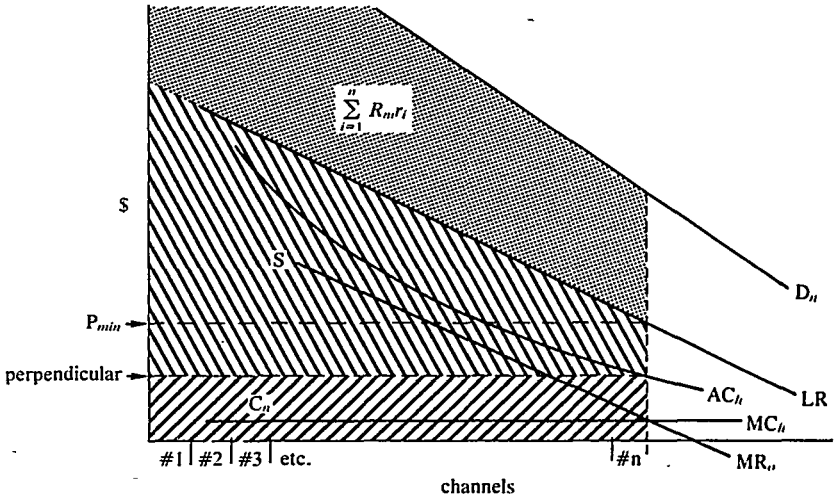
---

[103]It is debatable whether the cable operator could require programmers to enter only into leases for full-time control of an entire channel. *Cf.* United States v. Jerrold Elecs., 187 F. Supp. 545 (E.D. Pa. 1960), *aff'd per curiam*, 365 U.S. 567 (1961) (holding that a sale which required the purchase of a complete cable system would not violate antitrust prohibitions against tying only for as long as the cable industry was considered to be in a developing stage).

[104]Audience flow is the phenomenon of viewers staying tuned to the same channel unless they have a definite reason to switch. *See* E. EPSTEIN, NEWS FROM NOWHERE 93–100 (1973); Owen, *Structural Approaches to the Problem of Network Dominance*, [1979] DUKE L.J. 191, 224, 226.

[105]*See, e.g.*, section 1 of the Sherman Act, 15 U.S.C. § 1 (1976).

Figure 1 *Determination of channel capacity and allocation of revenues collected by the cable operator.*



All revenue amounts represent the present discounted value of the revenue streams over the franchise period and R_{ni} stands for the revenues anticipated by the programmers using cable channel i of an n-channel system. Revenues include both advertiser and subscriber fees.

- D_n derived demand by lessees for channels on an n-channel cable system: $\Sigma_{i=1}^{n} R_{ni}$.
- LR channel percentage lease rate.
- MC_h marginal cost of the cable system hardware.
- AC_h average cost of the cable system hardware.
- MR_o marginal revenues of the cable operator.
- n optimal number of channels (where $MR_o = MC_h$).
- P_{min} minimum fee charged to channel lessees.

represents the SOFTWARE COST: $\Sigma_{i=1}^{n} R_{ni}r_i$ where r_i is the percentage of revenues (or return on revenues[106]) that program distributors leasing channel i require to pay program producers a rental fee and also cover their own costs.

represents the HARDWARE COST: C_n. It is the present discounted value of the costs of building and maintaining the cable system during the franchise period, including a fair return on equity.[107]

represents SURPLUS: S. It is the remainder of the revenues after the hardware and software costs have been paid. On the graph it is the area between the LR curve and the AC perpendicular (minus any overlap). It is the monopoly value of the franchise license and as a percentage of the area under the demand curve (from channel #1 to #n) it represents the upper limit for franchise bids.

[106]Although required return is initially calculated as the opportunity cost of the capital investment, the multiple distribution outlets for programming requires producers to translate their required return on investment into a return on revenues. *See infra* note 113. In the motion picture industry, this return, quoted to film distributors as a rental fee, is given in terms of a percentage of gross revenues. *See* Guback, *Theatrical Film*, in B. COMPAINE, WHO OWNS THE MEDIA? 199, 219, 226 (2d ed. 1982) (discussing the economics of the film industry).

[107]The problem of how to treat and what to do with the cable hardware when the franchise lease expires so that investment incentives are not distorted is discussed *supra* notes 89–92 and accompanying text.

(channel revenues) − (SOFTWARE COSTS) − (HARDWARE COSTS) = SURPLUS, or:

$$\text{Equation 1: } \sum_{i=1}^{n} R_{ni} - \sum_{i=1}^{n} R_{ni}r_i - C_n = S \geq 0$$

This type of discriminatory pricing[108]—where the price paid varies directly with the benefits received—occurs infrequently because it is usually difficult, if not impossible, both to ascertain the value placed on a good by each individual buyer and to prevent resale.[109] The resale problem is especially troublesome because an individual charged a low price for a good can normally resell it to someone willing to pay a greater price. This prevents the original seller from being able to collect the full value of the sale.

These difficulties, though, do not arise in the leasing of cable channels. First, the value each programmer places on a channel is approximately proportional to its revenues and is thus fairly easy to compute. Second, even if a programmer were to resell its channel, the operator would still receive either P_{min} or a fee representing the actual value of the usage (i.e., the lease rate), whichever was greater, from the programmer actually using the channel.

Permitting the operator to set discriminatory prices also would have two significant advantages over imposing the regulated rate alternative. A monopolist such as a cable operator, with marginal (or variable) costs significantly less than average total costs, could not sustain prices at marginal cost. To permit the huge fixed cost of building the system to be recovered, a regulator would be forced to set access rates at the average cost of each channel (derived by dividing total costs by the number of channels in the system). This would force the regulator to un-

---

[108]Price discrimination occurs when a seller charges individuals according to what they are willing to pay, rather than a single, uniform, "non-discriminatory" price. Such conduct does not include favoring one buyer over another based on irrational or improper criteria, as does race or sex discrimination. *See* Baumol & Bradford, *Optimal Departures from Marginal Cost Pricing*, 60 AM. ECON. REV. 265 (1970).

This discriminatory price scheme would not violate the Robinson-Patman Act because the act applies only to the sale of goods, not to leases of cable channel space. *See* Export Liquor Sales v. Ammex Warehouse Co., 426 F.2d 251 (6th Cir. 1970), *cert. denied*, 400 U.S. 1000 (1971).

[109]Posner, *supra* note 1, at 571; *see also* Competitive Carrier Rulemaking, 77 F.C.C.2d 308, 336 (1980) ("In order to sustain a profitable price discrimination scheme, three essential elements must be present: (1) the entity must possess market power; (2) the entity must be able to segregate its customers into groups with differing demand-elasticities, and (3) arbitrage must be prevented.").

dertake the burdensome and expensive task of allocating the huge fixed cost over the duration of the license (i.e., amortizing high joint and common costs, a particularly difficult task in an industry experiencing such rapid, technological change). More importantly, though, this regulated price would prevent programmers able to pay only marginal cost from gaining access to the system.[110]

Discriminatory pricing, on the other hand, would enable the operator to offer access to lessees who could afford to pay only the marginal cost. This would be accomplished without any regulatory expense and without forcing the operator to lower its prices for those able to pay more.[111] Also, because the franchise bidding scheme would return the monopoly rent value of the system to subscribers, discriminatory pricing would not yield excessive profits for the operator.

Lease rates based on a percentage of revenues also would be easy to administer because the value of the channel to the lessee is reflected by the revenues that it earns from subscriber and advertising fees. The fees paid by subscribers are tabulated by the operator itself for billing purposes, and the advertiser fees would simply be the amount negotiated between the programmer and advertising agencies. If the operator believed that the advertising fees reported to it were understated, it could choose to collect its lease fee directly in the form of a percentage of the commercial minutes provided (which it then could resell).

One major objection to this uniform percentage lease rate[112] is that it would force individual lessees with very different cost structures to pay the cable operator identical lease rates. Certainly it would be more accurate and efficient to charge lessees according to net profits rather than revenues; but, this solution would seem prohibitively expensive, if not impossible, to administer.[113]

---

[110]The inability to provide such access is the central criticism of leased access made in S. BESEN & L. JOHNSON, *supra* note 38, at 71, 91.

[111]Posner, *supra* note 1, at 564–65.

[112]Just as the Robinson-Patman Act, 15 U.S.C. §§ 13, 21 (1976), permits sellers to meet the price of competitors, so COMCAR would permit cable operators to meet the lease rates offered by alternative media. This might be done by allowing the operator to structure its rates as a regressive schedule of prices, e.g., 50% for the first $20,000 in revenue, 30% for the next $20,000, and 5% of the amount over $40,000. Alternatively, operators would be able to subsidize those lessees through express "grants."

[113]If easily ascertainable expenses could be netted, that would of course be desirable and permit more efficient pricing and allocation. Net profits, however, would have to be determined by allocating the programmer's production costs among different distri-

In establishing its percentage lease rate, the operator, presumably, would be unable to charge more than competing video media.[114] Yet, from Equation 1 it can be seen that the cable operator would be forced to charge exactly that amount because it would have to recover its costs (C_n) and its franchise fee (S).[115]

A high percentage lease rate still appears to be prohibitively expensive to programmers who could afford the minimum but whose production costs require that they retain a high percentage of their gross revenues. These programmers could seek grants from private or public sources or from the cable public interest fund discussed below; more importantly, they could seek grants from the cable operator itself. As the operator would be assured of lease payments of at least its minimum fee from these services, it presumably would be willing to offer the sufficiently large grants to keep them as paying lessees. Its alternative would be to lease the channels to service suppliers who could not afford even P_{min}.

Admittedly, such grants could be used by the operator to favor programs that it preferred over those that it found displeasing. As there is no law against spending to promote one's viewpoint,[116] though, and as all parties who could afford the market rate minimum would be assured access, there would appear to be little danger of any violations of the First Amendment.[117]

The most powerful channel lessees would not be prevented from demanding grants also, especially if the single percentage rate was greater than what a competitive market would dic-

---

bution media such as movie theaters, network and syndicated television, and videodiscs or videocassettes. The film industry has adopted the only practical solution. *See supra* note 106.

[114]*See supra* note 35.

[115]"S" also represents the monopoly rent of the license or the cost advantage that cable distribution would have over competitors. Therefore (C_n + S) would represent the cost of cable's most efficient competitor. The operator would be prevented from charging an exorbitant rate because, in addition to attracting competing media entry, he could face retaliatory tactics from disgruntled lesees. These lesees could lease channels at the minimum price (P_{min}) and then offer free programming to viewers, merely to harm the operator by drawing viewers away from other channels and decreasing the revenues earned by the operator.

[116]In Buckley v. Valeo, 424 U.S. 1, 48–49 (1976), the Supreme Court held that the First Amendment does not permit restrictions to be placed on the speech of some in order to enhance the relative voice of others.

[117]The First Amendment would appear to forbid the government from granting economic preferences to speakers based on their viewpoints, whether directly or indirectly, through designated licensees. *See* Nadel, *supra* note 52, at 168.

tate.[118] Despite the possibility of such "subsidies," the pricing structure still would serve to frustrate any vertically integrated cable operator seeking to favor its own programming networks by providing them with access at disguised, but implicitly predatory, rates.

Under COMCAR, any predatory pricing would be clearly visible. If the system operator or any other firm were to try to lease channels merely to exclude others, its losses would be easy to document. Lease fees would be public knowledge, and the lessee's revenues could be computed easily by totaling the subscription rates billed (and conveniently stored) in the operator's computer and the advertising revenues billed (and relatively easy to corroborate in the marketplace).[119] The Justice Department or competing suppliers could force any suspected predator to justify its losses.

D. *Expansion of Channel Capacity*

In the ideal world, cable systems would provide programmers with new channels whenever the benefits of the new service exceeded the costs of operating the channel. Unfortunately, however, as there are many ways to define costs and benefits, it is difficult to define "optimal" channel capacity.[120]

Presently cable operators measure the benefits of a new service by the additional revenues generated from subscribers, suppliers,[121] and advertisers.[122] The addition of a new service may, however, only fractionalize the existing audience, rather than add new viewers. In fact, increased price competition from

---

[118]E.g., if the rate was not sufficiently regressive.

[119]Unfortunately, disguised subsidies would still be possible: (1) if a firm were to produce two "pay" channels of programming, e.g., Time, Inc.'s HBO and Cinemax, and offer subscribers a special price for the pair and then intentionally misallocate the revenues to show that neither was being offered merely to block the entry of others, or (2) if a similar misallocation were produced by a single firm with two advertiser-supported channels, e.g., Turner Broadcasting's WTBS and CNN, by offering a special rate to advertisers who sponsored programs on both channels.

[120]This is similar to the debate over the optimal diversity of television programming. *See* Noam, *supra* note 35, at 222–24.

[121]Many of the newer satellite networks now are offering to pay cable operators for access: e.g., National Spanish Television (SIN) pays 10 cents per subscriber with a Spanish surname. S. BESEN & L. JOHNSON, *supra* note 38, at 31; *Programmer Compensation Plans*, TVC, Nov. 1, 1982, at 54.

[122]Advertiser-supported networks may offer cable operators commercial time to sell as part of the network affiliate contract. S. BESEN & L. JOHNSON, *supra* note 38, at 31–33.

additional channels could actually lead to a reduction of total system revenues. This would be true for both advertiser-supported and subscription ("pay") networks.

From an economic perspective, an increase of services, while perhaps aesthetically desirable, might not be economically efficient unless the new service could increase the total revenues of the system. Thus, what a viewer perceives as an "optimal" level of service is likely to differ significantly from the dollar-conscious evaluation of the system operator. For the system operator, the determinant will be whether the revenue generated by a new service comes from the redirection of money already being spent on cable services or from the influx of "new money," not previously used to pay for cable television.

Even from the cable operator's perspective, however, the determination of an "optimal" level becomes difficult because of the existence of hidden costs. For example, even if the operator were to find a service to be profitable initially, these profits might actually lead to a long-term loss of revenues. Increased profits could be cited by a regulatory agency to justify denying future rate increases for the system's "basic" rates and could even provoke a demand for rate decreases.

COMCAR provides an incentive for expansion of channel capacity that can be considered economically optimal if the marginal societal benefit of an additional service is defined as the additional amount that subscribers will spend overall, given the opportunity to choose the new service in addition to any of the already existing channels. Their expenditures, thus, would be used as a proxy for their valuation of the benefits provided to them.

This definition assumes that the marginal benefit produced by the new service comes from the expenditure of "new money," and not from the effects of fractionalization.[123] If the marginal benefit of an additional channel were defined instead as the revenues generated by that channel alone, regardless of the source of the reveunes, COMCAR would not provide an incentive for expansion to optimal capacity. If this latter definition, which most closely approximates the viewer's conception of optimality, is adopted, COMCAR must be supplemented by the establishment of a specific channel capacity in franchise awards,

[123]If fractionalization actually produced an increased benefit to society the specialized services normally would be able to charge higher prices (attracting "new" money).

during renegotiations, and possibly, during the franchise term itself.[124]

V. Public Interest Options

Present franchise agreements normally tax away the franchisees' monopoly rent by requiring that a cable operator offer universal service,[125] provide institutional and public access channels,[126] and build studios for local production,[127] in addition to paying a franchise fee.[128]

COMCAR does not automatically favor in-kind services, but rather returns monopoly profits to the public directly—in the form of cash.[129] If the government and franchising authorities think that a service is valuable, they must act explicitly by imposing a tax. This alternative forces the costs and benefits of each service to be carefully—and publicly—weighed. This method seems preferable because the actual costs of these services are ultimately borne by subscribers,[130] who deserve to know how much they are paying.[131] Presently, costs are never

---

[124]The franchise agreement might, for example, require that additional channels be added whenever the present capacity was "full."

[125]*See* 47 C.F.R. § 76.31 (1981) (FCC Recommendation).

[126]*See* Note, *supra* note 66, at 613–30 (1981).

Lawsuits recently have been filed claiming that public access rules violate the First Amendment rights of cable operators. *See supra* note 74 and accompanying text; *see also* Century Cable v. City of San Buenaventura, No. 82-5274 (C.D. Cal. filed Oct. 12, 1982).

Access requirements, however, still exist in many states. *See, e.g.*, Rules Governing Community Antenna Television Sys., R.I. Pub. Utils. Comm'n, Div. of Pub. Utils. & Carriers, ch. 14 (Jan. 13, 1981) (Order No. 10,369) (available from the R.I. Pub. Utils. Comm'n) [hereinafter cited as R.I. Cable Rules]. The access requirements for cable systems in California, Connecticut, New York, and Rhode Island are discussed generally in Note, *supra* note 66, at 621–30. Rhode Island, for example, presently requires cable operators to provide a minimum of 20% of their channels for such access, assuming there is sufficient demand. R.I. Cable Rules § 14.1(a).

[127]Rhode Island, for example, specifically requires the operator to provide a full-color studio and two portable videotape units. It also requires the provision of personnel to train and assist certain institutional programmers such as schools and hospitals. R.I. Cable Rules §§ 14.2, 14.3.

[128]*See supra* note 26.

[129]Limited use of special cable options has led some to conclude that "most people want cable T.V. to provide a clear picture, extensive movies and sports coverage, and little else." MacKenna, *The Cabling of America: What About Municipal Ownership?*, 70 Nat'l Civic Rev. 307, 309 (1981).

[130]*See* R. Posner, *supra* note 11, at 18.

[131]The N.Y. State Comm'n on Cable Television, however, ordered a Long Island cable system to stop listing a one dollar per month charge for regulatory costs on each subscriber's bill. *Franchise Fees Flap*, Cablevision, Sept. 6, 1982, at 20.

explicitly stated, and benefits are—without careful study—often assumed to exist.[132]

The cable operator might be allowed to use its own discretion and service only those households that contracted to spend some minimum monthly amount on cable programming. Alternatively, public policy might mandate free universal service, subject only to a minimum installation fee or possibly a carrying charge to cover the administrative costs of such service.[133] Nevertheless, to impose a universal service obligation on cable operators and not on their competitors in the communications marketplace may be unfair. If cable operators are to be singled out for this responsibility, they deserve to be provided with some source of "excess" revenues to finance the subsidy. If operators are expected to overcharge some subscribers to enable them to cross subsidize others, they must be protected against competitors who are not required to overcharge and thus can undercut the cable franchisee's higher prices in profitable markets.

One method of providing the necessary protection would be to exclude all competitors.[134] But in the face of the failure of such restrictive measures in the telephone industry,[135] it would seem better to impose a special tax on all distributors who did not offer universal service. Alternatively, a locality could eschew cross subsidies and directly subsidize such service with general tax revenues or funds raised through an "entertainment" tax that applied to all distributors of video programming.[136] Pro-

---

[132]In Midwest Video Corp. v. FCC, 571 F.2d 1025 (8th Cir. 1977), *aff'd*, 440 U.S. 689 (1979), the Eighth Circuit did not feel it could find the access rules to be in the "public interest." 571 F.2d at 1046.

According to R. Peterson & A. Pierce, *supra* note 11, only 3.7% of the access time available on the nine public access channels studied was programmed.

[133]Cablevision, for example, has contracted to provide Boston subscribers with 52 channels of basic service for two dollars per month. Multichannel News, Aug. 24, 1981, at 1, col. 1.

[134]The states were almost unanimous in taking this position with the telephone industry. They recognized that the introduction of competition would prevent companies from overcharging for some services in order to subsidize low-cost universal service. This issue is discussed in Southern Pac. Communications Co. v. American Tel. & Tel., No. 78-0545, slip op. at 26–38 (D.D.C. Dec. 21, 1982).

[135]In the face of explosive technological changes, the FCC was pressed gradually to remove the entry barriers from many markets, permitting new competitiors into such areas as telephone equipment and long distance service. *See* Noam, *Federal and State Roles in Telecommunications: The Effects of Deregulation*, 36 VAND. L. REV. (forthcoming in 1983).

[136]Burbank, Ill., has adopted an entertainment tax. Huffman, *City Imposes Entertainment Tax on Top of Cable Franchise Fee*, Multichannel News, July 5, 1982, at 17, col. 1.

duction studios and institutional and public access channels should also be publicly evaluated, as they presently appear to be underutilized.[137]

Finally, it is necessary to consider the impact of COMCAR on present broadcasters and on those viewers not subscribing to cable. Under COMCAR, absent the must-carry rules,[138] local broadcasters would be forced to pay a substantial portion of their advertising revenues to cable operators for the use of a second distribution medium that they now receive at no charge.[139]

If networks leased channels directly from cable operators, local broadcasters would see the value of their "exclusive" distribution rights approach the value of a single channel on a cable system. In addition, their advertising revenues and network fees would decline to reflect the down-scale (non-cable) audience that they would then serve. Broadcast license holders might then be deserving of some compensation.

A partial solution might be to subsidize local broadcasters by allowing them a preferential channel lease rate, perhaps fifty percent of the ordinary rate. This cross subsidy would create an incentive for them to continue to serve non-cable homes and for networks to use them as intermediaries when dealing with

---

[137]*See supra* note 132. In general, a more efficient utilization of resources might be encouraged if there were greater public awareness of these expenditures. This could be accomplished if, in lieu of obligating the operator to provide specific services fixed at the time the franchise contract was signed, the franchisee were required to set aside the monetary value of such services for more discretionary public use. A locality might even impose the aforementioned entertainment tax for this purpose as well. These funds would be segregated from general government revenues and earmarked for use only in connection with media-related public services. *See, e.g., Warner Amex Gives $70,000 to City for Creation of Local Channels*, TVC, June 1, 1981, at 60. California has established a Foundation for Community Service Television. CALIF. GOV'T CODE § 53066.1 (West Supp. 1982). *But see* Harris, *Local Government Programming 'Boon' a Bust, Critics Claim*, L.A. Times, Feb. 28, 1982, § 4, at 1, col. 6. COMCAR's funds would not be administered statewide, but rather on a system-by-system basis similar to the decentralized structure of the New York City franchise plan. *See* Proposed Cable Television Franchise Agreement & Appendices app. F (Jan. 5, 1983) (available from the N. Y. City Bd. of Estimate).

[138]47 C.F.R. §§ 76.51–76.65 (1981). These rules require carriage of "significantly viewed" (as defined in § 76.5(k)) local broadcast signals at no charge. These signals cannot be unbundled or charged for individually. Clear Television Corp., 46 F.C.C.2d 74 (1974).

[139]Cable operators and networks are seeking repeal of the must-carry rules, but they face stiff competition from broadcasters. *See* Metz, *Must-Carry War Between Cable, Broadcasters Escalates*, CABLEAGE, Feb. 14, 1983, at 19. A particularly vocal opponent of the rules is the president of the Black Entertainment Television (BET) cable network, who feels that the rules have denied his network access to much of its potential market. Friedman, *How Many Must-Carries Must a Cable System Carry?*, VIEW, Aug. 1982, at 38, 41.

cable operators. This would slow the diminution of value of their broadcast licenses. Local broadcasters would survive as long as their operating revenues exceeded their operating costs, after lump sum capital losses had been written off.[140]

VI. Conclusion

By preserving the dynamics of a competitive marketplace within the franchise monopoly, COMCAR would transform the monopolistic/monopsonistic cable distribution medium into a fragmented marketplace resembling the print media. Although this would add the uncertainties inherent in a highly competitive marketplace to an already speculative industry, it would appear to serve the best interests of all parties who favor a competitive environment for business and an unregulated marketplace of ideas. COMCAR appears flexible and efficient and is both applicable to and practical for franchises being awarded and renegotiated today.

---

[140]*See* P. McAvoy, Deregulation of Cable Television 33–39 (1977).

COMMENT

REAGAN ADMINISTRATION HEALTH LEGISLATION: THE EMERGENCE OF A HIDDEN AGENDA

MICHAEL G. MICHAELSON*

The Reagan administration, under increasing criticism from friend and foe alike for failing to deliver on its promises to reform federal health programs,[1] has begun to act on health-planning and cost-control legislation, albeit somewhat indirectly. Unlike the comprehensive proposals of recent years,[2] the Reagan initiatives on the surface appear unambitious. Indeed, they are contained in documents the titles of which tend to obscure their health-planning objectives rather than to advertise them. The first set of such initiatives is found in a law passed with remarkable speed: the Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA).[3] The second set of provisions is embedded in the budget of the United States government for fiscal year 1984.[4] This Comment analyzes these two well-hidden but

---

*A.B., Johns Hopkins University, 1968; M.D., University of Pennsylvania, 1973; member, Class of 1984, Harvard Law School.

[1]*See, e.g.*, Meyer, *Health Care Reform and Market Discipline--Federalism Strikes Back*, REGULATION, Nov.-Dec. 1982, at 16, 18-20; Brown, *Washington Report: The Regulatory Assault*, 7 J. HEALTH POL. POL'Y & L. 772 (1982). For a sense of the high expectations among certain students of health policy stimulated by the advent of the Reagan administration, see R. RICARDO-CAMPBELL, THE ECONOMICS AND POLITICS OF HEALTH (1982), and C. HAVIGHURST, DEREGULATING THE HEALTH CARE INDUSTRY (1982).

[2]For a concise review of "five generic types of plans" that emerged in the 1970's, including those of the Carter and Nixon administrations, the Long-Ribicoff bill, and several proposals cosponsored by Senator Kennedy (D.-Mass.), see Altman, *The Design of a National Health-Insurance System for the United States*, in FEDERAL HEALTH PROGRAMS 205, 211-14 (S. Altman & H. Sapolsky eds. 1981) [hereinafter cited as FEDERAL HEALTH PROGRAMS]. The Gephardt-Stockman bill was introduced in 1981 as "The National Health Care Reform Act of 1981," H.R. 850, 97th Cong., 1st. Sess., 126 CONG. REC. H116 (daily ed. Jan. 16, 1981).

[3]Pub. L. No. 97-248, 1982 U.S. CODE CONG. & AD. NEWS 3 (96 Stat. 324) (codified as amended in scattered sections of 26 U.S.C., 42 U.S.C., and elsewhere). TEFRA set into motion a legislative process culminating in Title VI of the Social Security Amendments of 1983, reforming the Medicare reimbursement mechanism along the lines described in this Comment. *See infra* note 131.

[4]The budget was delivered to Congress on Jan. 31, 1983. In his State of the Union Address to a joint session of Congress six days before, President Reagan made reference to his health-care objectives, but devoted just two sentences to the subject. 41 CONG. Q. 228 (1983).

highly significant health-policy initiatives and their implications for medical care in the United States in this decade and beyond.

Section I briefly reviews the development and changing definition of the health-care crisis and of the government's response to that crisis, so that current proposals may be viewed in an appropriate context. Section II identifies and discusses the ways in which the provisions of TEFRA operate to impose limits on hospital reimbursement under Medicare. Section III summarizes and analyzes in detail the report on prospective reimbursement under Medicare for hospitals, which was submitted to Congress by Secretary of Health and Human Services Richard S. Schweiker in December 1982. Section IV catalogues the other elements of the Reagan administration's "Health Care Incentives Reform" package as revealed by the Fiscal Year 1984 Budget for the Department of Health and Human Services and the accompanying *HHS Fact Sheet*. Section V concludes that the Reagan initiatives may begin to reduce the growth of federal health-care expenditures, but suggests changes to reinforce their strengths and to mitigate their weaknesses.

I. WHAT'S PAST IS PROLOGUE: FEDERAL HEALTH POLICY IN TRANSITION

An understanding of the Reagan administration's emerging health-policy agenda depends upon an appreciation of how deeply the federal government has become involved in the provision of health-care services since the enactment of the Medicare[5] and Medicaid[6] programs in 1965, and of the inflationary consequences of that involvement. It depends, too, upon an analysis of why the health-care sector of the economy has been particularly prone to rapidly escalating costs.

A. *Health-Care Economics in a Nutshell*

That health-care costs need to be contained, and urgently, seems undeniable. Since World War II the rate of inflation in health-care prices has consistently exceeded the overall rate of

---

[5]Social Security Amendments of 1965, Pub. L. No. 89-97, § 102(a), 79 Stat. 291 (current version at 42 U.S.C. §§ 1395–1395tt (1976 & Supp. V 1981)).

[6]*Id.* § 121(a), 79 Stat. 343 (1965) (current version at 42 U.S.C. §§ 1396–1396m (1976 & Supp. V 1981)).

inflation in the United States. For example, in 1982, while the national rate of inflation was 5%, hospital costs rose 15.5%.[7] There has been a steady increase in the proportion of resources that is devoted to health care: in 1950, 4.6% of the Gross National Product (GNP) was expended on health care; in 1965, 5.9%; in 1980, 9.7%, or "roughly one of every ten dollars of wealth being generated in our society."[8] In 1982, health-care costs consumed 10.5% of the GNP,[9] threatening to make the "one out of every ten dollars" illustration, dramatic as it is, obsolete.

As health-care costs have increased, so has our understanding of health-care economics.[10] Although theoretical and empirical approaches have varied, students of the extraordinary inflation[11] in the health-care sector have reached a high degree of consensus as to the origins of the crisis. A primary factor, economists agree, is the fact that the consumer of health-care services is insulated by health insurance and by federal subsidies from the cost of those services and is, therefore, indifferent to price.[12] From the standpoint of consumer cost-consciousness, two subsidies are most important. The first is the effect of federal tax provisions that exempt employer-financed health insurance benefits from the taxable income of employees and that allow deductions for the cost of health insurance premiums.[13] The second and most direct subsidy consists of the Medicare and Medicaid programs.[14] As a consequence of these tax subsidies, an estimated ninety-five percent of Americans currently have some

---

[7]U.S. DEP'T. OF HEALTH & HUMAN SERVS., REPORT TO CONGRESS: HOSPITAL PROSPECTIVE PAYMENT FOR MEDICARE 2 (1982) [hereinafter cited as PROSPECTIVE PAYMENT].

[8]Schramm, *A State-Based Approach to Hospital Cost Containment*, 18 HARV. J. ON LEGIS. 631 (1981).

[9]U.S. Dep't. of Health & Human Servs., HHS Fact Sheet 2 (Jan. 31, 1983) [hereinafter cited as HHS Fact Sheet].

[10]*See generally* P. FELDSTEIN, HEALTH CARE ECONOMICS (1979); R. RICARDO-CAMPBELL, *supra* note 1.

[11]"As many analysts have noted, the term inflation may be misleading in this [hospital cost] context. Inflation normally refers to increases in the price or cost of an unchanged product, but much of the increase in hospital costs has been related to changes in the nature of the product." P. FELDSTEIN, *supra* note 10, at 197.

[12]*See id.*; R. RICARDO-CAMPBELL, *supra* note 1. This federal largesse resulted, ironically, as a conscious federal response to shortages of hospitals and physicians in rural areas, to socioeconomic inequalities of access to medical services, and to a social welfare ethos born of the Depression's ravages. *See* P. STARR, THE SOCIAL TRANSFORMATION OF AMERICAN MEDICINE 235–378 (1983).

[13]I.R.C. §§ 106, 213 (1976).

[14]*See infra* text accompanying notes 30–33.

form of health insurance coverage.[15] More than two-thirds of all health-care expenditures today are paid by third parties.[16]

A second factor upon which analysts of the health-care marketplace have focused is the distortion of the normal supply and demand relationship. One distortion occurs because consumers are not and cannot be knowledgeable about their actual medical needs, and must rely upon their physicians as "purchasing agents" for health care.[17] The physician, aware that third-party payment is guaranteed and that providing additional care to the patient normally results in additional income to the provider, has little or no incentive to withhold medically justifiable care.[18] An incentive to overutilize may exist because professional liability is minimized by erring on the side of ordering any diagnostic procedure or therapeutic intervention that might conceivably, however improbably, benefit the patient.[19] Contrary to what classical economic theory predicts, it appears well-settled that in the health-care field an increased supply of provider-resources (whether hospital beds or new technologies) creates increased demand.[20]

A third factor to which economists point is that conventional competitive forces do not work freely in the health-care market.[21] Physicians have traditionally resisted alternatives to fee-

---

[15]Schramm, *supra* note 8, at 610 n.23 (citing Congressional Budget Office, Profile on Health-Care Coverage: The Haves and Have-Nots (Mar. 1979) (background paper, Washington, D.C.)).

[16]U.S. PUB. HEALTH SERV., HEALTH: UNITED STATES 270 (1981). *See generally* Phelps, *The Effects of Insurance on Demand for Medical Care*, in EQUITY IN HEALTH SERVICES: EMPIRICAL ANALYSIS IN SOCIAL POLICY 142 (R. Andersen, J. Kravits, & O. Anderson eds. 1975); Rosenthal, *Controlling the Cost of Health Care*, in ECONOMICS AND HEALTH CARE 203, 215–18 (J. McKinlay ed. 1981).

[17]*See* R. RICARDO-CAMPBELL, *supra* note 1, at 36–103.

[18]*See* P. FELDSTEIN, *supra* note 10, at 84–90.

[19]*See* J. GUINTHER, THE MALPRACTITIONERS 22–24 (1978). *See generally* THE ECONOMICS OF MEDICAL MALPRACTICE (S. Rottenberg ed. 1978). A related problem is generated by the diagnostic and therapeutic procedures themselves, which in recent decades have become increasingly complex, sophisticated, technology intensive, and expensive, creating more, rather than less, work for physicians and highly trained technicians. *See generally* L. RUSSELL, TECHNOLOGY IN HOSPITALS: MEDICAL ADVANCES AND THEIR DIFFUSION (1979); Wagner & Zubkoff, *Medical Technology and Hospital Costs*, in TECHNOLOGY AND THE FUTURE OF HEALTH CARE 104–30 (J. McKinlay ed. 1982).

[20]*See* Shain & Roemer, *Hospital Costs Relate to the Supply of Beds*, 9 MODERN HOSPITAL 71 (1959); *see also* P. FELDSTEIN, *supra* note 10, at 84–90. This appears to be related to the purchasing-agent function of the physician, which allows the provider of services to define the medical necessity (and therefore the demand) for the health-care services provided.

[21]*See, e.g.*, R. RICARDO-CAMPBELL, *supra* note 1, at 36–103; *see also* Roberts & Bogue, *The American Health Care System: Where Have All the Dollars Gone?*, 13 HARV. J. ON LEGIS. 635, 663–66 (1976).

for-service practice and shunned colleagues who have worked as employees of corporations or of prepaid group practices.[22] Organized medicine has steadfastly opposed advertising to the public by individual physicians or by groups, and has imposed sanctions on errant doctors.[23] Entry to the profession has, of course, been restricted by means of licensing and certification. In addition, hospitals must be accredited and certified by the state.[24] The nonprofit structure of most hospitals historically has caused cost-effective management to develop slowly, if at all.[25]

B. *Federal Health Programs and Policy, 1946–1976*

To appreciate the new administration's proposals, it is necessary to outline the legacy of federal health programs inherited from preceding administrations, and the relationship of those programs to the economic theory summarized above.

Not until the nation began to refocus on domestic needs in the aftermath of World War II did a United States President call forcefully for a national health insurance program.[26] Promptly denounced by the American Medical Association as a plan for making doctors "slaves" and by Republican Senator Taft of Ohio as coming "right out of the Soviet Constitution," Truman's health insurance plan was frozen and killed by the icy winds of

---

[22]*See* L. BROWN, POLITICS & HEALTH CARE ORGANIZATION: HMOS AS FEDERAL POLICY 104 (1983).

[23]Canby & Gellhorn, *Physicians Advertising: The First Amendment and the Sherman Act*, 1978 DUKE L.J. 543, 546–47.

[24]*See* P. STARR *supra* note 12, at 102–12; *see also* Cohen, *Professional Licensure, Organizational Behavior, and the Public Interest*, in HEALTH CARE CONSUMERS, PROFESSIONALS, AND ORGANIZATIONS 232–47 (J. McKinlay ed. 1981).

[25]*See* Clark, *Does the Nonprofit Form Fit the Hospital Industry?*, 93 HARV. L. REV. 1416–89 (1980).

[26]The United States Congress had established a system of compulsory hospital insurance for merchant seamen in 1798 and had created a National Board of Health in 1879. However, the former was a consequence of commercial expediency and the latter a four-year flash in the pan extinguished, ironically, in 1883, the same year Germany pioneered the world's first national system of compulsory sickness insurance. *See* P. STARR, *supra* note 12, at 237, 240. The German experiment was imitated elsewhere in Europe, but not in the United States, where the defeat of Theodore Roosevelt's Progressive candidacy in 1912 delayed for at least two decades consideration of social insurance, including health insurance, which the Progressives had advocated. *Id.* at 243–49. Even then, the pressures generated by the Depression and New Deal years were focused more intensely upon social welfare insurance than upon national health insurance per se, so that the most significant developments of those years—the birth of the Blue Cross associations and the passage of the Social Security Act—were largely indirect. *See* S. LAW, BLUE CROSS: WHAT WENT WRONG 6–9 (1974); *see also* P. FELDSTEIN, HEALTH ASSOCIATIONS AND THE DEMAND FOR LEGISLATION: THE POLITICAL ECONOMY OF HEALTH 173–77 (1977); P. STARR, *supra* note 12, at 266–80.

the Cold War. One of the plan's recommendations was, however, signed into law with the approval of the medical and hospital establishments: the Hospital Survey and Construction Act of 1946 (Hill-Burton Act).[27]

The Hill-Burton program, which provided extensive funding for hospital construction, required state governments to administer eligibility rules under which applicant hospitals certified that they would not discriminate in providing care and would provide "a reasonable volume of services to persons unable to pay."[28] In return, between 1947 and 1973, the federal government granted nearly four billion dollars to almost six thousand hospitals, and generated some nine billion dollars in state and local matching funds.[29]

Even more massive health-care programs were contained in the Social Security Amendments of 1965: Medicare and Medicaid.[30] Part A of Medicare was a compulsory hospital insurance plan for the elderly. Part B of Medicare provided government-subsidized voluntary insurance to cover physicians' services for the elderly. Medicaid, financed jointly by the federal government and the states, covered medical and hospital care for welfare recipients.[31] These programs undoubtedly succeeded in increasing the access of the old, the disabled, and the poor to high-quality medical care, but they did so at enormous cost. The economic projections presented to Congress in 1965 suggested, for example, that by 1970 federal outlays for Medicare would total $2.46 billion; in fact, 1970 expenditures were $4.95 billion.

---

[27]Title VI of the Public Health Service Act, Pub. L. No. 79-725, 60 Stat. 1041 (1946) (codified as amended at 42 U.S.C. §§ 291–291o (1976)); *see* P. STARR, *supra* note 12, at 280–83; *see also* Comment, *The Hill-Burton Act, 1946 - 1980: Asynchrony in the Delivery of Health Care to the Poor*, 39 MD. L. REV. 316 (1979).

[28]42 U.S.C. §§ 291–291o. The nondiscrimination provision is at § 291c(e)(1); the free-care provision is at § 291c(e)(2). *See generally* Rosenblatt, *Health Care Reform and Administrative Law: A Structural Approach*, 88 YALE L.J. 243 (1978); Comment, *supra* note 27.

[29]Schramm, *supra* note 8, at 617. The program provided

> for the addition of 358,000 beds, or roughly one third of all hospital capacity in existence in 1973. Even though the free-care obligation was at best loosely complied with by many hospitals, it established the federal government as a major albeit indirect source of health care for the poor . . . . To the extent to which hospital care is supply-induced, the inflationary pressures on hospital costs were activated. Moreover, the Hill-Burton program established the federal government as a permanent presence in the hospital industry.

Id.; *see also* P. STARR, *supra* note 12, at 349–50.

[30]Social Security Amendments of 1965, Pub. L. No. 89-97, § 102(a), 79 Stat. 291 (current version at 42 U.S.C. §§ 1395–1395tt (1976 & Supp. V 1981)).

[31]*Id.* § 121(a), 79 Stat. 343 (current version at 42 U.S.C. §§ 1396–1396m (1976 & Supp. V 1981)).

The original estimate for 1975 was $2.7 billion; the reality was $10.6 billion.[32] Health-care funding is the fastest-rising expenditure in the federal budget, with combined federal Medicare and Medicaid spending up nearly 600% since 1970 to $64 billion.[33]

Congressional response over the last decade to skyrocketing federal health expenditures has been resourceful but only moderately effective. The Social Security Amendments of 1972[34] established Professional Standards Review Organizations (PSROs), panels of physicians empowered to monitor their peers through "utilization review" procedures. These PSROs were designed to eliminate unnecessary hospital admissions and to ensure that patients were hospitalized only as long as medically required.[35] The Health Maintenance Organizations Act of 1973,[36] passed with the strong backing of the Nixon administration, provided federal funding to prepaid group practices that furnish care for a fixed fee to enrolled individuals. The theory was that such Health Maintenance Organizations (HMOs) would prove more cost effective than traditional fee-for-service practice and that they would result in less hospitalization and therefore significantly lower total cost overall.[37]

When it became clear that various federal enactments designed to increase the supply of physicians and paraprofessionals[38] were, paradoxically, increasing rather than decreasing the demand for and cost of medical services, the federal government shifted into reverse with its most ambitious regulatory program to date, the National Health Planning and

---

[32]Schramm, *supra* note 8, at 621 n.87. "This year, Medicare and Medicaid will spend as much *every month* as they did during the *entire year* of 1966, their first full year of operation." HHS Fact Sheet, *supra* note 9, at 2.

[33]HHS Fact Sheet, *supra* note 9, at 3.

[34]Pub. L. No. 92-603, 86 Stat. 1329 (1972) (codified as amended in scattered sections of 42 U.S.C.).

[35]*Id.* § 249f(b), 86 Stat. 1429 (codified at 42 U.S.C. § 1320c (1976 & Supp. V 1981)); *see* Berman & Gertman, *Cost Containment and Quality Assurance: The Potential and Performance of the Professional Standards Review Organization Program*, in FEDERAL HEALTH PROGRAMS, *supra* note 2, at 43–61; *see also* H. GOSFIELD, PSROs: THE LAW AND THE HEALTH CONSUMER (1975); Havighurst & Blumstein, *Coping with Quality/Cost Trade Offs in Medical Care: The Role of PSROs*, 70 NW. U.L. REV. 6 (1975).

[36]Pub. L. No. 93-222, § 2, 87 Stat. 914 (codified at 42 U.S.C. § 300e (1976 & Supp. V 1981)).

[37]*See* McNeil & Schlenker, *HMOs, Competition, and Government*, in HEALTH MAINTENANCE ORGANIZATIONS 31–60 (J. McKinlay ed. 1981); *see also* L. BROWN, *supra* note 22, at 3.

[38]*See* S. Wallack, *Federal Health Professional Training Programs: The History and Impact*, in FEDERAL HEALTH PROGRAMS, *supra* note 2, at 3–39.

Resource Development Act of 1974.[39] The Act created new regulatory and planning agencies at federal, state, and local levels of government with the aim of reviewing major capital expenditures by hospitals in an effort to constrain, in a rationally planned manner sensitive to local needs, major expansions or replacements of health-care facilities.[40] Unlike the PSRO or the HMO programs, the Health Planning legislation was avowedly regulatory. Indeed, it systematically incorporated consumer representation into local Health Systems Agencies (HSAs), some two hundred of which were established to pass upon Certificate of Need applications from institutions that were required to obtain these certificates to make significant capital expenditures.[41]

The effectiveness of this new alphabet soup of regulatory schemes—PSRO, HMO, HSA, CON—has been difficult to measure. In each case, program implementation has varied widely, and conflict and confusion have been substantial. Moreover, results have been contradictory and subject to the vagaries of study design choices, inadequacies of data, and problems of interpretation.[42] In the latter half of the 1970's, studies appeared that suggested one program or another had effected some savings or efficiencies. Frequently, however, these were contradicted by studies suggesting no benefits at all, or benefits that were significantly outweighed by the costs of the program itself, in terms of administrative resources, quality of care, and money.[43] One aspect of the programs, however, was indisputable: whatever they had achieved, they had not prevented federal

---

[39]Pub. L. No. 93-641, 88 Stat. 2225 (1975) (codified at 42 U.S.C. § 300k (1976)); *see* P. STARR, *supra* note 12, at 402; *see also* Marmor & Morone, *Representing Consumer Interests: Imbalanced Markets, Health Planning, and the HSAs*, in HEALTH SERVICES RESEARCH, PLANNING, AND CHANGE 223, 228 (J. McKinlay ed. 1981). This program replaced the Hill-Burton Act.

[40]*See generally* L. BROWN, *supra* note 22.

[41]Altman, *Creating a National Health-Planning System*, in FEDERAL HEALTH PROGRAMS, *supra* note 2, at 121–31.

[42]*Id.*; *see also* L. BROWN, *supra* note 22; Berman & Gertman, *supra* note 35.

[43]Altman, *supra* note 41. The author of an exhaustive recent study of the HMO legislation concludes, for example: "As a policy strategy designed to correct the faulty incentives of the fee-for-service system and thereby reorganize American medicine, HMOs have been of little importance." L. BROWN, *supra* note 22, at 488. With respect to PSROs: "As a cost-control mechanism, the PSRO program is only marginally successful." Berman & Gertman, *supra* note 35, at 57. As to the health planning system: "To date, there has been wide variation across the country . . . . Generally, however, neither HSAs nor state agencies are likely to force signficant reductions in the number of existing local beds." Altman, *supra* note 41, at 129–30.

health expenditures from continuing to rise at a rate far exceeding that of other goods and services in the economy.[44]

II. Reagan's Hidden Health Agenda: The Tax Equity and Fiscal Responsibility Act of 1982

Notwithstanding the identification of important administration officials with health issues,[45] or the obvious fit between President Reagan's conservative ideology and the failure of the liberal health agenda, the new administration appeared to assign health issues a low priority. When Secretary of Health and Human Services Richard Schweiker announced his intention to discontinue the PSRO and HSA initiatives, much to the relief of organized medicine, and to cut back entitlement programs generally, the initial thrust was negative, not affirmative.[46] A "benign neglect" of health issues seemed to prevail, despite the urgings of conservatives to deregulate health care through vigorous antitrust enforcement.[47] The administration endorsed no health legislation and presented none. It seemed, in fact, to be backing away from the complex and controversial subject altogether.

---

[44]*See supra* text accompanying notes 7–11, 32–33. In his recent account of "the social transformation of American medicine," sociologist Paul Starr observes:

> When the blizzard of regulation stopped, the federal government found itself snowed in. Between 1971 and 1974 Congress had passed a great deal of complicated legislation. The laws were especially detailed because of Democratic reluctance to trust the Nixon administration with much discretion. Some were so severely compromised in passage as to be nearly unworkable. And each provoked bureaucratic conflict and litigation that took years to resolve. Meanwhile, little was accomplished and the impression was conveyed that the reforms were a failure.

P. Starr, *supra* note 12, at 405.

[45]Before his appointment as Director of the Office of Management and Budget, David Stockman was a member of the House of Representatives whose most prominent policy role was his co-sponsorship of health-care reform legislation. The Gephardt-Stockman bill was introduced in 1981 as "The National Health Care Reform Act of 1981," H.R. 850, 97th Cong., 1st. Sess., 126 Cong. Rec. H116 (daily ed. Jan. 16, 1981). Martin Feldstein, a Harvard economics professor currently serving as Chairman of the President's Council of Economic Advisors, has published numerous articles and a recent book on the subject of health economics. *See, e.g.*, M. Feldstein, Hospital Costs and Health Insurance (1981).

[46]P. Starr, *supra* note 12, at 419.

[47]*Id.* Clark Havighurst and other advocates of "procompetitive" approaches to health planning, such as Alain Enthoven, clearly had the President's ideological ear. *See* C. Havighurst, *supra* note 1; *see also* A. Enthoven, Health Plan: The Only Practical Solution to Soaring Health Costs (1980); Enthoven, *How Interested Groups Have Responded to a Proposal for Economic Competition in Health Services*, 70 Am. Econ. Rev. 142–48 (1980); Havighurst, *Competition in Health Services: Overview, Issues and Answers*, 34 Vand. L. Rev. 1115–78 (1981).

In 1983, the outline of the administration's health agenda is beginning to emerge. However, either because the administration sees federal health policy as merely one manifestation of an overall domestic fiscal austerity program, or because of a deliberate strategy to accomplish meaningful change without arousing controversy and opposition, the Reagan health initiatives have been disguised. To understand the new direction of health policy for the 1980's may, therefore, require analysts to read tax laws as if they were tea leaves, beginning with a close look at TEFRA.[48]

The health-planning significance of TEFRA is contained in three provisions that revise the way hospitals are to be reimbursed for their Medicare patients. In an effort to remedy the situation, TEFRA (1) shifted the basis of Medicare reimbursement from per-diem to per-case costs and included ancillary and special care costs in the calculation;[49] (2) established a "target rate reimbursement system" that limits, for three years, the allowable rate of increase in Medicare payments for hospital operating expenses;[50] and, (3) directed the Secretary of Health and Human Services to develop and submit to the Congress by December 31, 1982, legislative proposals for reimbursing hospitals and other Medicare Part A providers on a prospective rather than a retrospective basis.[51] In order to understand the rationale behind and the significance of these rather technical changes, a brief review of the history and evolution of the Medicare reimbursement system is needed.

At the risk of oversimplifying a complex and fascinating story, it may safely be said that in establishing the Medicare program in 1965 Congress and the Johnson administration were almost obsessively concerned with accommodating the interests of the hospital and insurance industries and of organized medicine. This concern resulted in at least four rather simple but profoundly significant decisions that were incorporated into the Social Security Amendments of 1965.[52] The most important de-

---

[48]Pub. L. No. 97-248, 1982 U.S. CODE CONG. & AD. NEWS 3 (96 Stat. 324) (codified as amended in scattered sections of 26 U.S.C., 42 U.S.C., and elsewhere).

[49]TEFRA, § 101(a)(1).

[50]*Id.*

[51]*Id.* § 101(b)(3).

[52]Social Security Amendments of 1965, Pub. L. No. 89-97, 79 Stat. 286 (codified as amended in scattered sections of 26 U.S.C. and 42 U.S.C.). *See generally* H. SOMERS & A. SOMERS, MEDICARE AND THE HOSPITALS (1967); P. STARR, *supra* note 12, at 374–78. First, the program was designed as an extraordinarily broad entitlement, covering all persons over age 65 regardless of personal resources, physical health, or other

cision required that "the amount paid to any provider of services with respect to services for which payment may be made under this [Title] shall . . . be the reasonable cost of such services."[53]

This decision to reimburse Medicare providers on the basis of retrospectively determined "reasonable cost" rather than on the basis of charges determined through negotiations with hospitals, as Blue Cross plans generally had in the past, was a momentous one. The initial regulations, promulgated to define reasonable cost and to provide the rules to calculate it, made the implications of that decision immediately clear[54] as the implementing regulations specified that virtually all costs incurred by a hospital would be reimbursed, with no constraints on cost increases and with no requirements for efficiency, regardless of the magnitude of the cost.[55] The only possible exception was for costs found to be "substantially out of line with other institutions in the same area which are similar in size, scope of services, utilization, and other relevant factors."[56] The original regulations thus created a "cost-plus" formula, the most powerful "plus" being generous accelerated depreciation on all hospital assets, including Hill-Burton assets that had been provided by the government.[57] To that "plus" was added icing on an already amply iced cake: an "allowance in lieu of specific recognition of other costs," and an add-on of two percent of all

---

insurance coverage. Second, the law provided for "fiscal intermediaries" as buffers between the government and the health-care industry, which ensured that the administration of the program would be lodged in a provider-dominated private insurance industry. Third, in an effort to eliminate the "two-tier" (ward versus private) system of hospital care and to ensure beneficiaries freedom of choice in choosing providers, the program sought to provide benefits essentially equal to those enjoyed by paying and privately insured patients. Schramm, *supra* note 8, at 619.

[53]42 U.S.C. § 1395f(b)(1) (1976 & Supp. V 1981).

[54]Weiner, *Paying for Hospital Services under Medicare: Can We Control Hospital Costs?*, in FEDERAL HEALTH PROGRAMS, *supra* note 2, at 135–39.

[55]*See* 42 C.F.R. § 405.451(c)(3) (1981) ("The reasonable cost basis of reimbursement contemplates that providers of services would be reimbursed the actual costs of providing quality care however widely the actual costs may vary from provider to provider and from time to time for the same provider.").

[56]*Id.* § 405.451(c)(2).

[57]*Id.* § 405.451(a)(3). Moreover,

> hospitals were given the right to use one method of depreciation with respect to one asset or group of assets and a different method for others. Depreciation was allowed on assets originally financed with federal or other public funds, even though the hospital incurred no cost in acquiring the asset. Finally, depreciation was allowed on assets in use at the time the hospital entered the program, even though the assets may have been fully depreciated on the hospital's books . . . .

Weiner, *supra* note 54, at 137–38; *see also* P. STARR, *supra* note 12, at 375 ("[D]epreciation for a nonprofit institution is a peculiar idea; when a community donates capital to an institution, it does not necessarily agree to replace it.").

other allowable costs.[58] In essence, the federal government "effectively abdicated all responsibility for evaluating—or assisting the hospitals in determining for themselves—the necessity or efficiency of costs incurred by hospitals."[59]

As Medicare expenditures outstripped initial projections, Congress passed three significant amendments. The amendments authorized alternative reimbursement experiments in 1967,[60] replaced the two percent add-on with a more modest nursing cost allowance in 1969,[61] and eliminated the accelerated depreciation allowance in 1970.[62] By far the most important legislative response to the open-ended, cost-based, and cost-generating reimbursement provisions of the original Medicare program, however, was section 223 of the Social Security Amendments of 1972.[63]

The provisions of section 223 represented the government's first real attempt to define what reasonable costs were. For the first time, hospitals were grouped according to size and geographic region in order to set ceilings on the amount Medicare would reimburse hospitals for their average routine costs per day "that would be incurred," in the words of the House Report, "by a reasonably prudent and cost-conscious management."[64] By giving content to the concept of reasonable costs, section 223 became "the keystone of the federal attempt to contain Medicare expenditures . . . [and] the most important federal restraint on cost increases."[65] As the record of hospital and Medicare cost increases since 1972 demonstrates, however, that restraint was insufficient.

Section 101(a)(1) of TEFRA may be best understood as an attempt to give section 223 important new teeth. By shifting the basis of Medicare reimbursement from per-diem to per-case costs, the possibility of meaningfully comparing similar hospitals will be enhanced; their costs incurred in treating patients with

---

[58]*See* Weiner, *supra* note 54, at 137–38.

[59]*Id.* ("[The] reasonable-cost formula represented a guarantee to hospitals that their actual patient-related costs would be reimbursed in full, and that they would receive payments even in excess of actual cost . . . ."). For the rationale behind what Professor Starr terms "the politics of accommodation," see P. STARR, *supra* note 12, at 290–334.

[60]*See* Weiner, *supra* note 54, at 139.

[61]*Id.*

[62]*Id.*

[63]Pub. L. No. 92-603, 86 Stat. 1329 (1972) (codified as amended in scattered sections of 42 U.S.C.); *see* Weiner, *supra* note 76, at 139.

[64]H.R. REP. No. 231, 92d Cong., 2d Sess. 101, *reprinted in* 1972 U.S. CODE CONG. & AD. NEWS 4989, 5088.

[65]Schramm, *supra* note 8, at 630.

a particular diagnosis should be the same. The Medicare program will be able to deny reimbursement in excess of the average cost. By including in the formula not only "routine" operating cost but "ancillary" costs as well, that is, the cost of x-ray procedures, laboratory and diagnostic procedures, drugs and special (e.g., intensive care) costs, TEFRA also closes a significant loophole in section 223. Such "ancillary" costs, despite their innocuous label, have been increasing with extraordinary speed in this era of high technology, defensive medicine, and sophisticated hospital administration and accounting. And by instituting a "target rate reimbursement system" to limit the allowable rate of annual reimbursement increase for three years beginning October 1, 1982, TEFRA supplies the Medicare program with crushing molars to ensure against the possibility that unpalatable cost increases might get by the incisors and canines of the new cost formula.[66]

III. The Secretary's Report: The Prospects for Prospective Reimbursement

As early as 1971, the House Ways and Means Committee expressed its belief that a prospective reimbursement system[67]

> offers the promise of encouraging institutional policymakers and managers, through positive financial incentives, as well as the risk of possible loss inherent in that method, to plan, innovate, and generally to manage effectively in order to achieve greater financial reward for the provider, as well as a lower total cost to the programs involved . . . . Theoretically, this approach to reimbursement introduces incentives not present under the existing reimbursement method which, since it tends to pay whatever the costs turn out to be, provides no incentives for efficiency.[68]

As a consequence, Congress included in its 1972 amendments to the Social Security Act an explicit authorization for experi-

[66]Under the new law, reimbursement increases would be limited to "the percentage increase in the hospital wage and price index plus one percentage point." S. Rep. No. 530, 97th Cong., 1st Sess. 420, *reprinted in* 1982 U.S. Code Cong. & Ad. News 12. Under the new section, 25% of costs in excess of the target would be allowed the first two years, but not in the third. In addition, there is an incentive payment to hospitals to keep their costs below the specified target rate (set at half the difference between the limit and their actual cost), although the incentive payment is capped at five percent of the target rate. *Id.*

[67]*See infra* text accompanying notes 73–76.

[68]H.R. Rep. No. 231, 92d Cong., 2d Sess. 80, *reprinted in* 1972 U.S. Code Cong. & Ad. News 4989, 5066.

ments and demonstration projects designed to test the potential of prospective reimbursement mechanisms to achieve cost savings.[69] Thus, in December 1982 when the Secretary of Health and Human Services reported to Congress on the subject, pursuant to section 101(b)(3) of TEFRA, he did so on the basis of a decade of experience and analysis within his department.[70]

The report, a comprehensive document of nearly three hundred pages, is unabashedly optimistic: "The Department believes that the prospective payment system proposed here will provide hospitals an incentive to improve efficiency, will establish Medicare as a prudent buyer of hospital services, will reduce the administrative burden on hospitals, and will assure beneficiary access to quality health care."[71] That optimism is based both upon the government's experience with demonstration projects in Connecticut, Maryland, Massachusetts, New Jersey, New York, Rhode Island, Washington, and Wisconsin,[72] and upon abstract theoretical grounds. Whether or not the report's optimism is justified, the force with which the Department of Health and Human Services is making the argument for a nationwide shift to prospective reimbursement, and its evident eagerness for doing so as soon as possible, are important indications that a Reagan administration health policy is emerging.

Chapter I of the report describes the Prospective Payment System (PPS).[73] The report declares that the ultimate objective of PPS is to set a reasonable price for known health-care products in order to provide incentives for hospitals to produce those products more efficiently. Thus, health-care providers will be confronted with "strong, lasting incentives to restrain costs for the first time in Medicare history."[74] The chapter outlines the history and distinctive economics of health care, and approves of the changes instituted by TEFRA, but notes that "even under this system, Medicare reimbursement remains a retrospective cost-based limit system."[75] Prospective payment rates, by contrast, "are determined in advance and fixed for the fiscal period to which they apply" and constitute "payment in full for the

---

[69]Social Security Amendments of 1972, Pub. L. No. 92-603, § 222(a), 86 Stat. 1390 (codified at 42 U.S.C. § 1395f (1976 & Supp. V 1981)).

[70]PROSPECTIVE PAYMENT, *supra* note 7.

[71]*Id.* at i.

[72]*Id.* at 19–33.

[73]*Id.* at 1.

[74]*Id.* at 3.

[75]*Id.* at 13.

specified unit of service," as a consequence of which "the hospital keeps the difference between the payment rate and its cost of providing the service and is at risk for exceeding the payment rates."[76]

Chapter II of the report summarizes the lessons learned from the eight states that have experimented with various versions of prospective payment.[77] The first of these lessons is that mandatory systems are more effective than voluntary ones.[78] A second is that such systems reduced cost increases by two to six percent.[79] A third is that differences among hospitals, such as between rural hospitals and high-technology university teaching hospitals, must be considered.[80] Overall, the results are "encouraging," the report concludes, but it includes some findings that must be considered caveats.[81] For example, "not all systems . . . tested are applicable to a Medicare-only system, and other systems do not meet the criteria set forth by the Department for a prospective payment system."[82] Still, the report insists its experiments "support the contention that there are feasible alternatives"[83] that "have in fact been demonstrated in real life settings."[84]

Chapters III and IV define and describe the Medicare prospective payment system proposal under which the Health Care Financing Administration (HCFA) "will establish payment for inpatient care to hospitals which participate in the Medicare program at a predetermined rate for each type of Medicare discharge in accordance with a Federal payment schedule for standard types of patient cases."[85] Those predetermined prospective rates will be based on a cost per discharge for each "Diagnosis Related Group" (DRG).[86]

---

[76]*Id.* at 15.
[77]*Id.* at 19.
[78]*Id.*
[79]*Id.* at 32.
[80]*Id.*
[81]*Id.* at 33.
[82]*Id.*
[83]*Id.* at 19.
[84]*Id.* at 34–40.
[85]*Id.* at 40.
[86]*Id.* at 41–42. These per-case rates will be adjusted for local variations in hospital labor costs. Capital costs and medical education costs also will be excluded from the initial rate calculations and reimbursed separately on a reasonable cost basis. Outpatient costs will be reimbursed independently, and as time passes the inpatient rates "will be updated by the Secretary, who will take into account such factors as inflation of hospital input costs (the hospital marketbasket index), improved industry productivity, and technology." *Id.* at 42.

Developed at Yale University, the DRG system groups patients' treatments into 467 categories.[87] Factors taken into account by the DRG system include the patient's age and discharge status, the primary diagnosis of the patient, the secondary diagnosis of the patient, the primary surgical procedure utilized (if any), secondary procedures, complications, and comorbidities.[88] As a consequence, skilled labor and technology-intensive cases such as kidney transplants (DRG 302) will receive a much higher payment than simpler cases such as hernia repair (DRG 161), and a myocardial infarction (heart attack) with complications (DRG 121) will be reimbursed at a higher rate than an uncomplicated myocardial infarction (DRG 122).[89]

In designing the DRG system, HCFA and the Yale group sought to fashion a method that was medically interpretable, compatible with existing Medicare data and medical record abstracts, and simple enough to be administrable. Yet the DRG system had to be comprehensive enough to be representative of the entire range of hospital patients, and fair.[90] The detailed report of the development of the DRG concept, with extensive illustrations of how a typical set of DRGs was developed, clinically reviewed, and refined, along with practical experience and evidence of administrability from the New Jersey prospective reimbursement demonstration project, are convincing.[91] The DRG-based PPS, as presented in the pages of the report, makes special provision for atypical or "outlier" cases "which, although classifiable into a specific DRG, have an extremely short or extremely long length of stay relative to most cases in the same DRG."[92] The care with which the report addresses significant related issues also reinforces the impression that the prospective reimbursement system proposed in this report can, and should, be implemented.[93]

---

[87]*Id.* at 43. The categories are "derived from a multi-stage process applied in conjunction with a nationally representative sample of 1.4 million patient discharge records." The objective of the system is to classify patients into groups (types of cases) that are "clinically coherent and homogeneous with respect to resource use," so that comparable services in comparable hospitals can be reimbursed identically. *Id.* at 66.

[88]*Id.* at iv, 72. "Comorbidities" are concurrent illnesses.

[89]*Id.* at iv.

[90]*Id.* at 71.

[91]*See id.* at 66–101 & app. F.

[92]*Id.* at 51–53.

[93]*Id.* at 62–63 (summarizing specific points made throughout chapters II and III of the report). For example, the allocation of costs is broken down into direct capital costs (initially passed through and reimbursed on a reasonable cost basis), direct medical education costs (passed through), indirect medical education costs (estimated and paid

In chapters V and VI, the report specifies in technical detail its mechanism for setting initial PPS prices and predicting Medicare hospital revenues under the new system, considers the limitations of existing Medicare data and the potential sources of error, and concludes, reasonably, that the Department is ready to introduce the new Medicare reimbursement system as soon as Congress approves. In short, the Reagan administration has not only moved, but moved boldly in the field of federal health policy.

IV. THE BUDGET AS BACKDOOR: "HEALTH CARE INCENTIVES REFORM"

Although continuing to eschew a frontal assault on health-care issues in the traditional comprehensive legislative form, President Reagan nonetheless proposed dramatic changes in federal health policy in his Fiscal Year 1984 Budget.[94] Characterized by the *HHS Fact Sheet* (issued in January to accompany the Department's portion of the budget) as a "Health Care Incentives Reform" plan, the proposed program incorporates prospective reimbursement and, according to HHS,

> —provides Medicare coverage for catastrophic illness involving lengthy hospital stays;
> —improves cost-sharing provisions in Medicare, encouraging efficiency while reducing cost burden on the severely ill;
> —limits tax subsidy of higher-cost private health plans;
> —expands opportunities for Medicare beneficiaries to use their benefits to enroll in private health plans.[95]

In addition, the budget proposal contains provisions to which the *HHS Fact Sheet* does not allude. It would freeze physician reimbursement for one year,[96] reduce TEFRA's target rate of hospital cost increase by eliminating the one percent add-on allowance to the percentage increase in the "hospital market

---

on a "lump sum" basis), outpatient care costs ("reimbursed on a reasonable cost basis until methods are developed to pay for this prospectively"), and the special needs of psychiatric, pediatric, long term care and teaching hospitals, as well as rural hospitals and HMOs. *Id.*

[94] *See supra* note 4. The President devoted just two sentences in his State of the Union Address last January to the need to curb "the skyrocketing cost of health care" and to "provide catastrophic illness insurance coverage for older Americans." 41 CONG. Q. 228 (1983).

[95] HHS Fact Sheet, *supra* note 9, at 1.

[96] U.S. DEP'T. OF HEALTH & HUMAN SERVS., FISCAL 1984 BUDGET 29–30 [hereinafter cited as HHS BUDGET].

basket" currently allowed,[97] increase the premium currently paid by Medicare Part B beneficiaries from a formula currently equalling twenty-five percent of program costs to one equalling thirty-five percent,[98] increase the deductible paid by beneficiaries under Part B from the current seventy-five dollars (up from sixty dollars in 1980) to eighty dollars in 1984,[99] defer Medicare eligibility of new beneficiaries by one month,[100] and eliminate mandatory utilization review (the PSRO program). Other aspects of the budget proposal are aimed at reducing regulatory burden;[101] lowering reimbursement to home health agencies for durable medical equipment;[102] authorizing the Department of Health and Human Services to employ competitive purchasing procedures for the procurement of laboratory services, durable medical equipment and other medical supplies for beneficiaries;[103] eliminating the waiver of provider liability for uncovered Medicare services;[104] and, requiring "nominal" cost sharing by Medicaid recipients.[105]

V. Assessments and Conclusions: Death, Taxes, and the Future of Health Care in the United States

"Never before," according to a January 31, 1983, statement by Secretary Schweiker, "has any administration proposed such a balanced, sensible and comprehensive plan to ease health cost inflation."[106] That appraisal is, to say the least, open to question. There can be little doubt, however, that the Reagan health policy agenda, because of the manner in which some of its most important provisions are tucked away in budgetary minutiae, makes sense as a strategic matter. The recent history of health-care legislation suggests that the more comprehensive a bill is, the more likely it is to attract opposition, and eventually to be watered down or defeated. Moreover, it has become increasingly clear since 1965 that something needs to be done to restrain

---

[97] *Id.* at 30.
[98] *Id.* at 30–31.
[99] The Part B deductible amount would be indexed to the Medicare Economic Index beginning in 1984. *Id.* at 31.
[100] *Id.*
[101] *Id.*
[102] *Id.*
[103] *Id.* at 32.
[104] *Id.*
[105] *Id.* at 34–35.
[106] HHS News, Jan. 31, 1983, at 2 (Statement of Secretary Richard S. Schweiker).

health-care expenditures. Not only do American taxpayers, through Medicare and Medicaid, pay an estimated forty percent of all hospital bills, but tax provisions granting favorable tax treatment to medical expenses and health insurance benefits have decreased government tax revenues and have provided incentives for Americans to over-insure and thereby to become effectively insulated from the actual costs of the medical and health services they purchase.[107]

Perhaps the most progressive of the administration's "Health Care Incentives Reform" budget proposals is its plan to limit the tax subsidy currently provided for expensive employer-provided health insurance coverage. Section 106 of the Internal Revenue Code presently excludes from gross income employer contributions to employee accident or health plans.[108] The administration proposal would restrict tax-free treatment to amounts of seventy dollars per month for individual coverage or $175 per month for family coverage. At present, approximately thirty percent of those with employment-based health coverage receive employer contributions above these limits.[109] "While individuals and companies would remain free to purchase as much health coverage as they desire, the new provision would eliminate the bias that now works in favor of high-priced coverage and against comparable higher wages."[110]

The Medicare reimbursement reforms included in TEFRA and in the Prospective Payment System proposed pursuant to TEFRA's section 101(b)(3) have genuine potential for initiating control of federal Medicare expenditures to hospitals. Any endorsement of the administration's initiatives in this area, however, must be qualified by several observations. First, and most obviously, because TEFRA and the prospective payment system apply *only* to Medicare, hospitals may attempt to shift their costs to private patients and to private insurance companies.

A second weakness of PPS is that university-affiliated and other teaching hospitals will receive extra payments for actual medical education costs. There is thus a danger that large, high-cost, high-technology institutions may become a "privileged class" under PPS. This not only raises the possibility of inequity and abuse, but reinforces the current system's bias towards

---

[107]*See supra* text accompanying notes 13–16.
[108]I.R.C. § 106 (1976).
[109]HHS Fact Sheet, *supra* note 9, at 8.
[110]*Id.*

high-cost, high-technology diagnostic and therapeutic procedures.[111]

A third problem inheres in the DRG concept as a prospective reimbursement standard.[112] The notion of paying the same amount per DRG, after taking local hospital labor conditions into consideration, is extremely attractive. A hospital, however, will be motivated to keep each patient's hospital stay short and to admit as many cases as possible because reimbursement is figured on a per-case rather than per-day basis. The HHS report, to its credit, recognizes this potential threat of "gaming" and "DRG creep,"[113] and appears to rely on "the constant threat of malpractice suits" as a deterrent.[114] One may question whether this deterrent will be strong enough to keep hospitals from discharging patients prematurely, which would result in otherwise unnecessary readmissions.[115]

Fourth, the DRG-based PPS does not affect the physician's "purchasing agent" role in the health-care system. Although the PPS gives hospitals an incentive to operate efficiently, that incentive must be transmitted to and incorporated by each hospital's medical staff. Physicians cherish their independent professional judgment, and some, perhaps substantial, conflict is likely.[116]

Fifth, and most important, a hospital working within rigid cost constraints may choose to sacrifice quality of care in order to minimize costs through the cumulative impact of many small and some large decisions. Premature discharges are one danger; a failure to perform important ancillary services may be another. This suggests that ultimately "quality" will need to be defined. Moreover, such regulatory initiatives as utilization review by peer group physicians as well as capital-expenditure review, now anathemas to the Reagan administration in its asserted

---

[111]*See* PROSPECTIVE PAYMENT, *supra* note 7, at 47–48.

[112]HHS Fact Sheet, *supra* note 9, at ii. Despite the essential uniformity of medical care created by a national system of medical education, examination, and specialty certification under the current cost-based retrospective payment mechanism, Medicare payments to providers vary dramatically. Medicare will pay $1500 for treating a heart attack at one hospital and $9000 for treatment at another, for example. Standard reimbursement for a hip replacement can range from $2100 to $8200, depending upon the provider; and, reimbursement for essentially indistinguishable cataract surgery can vary from $450 to $2800. *Id.*

[113]PROSPECTIVE PAYMENT, *supra* note 7, at 74.

[114]*Id.*

[115]*See id.* at 104–05.

[116]On the development of the American physician's "professional sovereignty," see generally P. STARR, *supra* note 12.

efforts to reduce the regulatory burden, may not only have to be maintained, but also enhanced.[117]

It is with these potential drawbacks in mind that TEFRA and the proposed PPS must be assessed. Is so radical a change in federal hospital financing advisable, when even the state demonstration projects upon which it is based have achieved a decrease in the annual increase of hospital costs of only two to six percent when compared with the national average?[118] On balance, the answer appears to be affirmative, partly because the state programs *have* had some success, and any success in containing cost is preferable to failure, especially if the quality of care in the demonstration projects does not seem to have suffered appreciably.[119] Moreover, it is difficult to imagine a system more inefficient and unwieldy than the retrospective cost-based reimbursement apparatus now in place. PPS's move toward national uniformity and ultimately even more federal regulatory control is likely to be one step in rationalizing health care in the United States.[120]

On the other hand, a review of the contemporary economics of health care makes it clear that costs can be significantly decreased only if the system is structured so as to keep people out of hospitals for as long as possible, and to make unavoidable

---

[117]The Reagan initiatives are internally inconsistent on the subject of such regulatory review. Although PSRO funding is excluded from the fiscal 1984 Department of Health and Human Services budget, the report to Congress on prospective reimbursement refers to the PSRO function as though it would have a role to play. In addition, it refers to an annual "recalibration" of DRG payment rates so as to account for inflation and "such matters as significant changes in specific diagnostic or treatment technologies . . . ." PROSPECTIVE PAYMENT, *supra* note 7, at 58–59.

[118]*Id.* at 32.

[119]*Id.* at 105.

[120]Symptomatic of this impetus towards a heavy federal regulatory hand is the prospective payment report's position on judicial review:

> Payment amounts, exceptions, adjustments, and rules to implement the prospective payment system would not be subject to any form of judicial review. Retroactive adjustment of the payment rates, as might result from judicial review, is inimical to the basic purpose of a prospective system. Moreover, the delays inherent in the judicial process, when coupled with the likelihood of annual revisions in the rates of payment, could lead to chaotic results, in which rates for a previous period may be overturned by a court, or remanded to the Department for further consideration, even though different rates had superseded the contested rates. The prospect of continuous litigation and reopened administrative proceedings related to supposedly prospective rates for past periods can be prevented by a complete preclusion of judicial review . . . .

Id. at 41. On February 24, 1983 the Health Subcommittee of the House Ways and Means Committee "retained most of the existing authority for administrative and judicial review of Medicare payment decisions, which the administration proposed to eliminate, but excluded DRG payment rates and classifications from administrative and judicial review." 41 CONG. Q. 421 (1983).

hospital stays as short and as inexpensive as is consistent with quality care. That, of course, is a tall order, and it raises large and difficult questions. To its credit, the Reagan administration has addressed some of these questions in its "Health Care Incentives Reform" proposals. The answers embodied in certain of those proposals, however, appear inadequate and inequitable.

No aspect of the budget proposal is more significant, or more problematic, than that to increase the cost sharing requirements for Medicare beneficiaries when they are hospitalized. Under current law, a beneficiary pays in full for his or her first hospital day; for 1984, this deductible amount will be approximately $350.[121] For days two to sixty, the beneficiary pays nothing. For days sixty-one to ninety, the beneficiary pays an amount equal to one fourth of the deductible amount (estimated at $87.50 per day in 1984). If hospitalization beyond ninety days is required, the beneficiary pays half the deductible amount ($175 per day in 1984) for the next sixty days, and the full cost of care thereafter.[122]

The Reagan proposal is based upon the assumption that once a patient is admitted and the $350 deductible is paid, the patient (and attending physician) has no financial incentive to leave the hospital for the next sixty days. This complete insulation from cost encourages excessively lengthy hospital stays. The Reagan administration proposes to make the patient feel the cost by requiring beneficiaries to share costs from the outset on a daily basis. Under the HHS "restructuring initiative," beneficiaries will pay a deductible equal to one day of hospital care (as they do now) but will be required to pay an amount equal to eight percent of the deductible for each day of inpatient hospital care from days two to fifteen and five percent of the deductible for each day of care from days sixteen to sixty.[123] After sixty days, the "catastrophic coverage" proposed by the administration would make further payments by the beneficiary unnecessary for the rest of the year.[124] The administration proposes, however, that a beneficiary be liable to pay the full first-day hospital deductible no more than twice per year (compared with limitless admissions deductibles currently), and that the skilled nursing facility co-payment (applicable from days twenty-one to one

[121]HHS Fact Sheet, *supra* note 9, at 6.
[122]*Id.*; *see also* HHS BUDGET, *supra* note 96, at 28.
[123]HHS BUDGET, *supra* note 96, at 28.
[124]*Id.*

hundred) be reduced from the current 12.5% of the deductible to 5%.[125] According to administration estimates, the new cost sharing and catastrophic provisions would, in combination, save the Medicare program more than $600 million in fiscal year 1984.[126]

The administration argues that the current system "places the greatest financial burden on the sickest beneficiaries."[127] That may be so, and its proposal for catastrophic insurance for the approximately 170,000 Medicare beneficiaries annually[128] who require more than sixty days in an acute care hospital seems unassailable on humanitarian grounds. On the other hand, more than ninety-eight percent of Medicare-financed hospital stays are for less than sixty days. Although for a significant portion of Medicare's thirty million beneficiaries an additional charge of twenty-eight dollars per day[129] may represent an inducement to socially and economically beneficial cost consciousness, for millions of America's aged it will represent an unbearable burden and a strong disincentive to entering the hospital—even when hospital care is a medical necessity.

In addition, a strong argument could be made for questioning one aspect of the status quo that the administration does not propose to change, evidently for budgetary reasons: the full deductible amount payable for the first hospital day ($350 in 1984). If that amount were reduced to, perhaps, $100, with the balance distributed over the next ten days or two weeks, patients would have an incentive to leave the hospital as soon as possible. They would not, however, have to overcome the high barrier to entry that is now in place and that would remain in place (at least for the first two admissions each year) under the administration's proposal. Similarly, the notion of *graduated*

---

[125]*Id.*

[126]HHS Fact Sheet, *supra* note 9, at 7.

—Under current law, the beneficiary hospitalized in 1984 for 150 consecutive days would owe $13,475 from his or her own pocket. The beneficiary would also bear the full cost of all subsequent hospital days.

—Under the new plan, the beneficiary's expenses would be $1,530, with no additional coinsurance after 60 days.

Id.

[127]HHS BUDGET, *supra* note 96, at 28.

[128]HHS Fact Sheet, *supra* note 9, at 6.

[129]A Medicare beneficiary's first hospital day would require a $350 payment, as it does under current regulations. For the first two weeks in hospital, now "free," the beneficiary under the administration's plan would pay $28 daily; for the next six weeks, also currently co-payment-free, the beneficiary will pay $17.50 per day. After the sixtieth day, however, just at the time beneficiaries now begin paying for their care, catastrophic coverage would take effect.

deductibles, keyed to an individual's ability to pay, should be explored.

Other benefit reductions may be viewed as characteristic more of the Reagan administration's general approach to entitlement programs than of particular health-policy objectives: the one-month postponement of Medicare eligibility, the decrease in federal subsidies for voluntary insurance premiums under Medicare Part B, and the continuing cutbacks in federal Medicaid contributions to the states.[130] Other parts of the package, however, are nicely tailored to be consistent with prevailing conceptions of the cost-insensitive health-care marketplace. Thus, where benefit reductions take the form of an increase in Part B deductibles, a reduction in home health agency durable medical equipment reimbursement from one hundred percent to eighty percent of cost, or the requirement of "nominal" cost sharing by Medicaid recipients, such reductions may be rationalized as attempts to increase cost consciousness among the beneficiaries of federal health programs.

In all, the Reagan initiatives constitute an ambitious program of health planning and health-care financing.[131] The initiatives

---

[130]HHS BUDGET, *supra* note 96, at 35. Similarly, the proposals contained in the budget to abolish funding for the PSRO and HSA programs and to establish a "Voluntary Voucher Program" to "provide beneficiaries a choice of electing to receive services through a private health benefits plan, rather than mandating participation in Medicare" have a distinctly "conservative" flavor. HHS BUDGET, *supra* note 96, at 29.

[131]As this issue went to press, President Reagan was expected to sign into law the Social Security Amendments of 1983. Title VI of this law, "Prospective Payments for Medicare Inpatient Hospital Services," provides for a three-year phase-in of the DRG payment mechanism described in this Comment. H.R. REP. NO. 47, 98th Cong., 1st Sess., 129 CONG. REC. H1724–87 (daily ed. Mar. 24, 1983). During the first year, 25% of a hospital's payment will be based on a combination of "national" and "regional" DRG rates (25% national, 75% regional); 75% will be based on retrospectively determined costs (the traditional system). In the second year, 50% of the payment will be based on a combination of national and regional rates (50% each); 50% will be based on each hospital's retrospective cost base. In the third year, 75% of the payment will be based on a combination of national and regional DRG rates (75% national, 25% regional); 25% will be based on each hospital's retrospective cost base. After the third year, 100% of the payment will be determined under the national DRG payment methodology. *Id.* at H1773. The new law also distinguishes between rural and urban hospitals in determining DRG rates. Congress' addition of regional distinctions and an urban/rural distinction across regions renders the DRG payment methodology significantly more complex than the administration's original proposal.

The new law also requires the Office of Technology Assessment to appoint a commission of independent experts to review DRG payment rates annually, beginning in 1986. The DRG payments will be adjusted to take into account inflation, changes in hospital productivity, and scientific and technological advances. *Id.* at H1774. In accord with the administration's proposal, the new law requires the Secretary of Health and Human Services to provide systematically for atypical or "outlier" cases, to reimburse medical education expenses of teaching hospitals on a reasonable cost basis, and to exempt "sole community hospitals." *Id.* at H1775–76. However, apparently recognizing

raise large and important questions, not only of tax policy and incentive structures, but also of the proper balance between regulation and the free market. They implicate issues not only of cost containment and quality of and access to care, but also—in this area of public policy it is not inappropriate to say so—of life and death. As the American "baby boom" generation comes not only of age but to age, and as advances in medical technology permit even catastrophically ill patients to exist on life-support machines for years at tremendous expense, important decisions will have to be made, in hospitals and, perhaps more significantly, in Congress. As the outlines of the Reagan initiatives emerge, they suggest the parameters of the debate on health legislation for this decade and beyond. They suggest, too, that the debate can be restrictcd to tax laws and budget proposals for only so long—that in the foreseeable future, plenary consideration and plenary legislation will be in order. Indeed, like death and taxes themselves, they seem inevitable.

---

the threat that a hospital working within rigid cost constraints might attempt to minimize costs by sacrificing quality of care. Congress rejected the administration's attempt to eliminate peer review of utilization and quality of care. *Id.* at H1776–77. The new law also leaves current mechanisms for administrative and judicial review of reimbursement decisions intact. *Id.* at H1778. In addition, the law requires that hospitals continue to file extensive cost reports, which will enable individual states to adjust their Medicaid reporting requirements, and will enable the federal government to revert to a retrospective cost reimbursement system should the prospective payment program be discontinued. *Id.* at H1773. The law, however, anticipates the ultimate success and continuation of the program and specifically mandates reports on the feasibility of extending prospective payment to capital expenditures, physicians' services, and certain currently exempted facilities such as skilled nursing care institutions. *Id.* at H1775–76. To facilitate congressional evaluation of the efficacy of the new prospective system, the new law requires the Secretary of Health and Human Services to monitor its implementation closely, *id.* at H1776–77, to continue certain demonstration projects, and to report regularly to Congress on the impact of the new system on specified classes of hospitals, providers, beneficiaries, and third-party payers. *Id.* at H1778.

The enactment of prospective payment received very little media attention compared with the efforts of Congress to "save" the Social Security system, which constituted the bulk of the legislation. The relative speed with which prospective payment passed and the bipartisan support it received testify to the effectiveness of the administration's "low profile" health-care strategy. Whether those aspects of the Reagan initiatives that more directly affect Medicare beneficiaries and the health insurance industry, such as accelerated cost-sharing and amendments to I.R.C. § 106, will meet with similar success remains to be seen.

COMMENT

THE RIGHT TO TRICK-OR-TREAT: CONSTITUTIONAL IMPLICATIONS OF HALLOWEEN ORDINANCES

JORDAN LIPKA*
FRANK GIORDANELLA**

The Tylenol contamination incident of September 1982 sparked a concern over food and drug tampering throughout the country.[1] In response to this concern, several communities banned, considered banning, or advised against trick-or-treating[2] this past Halloween.[3] This Comment will deal with several of the constitutional challenges that might be advanced against local laws prohibiting children from trick-or-treating, and will assess the likelihood of their success. Section I discusses the history of the holiday. Section II analyzes the Commerce

---

*B.A., Baruch College, 1981; member, Class of 1984, Harvard Law School.
**B.S.F.S., Georgetown University, 1981; member, Class of 1984, Harvard Law School.

[1]*See generally* Beck, *The Tylenol Scare*, NEWSWEEK, Oct. 11, 1982, at 32; Tifft, *Poisoning Madness in the Midwest*, TIME, Oct. 11, 1982, at 18.

[2]Trick-or-treat is the custom, practiced by children, of dressing up in costumes and masks (usually of characters in fairy tales, comic books, or television programs) on Halloween and traveling from door to door asking for candy. The children, often in groups or with parents, ring the doorbell or knock on the door and say "trick or treat" when the occupant answers. The person answering gives the children candy (or sometimes a toy or money), which the children collect in a bag. The activity is usually restricted to a specific neighborhood or apartment building. The exact procedure may vary in different areas.

[3]*See, e.g.*, Boston Globe, Nov. 1, 1982, at 10, col. 1; Daley, *Fear of Tainted Candy Prompts Wide Concern for Halloweeners*, N.Y. Times, Oct. 30, 1982, at A6, col. 1; Barron, *Poison Worries Lead to Precautions for Halloween*, N.Y. Times, Oct. 28, 1982, at B1, col. 1. Many cities warned against trick-or-treating, but 40 to 50 communities were reported to have banned the activity. Boston Globe, Nov. 1, 1982, at 10, col. 1; *Daley*, supra. Phone calls by the authors to town clerks in several Massachusetts towns reported to have prohibited trick-or-treating, however, revealed that most of the towns' selectmen merely warned against trick-or-treating and arranged substitutes such as parties. Telephone interviews with town clerks of Fitchburg, Methuen, Dudley, Millbury, Ashby, and Barnstable (Feb. 10, 1982). A telephone interview with Suzanne Daley (Feb. 18, 1982) revealed that the towns referred to in her article, *supra*, also did not pass any written ordinance if they did ban the activity.

Two Massachusetts towns whose boards of selectmen did pass written resolutions were Palmer and Holland. The Holland resolution reads, "There will be no trick or treat this year in Holland. This will be enforced by the Holland Police Department." Notice of Holland, Mass., Bd. of Selectmen, Oct. 1982. In Palmer, the Board of Selectmen's minutes state, "On a motion . . . it was voted to ban Trick or Treating in the Town of Palmer this year." Minutes of Board of Selectmen, Town of Palmer, Mass., Oct. 19, 1982, at 1.

Clause and arguments based on the commercial nature of Halloween. Section III discusses the liberty aspects of the prohibition. Part A of that section sets out a general framework of the protection afforded by the Fourteenth Amendment. Part B analyzes the specific liberty interests in the right of speech and assembly, while Part C deals with the constitutional freedom of religion. Part D focuses upon the assertion that there is a compelling state interest that overrides the liberty interests.

I. The History of Halloween

The celebration of Halloween dates back to the Celtic-Druid communities of northern and western Europe in the second century B.C.[4] The purpose of the holiday ceremony was to thank the sun god for the harvest of the past year and to honor Samhain, the lord of the dead.[5] During the ninth century, Christianity incorporated the holiday as All Hallows' Eve, the vigil preceding All Saints' Day.[6] Also at this time, demonic religions sprang up and designated Halloween, due to its close association with dead spirits, as one of their principal holidays.[7] In addition to these religious aspects, the connection to the spirit world brought more lighthearted forms of celebration, such as fortunetelling.[8] This amalgamation of customs and influences was brought to the United States by the Irish immigration of the mid-1800's.[9]

The major form of American celebration of Halloween is trick-or-treating. The exact origins of this tradition remain obscure. Several theories have been developed to explain the beginning of the practice. First, trick-or-treating has been linked to an Irish custom in which groups of people traveled from house to house asking for gifts to commemorate the Druid god Muck Olla

---

[4]J. Hatch, The American Book of Days 968 (1978); R. Linton & A. Linton, Halloween Through Twenty Centuries 4 (1950) [hereinafter cited as Linton].

[5]On October 31, Samhaim gathered all the souls of people who had died the previous year and who had been undergoing penance in the form of animals, to determine which of the souls would go to Druid heaven and which would be returned to animal form. With the Roman conquest came the integration of Roman religious rites, especially the November celebration in honor of the goddess Pomona, into Halloween. J. Hatch, *supra* note 4, at 968; Linton, *supra* note 4, at 4–5.

[6]J. Hatch, *supra* note 4, at 969; Linton, *supra* note 4, at 6 (All Saints' Day honors those saints who died unrecognized.).

[7]J. Hatch, *supra* note 4, at 969; Linton, *supra* note 4, at 8–9, 48–49.

[8]J. Hatch, *supra* note 4, at 969–70; Linton, *supra* note 4, at 23–25.

[9]J. Hatch, *supra* note 4, at 970; Linton, *supra* note 4, at 100–01.

or the Catholic saint Columba.[10] Generous givers were to be rewarded with prosperity in the coming year, while those who failed to give or gave too little were threatened with bad luck.[11] Another religious custom that may be the foundation of trick-or-treating is the English practice of "souling," in which people went from house to house on All Souls' Day asking for "soul cakes," in exchange for which prayers were offered for the giver's deceased relatives.[12] Finally, trick-or-treating may be based on a secular English tradition associated with Guy Fawkes Day, November 5, in which children roamed the streets asking for "a penny for the Guy."[13] Although it is uncertain from which of these ancient customs trick-or-treating is derived, the practice has become firmly rooted in the United States as the primary means of celebrating Halloween.

II. THE COMMERCE CLAUSE

One of the constitutional provisions with which Halloween ordinances must comply is the Commerce Clause.[14] The Commerce Clause may be an impediment to municipalities that wish to prohibit trick-or-treating, primarily because of the burden that such bans place on the high volume[15] of interstate commerce in candy that occurs in anticipation of the holiday.[16] The arguments in support of a dominant federal interest in this area, besides being based upon the sheer volume of sales, would focus upon the need for a uniform system of law to prevent the disruption that would be caused by a haphazard series of bans scattered across the country.[17] This disruption would primarily be the

---

[10]J. HATCH, *supra* note 4, at 971; LINTON, *supra* note 4, at 102–03.

[11]J. HATCH, *supra* note 4, at 971; LINTON, *supra* note 4, at 102–03.

[12]All Souls' Day is celebrated on November 2. It commemorates those people who, while not saints, died as Christians. J. HATCH, *supra* note 4, at 971; LINTON, *supra* note 4, at 16–17, 102.

[13]Guy Fawkes Day commemorates the foiling of the Gunpowder Plot of 1605 to kill King James I of England. J. HATCH, *supra* note 4, at 971.

[14]U.S. CONST. art. I, § 8, cl. 3 ("The Congress shall have Power . . . To regulate Commerce with foreign Nations, and among the several States, and with the Indian Tribes."); *see* Gibbons v. Ogden, 22 U.S. (9 Wheat.) 1 (1824).

[15]Confectionary wholesale sales rose from $2.48 billion in 1978 to $2.88 billion in 1979 and $3.11 billion in 1980. In all three years wholesale sales were highest in the month of October. U.S. BUREAU OF THE CENSUS, DEP'T OF COMMERCE, REVISED MONTHLY WHOLESALE TRADE: SALES AND INVENTORY, JAN. 1973–DEC. 1980, at 11–13 (1981).

[16]N.Y. Times, Nov. 2, 1982, at A14, col. 6 (retail sales of candy fell in 1982 between 20% and 50%).

[17]*See* Exxon Corp. v. Governor of Md., 437 U.S. 117, 128 (1978); Wabash, St. L. &

result of increased uncertainty and risk in industries that manufacture or distribute goods for Halloween.[18]

A town defending against a Commerce Clause attack in court is aided by two factors. First, the Halloween ordinances do not seek to favor local products by discriminating against products originating in other states and traveling through interstate commerce. The bans affect all candy sales regardless of the point of origin. This equality of burden is significant because it relieves the municipalities of the stricter standard usually applied by the courts to discriminatory regulations,[19] and instead triggers a "balancing" test.[20] As a result, the validity of the ordinances will be judged by weighing the federal interest in preserving the flow of interstate commerce against the state interest, operating through each individual municipality, in the health and safety of its children. Second, when implementing the balancing test there is a presumption in favor of ordinances that protect legitimate local interests and only incidentally affect interstate commerce; the burden on interstate commerce must be excessive for the Halloween ordinances to be held invalid.[21]

P. Ry. v. Illinois, 118 U.S. 557, 567 (1886); *see also* Cooley v. Board of Wardens, 53 U.S. (12 How.) 299, 319 (1851).

[18]For manufacturers, such as candy producers, there is a substantial risk of either overproducing or underproducing because they will not know the number of towns that will institute Halloween bans until a few days before trick-or-treating begins. In the case of distributors, primarily trucking companies, set delivery routes may have to be altered and longer or more circuitous routes required, which would increase costs and reduce profitability and efficiency.

[19]Lewis v. BT Inv. Management, 447 U.S. 27, 36–37 (1980); Exxon Corp. v. Governor of Md., 437 U.S. at 126–27; Breard v. Alexandria, 341 U.S. 622, 636–37 (1951); *see* Dean Milk Co. v. Madison, 340 U.S. 349, 354 (1951).

[20]The balancing test has been described using various terminology. At times the Court will explicitly use the word "balancing" in its analysis. In other cases, the Court uses a "direct-indirect" approach. In either case the Court balances the state's interest against the burden placed on interstate commerce. Great Atl. & Pac. Tea Co. v. Cottrell, 424 U.S. 366, 371–72 (1976) (balance); Pike v. Bruce Church, 397 U.S. 137, 142 (1970) (direct-indirect); H.P. Hood & Sons v. DuMond, 336 U.S. 525, 552 (1949) (Black, J., dissenting) (balance). *See generally* Dowling, *Interstate Commerce and State Power*, 27 VA. L. REV. 1 (1940).

[21]*See* Hunt v. Washington Apple Advertising Comm'n, 432 U.S. 333, 353 (1977); Pike v. Bruce Church, 397 U.S. at 142; Huron Portland Cement Co. v. Detroit, 362 U.S. 440, 443 (1960).

> State regulations seemingly aimed at furthering public health or safety . . . are less likely to be perceived as undue burdens on interstate commerce than are state regulations evidently seeking to maximize the profits of local businesses. Indeed, where the Supreme Court has held that the national interest in the free flow of commerce supercedes [sic] a state interest in public safety, it has generally seemed that the challenged statute contributed only marginally if at all to the public safety. In contrast, economically based state regulations have almost invariably been struck down.

L. TRIBE, AMERICAN CONSTITUTIONAL LAW § 6-12, at 340 (1978).

The municipalities also may defend their ordinances on the basis of the traditional police power vested in the states to protect the general health, welfare, safety, and morals of their citizens.[22] Bans on trick-or-treating are designed to protect children within a town from the danger of tampered treats, a danger made greater in 1982 by the Tylenol poisonings.[23] The power of the state and its subdivisions to enact this type of regulation has been recognized in cases concerning transportation,[24] incoming commerce,[25] and outgoing commerce.[26] There have been instances, however, when the police power rationale was not accepted by the courts.[27] The major reason for the denial of power in these cases was that the regulations were motivated by economic considerations and did not primarily focus on safety.[28] In contrast, the anti-trick-or-treating laws should not fail on this ground.

Moreover, the towns can bolster their arguments by specifically relying upon *Breard v. Alexandria*.[29] In *Breard*, the Supreme Court upheld a municipal ordinance that banned door-to-door, uninvited, commercial solicitations against a Commerce Clause attack.[30] The Court held that the municipality's interest in preserving the privacy rights of its citizens outweighed the Commerce Clause interest in the volume of interstate magazine sales.[31] The Halloween ordinances are similar to the *Breard* ordinance in that they ban door-to-door solicitations of treats for the important public policy reason of protecting the towns' children from poisoning. It follows from *Breard* that the incidental effect on the flow of interstate commerce would not be sufficient to invalidate the ordinances. Although the banned solicitations under the Halloween regulations often are encouraged by the homeowner and the Halloween ban is designed to

---

[22]*See, e.g.*, Great Atl. & Pac. Tea Co. v. Cottrell, 424 U.S. at 371; H.P. Hood & Sons v. DuMond, 336 U.S. at 531–32.

[23]*See* Beck, *supra* note 1; Tifft, *supra* note 1.

[24]*E.g.*, Raymond Motor Transp. v. Rice, 434 U.S. 429 (1978); South Carolina State Highway Dep't v. Barnwell Bros., 303 U.S. 177 (1938); Bradley v. Public Util. Comm'n, 289 U.S. 92 (1933).

[25]Breard v. Alexandria, 341 U.S. 622 (1951).

[26]Milk Control Bd. v. Eisenberg Farm Prods., 306 U.S. 346 (1939).

[27]*See, e.g.*, Philadelphia v. New Jersey, 437 U.S. 617 (1978); Baldwin v. G.A.F. Seelig, Inc., 294 U.S. 511 (1935).

[28]Philadelphia v. New Jersey, 437 U.S. at 624 (economic protection is per se invalid); *see* Baldwin v. G.A.F. Seelig, Inc., 294 U.S. at 511.

[29]341 U.S. 622 (1951).

[30]*Id.* at 633–41.

[31]*Id.* at 640–41.

protect the solicitor rather than the party solicited, these differences do not alter the basic proposition in *Breard* that solicitations can be prohibited to support an overriding local public policy interest without violating the Commerce Clause.

The opponents of the Halloween ordinances might try to distinguish *Breard* on the ground that the safety interest asserted by the municipalities is not as significant as the privacy interest in *Breard*. There is a series of cases that state that burdens on interstate commerce will not be allowed unless they substantially increase safety; mere marginal benefits are not enough to justify a burden on the flow of interstate commerce.[32] Bans on Halloween might fall into this category for several reasons. First, because the prohibitions were imposed by only a small handful of towns, the regulations could be evaded by children who live close to other municipalities that permit trick-or-treating. The municipalities may argue, however, that evasion is unlikely because most small children have a very limited traveling range and the distance between towns, especially in rural areas, is often very large. In addition, the municipalities may say that the likelihood of evasion is irrelevant because in other legal areas, most notably in drinking age statutes, states are able to act even though their citizens are able to avoid the regulations by going to other states. Second, few candy tampering incidents occur.[33] The towns may reply, however, that they should not have to wait until an incident occurs before they can act to prevent it. Finally, the broadly drawn ordinances prohibit children from soliciting treats from relatives or close family friends from whom there is even less of a chance of poisoning. The

[32]Kassel v. Consolidated Freightways, 450 U.S. 662, 670–71 (1981); Raymond Motor Transp. v. Rice, 434 U.S. 429, 443–44 (1978); Bibb v. Navajo Freight Lines, 359 U.S. 520, 524 (1959); Southern Pac. Co. v. Arizona, 325 U.S. 761, 775–76 (1945).

[33]*See* DUDLEY-ANDERSON-YUTZY, INDUSTRY SURVEY CONFIRMS ONLY TWO INJURIES FROM HALLOWEEN CANDY 3 (Nov. 1982) (available from the Nat'l Confectioners Ass'n, Chicago, Ill.). This report surveyed police stations in 24 major metropolitan areas to ascertain the total number of candy tamperings and injuries during the Halloween of 1982. Ten cities—Atlanta, Cleveland, Denver, Miami, Minneapolis-St. Paul, New Orleans, Portland (Or.), St. Louis, San Jose, and Washington—did not report any instances of tampering. Seventy-six occurrences of contamination were reported in the following 13 cities: Boston (five); Cincinnati (six); Dallas (two); Detroit (six); Fort Lauderdale (one); Hartford (one); Kansas City (Mo.) (two); Milwaukee (10); New York (30); Pittsburgh (seven—with three confirmed and four under investigation); San Diego (two); San Francisco (one); and Seattle (three). Only two actual injuries from this tampering were reported. *See also* Boston Globe, Nov. 1, 1982, at 10, col. 1 (informal count by Associated Press of 175 reports of sabotaged fruits and candies in 100 cities in 24 states); Daley, *Caution Can't Mask Holiday's True Face: Halloween's a Party*, N.Y. Times, Nov. 1, 1982, at B3, col. 1 (reporting at least 30 incidents of tampered candy in New York City between Oct. 29 and Oct. 31).

towns may contend that a general statute is needed because of the enforcement difficulties that would be caused by hard-to-define exceptions.

An additional factor, when weighed in the balancing test, may favor the interstate commerce interest over the safety interest of the municipalities. Courts adjudicating the validity of town ordinances under the Commerce Clause must evaluate the availability of alternatives that would fulfill the purposes of the ordinances while imposing less of a burden on interstate commerce.[34] There are several alternatives open to municipalities that may protect their children and do not hinder interstate candy sales as much as would a total ban on trick-or-treating.[35] Any alternative, however, will not decrease the risk of contamination as much as a total ban on trick-or-treating.[36] Because the safety of their citizens is a local matter, the towns may conclude that they have the responsibility to determine the acceptable level of risk.

A court faced with adjudicating a Commerce Clause attack on the Halloween ordinances thus will weigh the safety interest advanced by the municipalities against the federal interest in the large volume of interstate commerce in candy and the need for uniform legislation. Because the state interest in the safety of its children appears to be quite strong, it can be concluded that a ban on trick-or-treating will most likely withstand a Commerce Clause attack.

III. THE LIBERTY INTERESTS

A. *The Fourteenth Amendment and Substantive Due Process*

The Fourteenth Amendment protects individuals from the deprivation of their liberty by state governments without due process of law.[37] This protection extends to both the substantive

---

[34]Hunt v. Washington Apple Advertising Comm'n, 432 U.S. 333, 353 (1977); Great Atl. & Pac. Tea Co. v. Cottrell, 424 U.S. 366, 371–73 (1976); Pike v. Bruce Church, 397 U.S. 137, 142 (1970).

[35]*See infra* Section III, Part D.

[36]For example, even at government-sponsored parties, it is possible that contaminated treats will be given to children. *See Sixteen Children Hospitalized After Party*, Boston Globe, Nov. 1, 1982, at 10, col. 1.

[37]"[N]or shall any State deprive any person of life, liberty, or property, without due process of law." U.S. CONST. amend. XIV, § 1; *see* Youngberg v. Romeo, 102 S. Ct. 2452, 2458 (1982).

and procedural aspects of due process.[38] An ordinance prohibiting trick-or-treating may deprive the affected children of a liberty interest without substantive due process.

The Supreme Court has never precisely defined liberty,[39] but its decisions provide some guidance. In *Meyer v. Nebraska*, the Court spoke of "freedom from bodily restraint" and "enjoy[ing] those privileges long recognized at common law as essential to the orderly pursuit of happiness by free men."[40] In *Griswold v. Connecticut*, Justice Douglas noted that the rights protected by the Fourteenth Amendment are not merely those that have been enumerated in the Bill of Rights.[41] Justice Harlan, dissenting in *Poe v. Ullman*, wrote that appropriate limits on substantive due process come from the balance this country "has struck between . . . liberty and the demands of organized society," with "regard to what history teaches are the traditions from which [the liberty interest] developed as well as the traditions from which it broke."[42]

B. *Freedom of Speech and of Assembly*

Trick-or-treating involves two First Amendment issues.[43] Speech interests are implicated because the children say "trick or treat" in soliciting candy from their neighbors.[44] Speech and assembly interests are involved because the children and adults

---

[38]"[The Fourteenth Amendment] affords not only a procedural guarantee against the deprivation of 'liberty,' but likewise protects substantive aspects of liberty against unconstitutional restrictions by the State." Kelley v. Johnson, 425 U.S. 238, 244 (1976).

[39]*See* Board of Regents v. Roth, 408 U.S. 564, 572 (1972); Bolling v. Sharpe, 347 U.S. 497, 499–500 (1954); Meyer v. Nebraska, 262 U.S. 390, 399 (1923).

[40]262 U.S. at 399; *see also Youngberg*, 102 S. Ct. at 2458.

[41]381 U.S. 479, 482–83 (1965); *see also* Moore v. City of East Cleveland, 431 U.S. 494, 502 (1977) (noting the risks inherent "when the judicial branch gives enhanced protection to certain substantive liberties without the guidance of the more specific provisions of the Bill of Rights"); Roe v. Wade, 410 U.S. 113 (1973); Board of Regents v. Roth, 408 U.S. at 572 ("In a Constitution for a free people there can be no doubt that the meaning of liberty must be broad indeed."); Bolling v. Sharpe, 347 U.S. at 499 ("Liberty under law extends to the full range of conduct which the individual is free to pursue . . . .").

[42]367 U.S. 497, 542 (1961).

[43]Freedom of speech and assembly are liberties protected from state infringement by the Due Process Clause of the Fourteenth Amendment. *See, e.g.*, NAACP v. Claiborne Hardware Co., 102 S. Ct. 3409 (1982); NAACP v. Button, 371 U.S. 415 (1963); Gitlow v. New York, 268 U.S. 652 (1925).

[44]While the child could hold his or her bag open and not say anything, this would still be "asking" for candy, albeit symbolically. The speech would be entitled to the same protection. *Cf.* Tinker v. Des Moines Indep. Community School Dist., 393 U.S. 503 (1969); United States v. O'Brien, 391 U.S. 367 (1968).

meet each other on the doorstep for the purposes of exchanging candy and socializing. These speech and assembly interests are social[45] in nature, however, and whether or not they are protected as such must be examined.[46]

While the Supreme Court has stated that there is "no absolute right under the Federal Constitution to enter on the private premises of another and knock on a door for any purpose,"[47] it also has noted that charitable appeals for funds involve a variety of speech interests that are within the protection of the First Amendment.[48] The regulation of solicitation must not intrude upon the rights of free speech and free assembly,[49] and any restrictions must be narrowly drawn.[50]

Solicitation cases have involved the distribution[51] and sale[52] of religious literature, the distribution of political literature,[53] membership in unions,[54] and charitable appeals for funds.[55] The Court upheld the protection of the First Amendment interests in each case. While the Court in *Breard v. Alexandria* upheld the prohibition of canvassing for commercial purposes without permission of the homeowner,[56] the recent judicial protection given to commercial speech may indicate that this doctrine no longer holds true.[57] The commercial speech argument, in conjunction with the other liberty interests, may be enough to strike down the Halloween ordinances.

The social speech and assembly involved here also may warrant protection. Not only is assembly for the purpose of ad-

---

[45]For the purposes of this Comment, a social association or social speech refers to topics involving family, friends, hobbies, etc., and not to subjects in the political, governmental, or commercial areas.

[46]Although speech and assembly are not prohibited per se, they are so integrally involved in trick-or-treating that it cannot be said that they are only incidentally affected. However, even incidentally affected speech or assembly is protected. *See* Schad v. Borough of Mt. Ephraim, 452 U.S. 61 (1981); United States v. O'Brien, 391 U.S. 367 (1968); UMW, Dist. 12 v. Illinois State Bar Ass'n, 389 U.S. 217 (1967); Bates v. City of Little Rock, 361 U.S. 516 (1960).

[47]Hynes v. Mayor of Oradell, 425 U.S. 610, 619 (1976).

[48]Village of Schaumberg v. Citizens for a Better Env't, 444 U.S. 620, 636 (1980).

[49]*Hynes*, 425 U.S. at 616; Thomas v. Collins, 323 U.S. 516, 540 (1945).

[50]*Village of Schaumberg*, 444 U.S. at 637; *Hynes*, 425 U.S. at 620.

[51]Martin v. City of Struthers, 319 U.S. 141 (1943).

[52]Hynes v. Mayor of Oradell, 425 U.S. 610 (1976).

[53]Murdock v. Pennsylvania, 319 U.S. 105 (1943).

[54]Thomas v. Collins, 323 U.S. 516 (1945).

[55]Village of Schaumberg v. Citizens for a Better Env't, 444 U.S. 620 (1980).

[56]341 U.S. 622 (1951); *accord* Valentine v. Chrestensen, 316 U.S. 52 (1942).

[57]*Village of Schaumberg*, 444 U.S. at 632 n.7. *But see* Village of Hoffman Estates v. Flipside, Hoffman Estates, Inc., 455 U.S. 489, 496–97 (1982) (refusing to apply overbreadth doctrine to commercial speech).

vancing beliefs and ideas protected,[58] but the Court also has signaled that social meetings warrant protection. Assembly "for any lawful purpose" is a right of personal liberty.[59] Nor does speech have to involve ideas and opinions on political and commercial matters to be shielded. "[T]he rights of free speech and a free press are not confined to any field of human interest."[60] Moreover, since conversations often include social and political matters, a prohibition on social speech may chill the exercise of other freedoms protected by the Fourteenth Amendment.[61] It seems likely, therefore, that purely social speech is a liberty falling under the First and Fourteenth Amendments.

Social speech also conveys ideas and opinions. "Men are entitled to speak as they please on matters vital to them."[62] Family, friends, hobbies, and like topics that comprise social speech are important to individuals. For example, deciding which college to attend or what job to take are matters that greatly affect people's lives and deserve unencumbered discussion.

Advocates for the towns may argue that what is involved here, however, is not as important. Asking for candy is neither vital to our democracy nor as significant as college or work. The children could talk to a neighbor about an important matter without soliciting candy. Speech associated with trick-or-treating should not rise to a level that impedes the government's ability to protect its citizens.

This particular form of speech, however, may be important to the children as part of an event they look forward to and enjoy. The speech also disseminates the "idea" of Halloween and trick-or-treating. Furthermore, it is difficult to decide what social speech is significant enough to be protected. Much political speech is unimportant or superfluous, yet it is guarded.

The towns may try to weaken the speech interest involved by

---

[58]*See, e.g.*, Gibson v. Florida Legislative Investigation Comm., 372 U.S. 539 (1963).

[59]Hague v. Committee for Indus. Org., 307 U.S. 496, 519 (1939); *see also* Coates v. City of Cincinnati, 402 U.S. 611, 615 (1971); Schneider v. Smith, 390 U.S. 17, 25 (1968).

[60]Thomas v. Collins, 323 U.S. 516, 531 (1945); *see* Abood v. Detroit Bd. of Educ., 431 U.S. 209, 231 (1977); Winters v. New York, 333 U.S. 507, 528 (1948) (Frankfurter, J., dissenting) ("Wholly neutral futilities, of course, come under the protection of free speech as fully as do Keats' poems or Donne's sermons.").

[61]*See, e.g.*, Richmond Newspapers v. Virginia, 448 U.S. 555, 577 (1980) (Assembly is "regarded not only as an independent right but also as a catalyst to augment the free exercise of the other First Amendment rights.").

[62]Wood v. Georgia, 370 U.S. 375, 389 (1962).

asserting that it is commercial speech and therefore entitled to less protection than social speech.[63] The children are essentially begging for candy. Trick-or-treating involves a transaction wherein the child is given candy in return for not playing a prank on the adult or as a reward for wearing a costume. The adults must spend money to buy this candy.

The response to this argument is that trick-or-treaters are not in a business for profit. While they do want gifts of candy, most are not seeking a payoff for not playing a trick. Mere solicitation of funds does not make an activity commercial,[64] and the underlying basis of trick-or-treating is the celebration of a holiday.

C. *Freedom of Religion*

The Supreme Court has held that the Free Exercise Clause of the First Amendment[65] applies to the states through the Fourteenth Amendment.[66] As a result, any state action, including that of municipalities operating under a grant of authority from the state as agent of the state, which might directly or indirectly impinge on religious activity must be scrutinized. Although it has rather extensive religious roots, there is doubt as to whether Halloween retains its former religious connotations.[67] Many sociologists consider Halloween to be a "degenerate" holiday observed more for its revelry than for its original religious purpose.[68] There do not appear to be any modern religions that incorporate trick-or-treating as a religious rite and the practice seems void of any religious meaning. This lack of religious meaning would indicate that trick-or-treating is not protected by the Free Exercise Clause, as only those practices that are tied to religious beliefs are entitled to this protection.[69]

---

[63]*See, e.g.*, Metromedia v. City of San Diego, 453 U.S. 490 (1981); *see also* text accompanying notes 72–78.

[64]Village of Schaumberg v. Citizens for a Better Env't, 444 U.S. 620, 632 (1980); Murdock v. Pennsylvania, 319 U.S. 105, 111 (1943).

[65]U.S. CONST. amend. I ("Congress shall make no law respecting an establishment of religion or prohibiting the free exercise thereof . . . .").

[66]California v. Grace Brethren Church, 102 S. Ct. 2498 (1982); Cantwell v. Connecticut, 310 U.S. 296, 303 (1940).

[67]J. HATCH, *supra* note 4, at 868–69; LINTON, *supra* note 4, at 3–9.

[68]J. HATCH, *supra* note 4, at 972; LINTON, *supra* note 4, at 104.

[69]Wisconsin v. Yoder, 406 U.S. 205, 215 (1972).

D. *The Tests*

When laws do not offend liberties protected by the Fourteenth Amendment, they are constitutional if they are not arbitrary and they have a reasonable relationship to their purpose.[70] The ordinances banning trick-or-treating clearly meet this test.

Should trick-or-treating be deemed commercial speech, the law must meet the four-pronged test enunciated in *Central Hudson Gas & Electric Corp. v. Public Service Commission.*[71] The speech, to warrant protection, must concern a lawful activity and not be misleading. The government can regulate the speech if the regulatory statute implements a substantial governmental interest, directly advances that interest, and reaches no further than necessary to accomplish the given objective.

The test is stricter when other forms of speech and assembly are infringed. The Supreme Court has stated that, when free speech is infringed by an anti-solicitation statute, the law "cannot be sustained unless it serves a sufficiently strong, subordinating interest that the [towns] are entitled to protect."[72] Other cases have used the "serving a compelling state interest" test when a liberty interest protected by the Fourteenth Amendment was limited.[73] While it is not absolutely clear, "sufficiently strong" and "compelling" appear to be closely related. However, because social speech is generally not "[i]deological expression . . . integrally related to the exposition of . . . thought that may shape our concepts of the whole universe of man,"[74] it may

---

[70]*See, e.g.*, Duke Power Co. v. Carolina Envtl. Study Group, 438 U.S. 59, 83 (1978) (economic area); Whalen v. Roe, 429 U.S. 589, 597 n.20 (1977) (quoting New York State Ice Co. v. Liebmann, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting)).

[71]447 U.S. 557, 566 (1980). This test was followed in Metromedia v. City of San Diego, 453 U.S. 490, 507 (1981), and in *In re* R.M.J., 102 S. Ct. 929, 937 n.15 (1982).

[72]Village of Schaumberg v. Citizens for a Better Env't, 444 U.S. 620, 636. While the Court in Breard v. Alexandria, 341 U.S. 622, 644–45 (1951), stated that the privacy interest of homeowners outweighed the First Amendment interest in distributing magazines, there is no right of privacy at issue here. Furthermore, there were alternatives to door-to-door soliciting in *Breard*, such as telephone calls and newspaper advertisements, which do not exist here. For a discussion of these purposes in a case that invalidated an anti-solicitation law, see Martin v. Struthers, 319 U.S. 144, 144–45 (1943).

[73]Brown v. Socialist Workers '74 Campaign Comm., 103 S. Ct. 416, 420 (1982); Consolidated Edison v. Public Serv. Comm'n of N.Y., 447 U.S. 530, 540 (1980); First Nat'l Bank v. Bellotti, 435 U.S. 765, 786 (1978); Roe v. Wade, 410 U.S. 113, 178 (1973); Sherbert v. Verner, 374 U.S. 398, 404 (1963); Gibson v. Florida Legislative Investigation Comm., 372 U.S. 539, 554 (1963); Bates v. City of Little Rock, 361 U.S. 516, 524 (1960); NAACP v. Alabama, 357 U.S. 449, 463 (1958).

[74]Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, 425 U.S. 748, 779 (1976) (Stewart, J., concurring).

receive a lesser degree of protection, much like that of commercial speech. Nevertheless, the government cannot choose means that broadly stifle liberties when the same end may be more narrowly achieved.[75]

Although the Supreme Court has held that children "are protected by the Constitution and possess Constitutional rights,"[76] the Court also has held that the state has a somewhat broader authority to regulate the activities of children than of adults.[77] The Court has said that this broader regulatory authority is justified by the peculiar vulnerability of children, their inability to make critical decisions in an informed, mature manner, and the importance of the parental role in child rearing.[78] The exact test used to determine when a child's constitutional liberty interest has been infringed, however, has not yet been formulated.[79] This may lead to different standards in different courts.

Under any level of scrutiny, however, the town's "interest in safeguard[ing] the physical and psychological well-being of a child is compelling."[80] We must examine whether or not the laws in question serve that interest. The purpose of banning trick-or-treating is to protect children from injury that could occur from eating tampered candy.[81] A ban is the best way to protect children, because they will not be receiving *any* candy. Even if parents were to accompany children and to inspect the candy, it would be difficult for them to tell in every case whether candy had been contaminated. Parents, however, may argue that they should be able to decide for themselves, without governmental interference, whether to send their children trick-or-treating. Moreover, the laws may be overinclusive, because children are prohibited from trick-or-treating at every house, including those of relatives and close family friends, where there is little chance of receiving tampered candy. When dealing with

---

[75]Dunn v. Blumstein, 405 U.S. 330, 343 (1972); Braunheld v. Brown, 366 U.S. 599, 607 (1961); Shelton v. Tucker, 364 U.S. 479, 488 (1960).

[76]Planned Parenthood v. Danforth, 428 U.S. 52, 74 (1976); *see In re* Gault, 387 U.S. 1 (1967).

[77]*Danforth*, 428 U.S. at 74–75; *see also* Ginsberg v. New York, 390 U.S. 629 (1968); Prince v. Massachusetts, 321 U.S. 158, 170 (1944).

[78]Bellotti v. Baird, 443 U.S. 622, 633–39 (1979); *see* H.L. v. Matheson, 450 U.S. 398 (1981).

[79]Bykofsky v. Borough of Middletown, 401 F. Supp. 1242 (M.D. Pa. 1975), *aff'd mem.*, 535 F.2d 1245 (3d Cir.), *cert. denied*, 429 U.S. 964 (1976).

[80]Globe Newspaper v. Superior Court, 102 S. Ct. 2613, 2621 (1982).

[81]*See generally* Daley, *supra* note 3; Fox, *Trick or Treat Bans Grow*, Boston Globe, Oct. 21, 1981, at 1, col. 1.

children, though, there is a need for a generalized rule. Without one, the law would be harder to enforce; police would not know which houses children are allowed to go to, and the children might knock on doors that are off limits. The towns should not have to wait until children are harmed before enacting legislation to protect them.[82] In many small towns, however, the chance of injury is slight.[83] The danger may not justify the bans, particularly in that the prohibition encourages children to be suspicious and fearful of adults.[84]

Futhermore, there are alternatives to a total prohibition. While the state may not inhibit liberty interests broadly when narrower means are available and the same basic end can be achieved, it is unclear how this requirement will be applied in light of the state's greater ability to regulate the activity of children. Should a court take alternatives into account, there are several it could consider.

A) The town could hold a party, parade, or other event with incentives such as prizes and costume contests for attending, but allow the children to go trick-or-treating instead of prohibiting it. Many children in small towns will choose to attend the party and forego trick-or-treating.

B) The town could require parents to accompany children so that they will not eat any candy until their parents have inspected it.

C) The town or the state could increase the penalties for those who tamper with candy.

D) The town could either encourage parents who accompany children to make a list of candies received and the person who gave them so that any candy tampered with could be traced, or require people who give candy to put it in bags with their names on them.

E) The town could have an informational program to warn parents and children about the dangers and how to spot candy which may be adulterated.[85]

---

[82]For example, the town of Hardwick, Massachusetts, banned trick-or-treating several years ago after some children received candies with razor blades in them (Telephone interview with the Chief of Police, Mar. 7, 1983).

[83]*See* Minsky, *A Halloween That Was All Too Scary: Parents Were Wary of Dangerous Tricks*, Boston Globe, Nov. 1, 1982, at 1, col. 4 (less than a dozen tampering incidents in the entire New England area); *see also* DUDLEY-ANDERSON-YUTZY, *supra* note 33.

[84]Paye, *Halloween Ban Called Cause for Anxiety*, Boston Globe, Oct. 27, 1982, at 22, col. 3. The article quoted four psychologists who spoke of feelings of distrust that may be aroused in some children.

[85]For example, the Department of Health and Human Services advises parents to

F) The town could have a metal detector available at a public area such as a hospital or police station to spot any pins and needles hidden in candy.

G) The town could have incentives such as prizes at school for those children who choose not to go trick-or-treating (a phone call to the child's home could confirm whether the child stayed home, if children were required to answer).[86]

IV. CONCLUSION

While it is conceivable that a religious interest is affected by the ban, the lack of any evidence that children consider trick-or-treating a religious activity diminishes the likelihood that an attack on the regulations will be made on this ground. The activities of the children probably rise to the level of protected speech. This protection will attach to the speech whether it is social or commercial.

Although not as constitutionally important as political speech, social speech and assembly do warrant protection. Courts may, however, use a lower standard than a compelling state interest to judge the regulations because social speech has lesser constitutional significance. Yet, the interest of the towns in guarding the safety of their children is so important that this regulation would most likely pass the strictest test. Although the danger of harm may not be as great in every town, the greater deference given to the local authorities when dealing with children and the increased risk due to the recent Tylenol scare seem to be enough to justify the bans.

Banning trick-or-treating raises issues of constitutionality and local power. These issues are similar to those present in the regulation of video games, "head shops," curfews, and pornography. The proliferation of these types of regulations will force the courts to define specifically the limits of local power. The difficult constitutional questions demonstrate the complexity of the task confronting the judicial system.

---

examine unpackaged food for small punctures, to look at the color of food, to smell all candy, and to destroy packaged food that looks as if it has been tampered with. Dep't of Health & Human Servs., Press Release (Oct. 20, 1982) (statement by Edward N. Brandt, Jr., Assistant Secretary for Health) (on file at HARV. J. ON LEGIS.).

[86]*See* Barron, *Poison Worries Lead to Precautions for Halloween*, N.Y. Times, Oct. 28, 1982, at B1, col. 1.

COMMENT

PRE-ENFORCEMENT CONSTITUTIONAL CHALLENGES TO LEGISLATION AFTER *HOFFMAN ESTATES*: LIMITING THE VAGUENESS AND OVERBREADTH DOCTRINES

CHRISTINA L. JADACH*

In *Village of Hoffman Estates v. Flipside, Hoffman Estates, Inc.*,[1] the Supreme Court substantially limited the right to mount pre-enforcement[2] facial[3] challenges to statutes. By upholding the statutory language of an ordinance regulating the sale of drug paraphernalia, the Court reduced the opportunity for businessmen who wish to challenge the constitutionality of economic regulations.

The *Hoffman Estates* decision reviewed a municipal ordinance that required businesses to obtain licenses before selling "items . . . designed or marketed for use with illegal cannabis or drugs" and to maintain lists of purchasers of the drug-related objects.[4] A local vendor of drug paraphernalia, The Flipside, Hoffman Estates, Inc., made a facial challenge to the ordinance on the ground that it violated the First Amendment by imposing a "prior restraint" on speech and by infringing "protected symbolic speech."[5] Flipside also claimed that the ordinance violated the Due Process Clause of the Fourteenth Amendment because its vague drafting denied citizens any certainty as to the scope of restrictions.[6]

---

*B.S., Wharton School of Finance, 1981; member, Class of 1984, Harvard Law School.

[1]455 U.S. 489 (1982).

[2]A pre-enforcement challenge is raised by a party who, although not yet charged under a statute, sues to bar enforcement of that statute.

[3]A facial challenge is a claim that a statute cannot be applied validly in any context. *Hoffman Estates*, 455 U.S. at 494 n.5 (citing Steffel v. Thompson, 415 U.S. 452, 474 (1974)). A federal court must consider any limiting construction applied by state courts and enforcement agencies when evaluating a facial challenge. *Hoffman Estates*, 455 U.S. at 494 n.5 (citing Grayned v. City of Rockford, 408 U.S. 104, 110 (1972)).

[4]455 U.S. at 492. Although the challenge was to a municipal ordinance, the court's language clearly applies to all pre-enforcement challenges of legislation. *See, e.g., id.* at 494–99, 503–05.

[5]*Id.* at 495–96.

[6]*Id.* at 493, 495.

The Supreme Court upheld the ordinance,[7] finding that the overbreadth doctrine does not apply to commercial speech and that infringement of noncommercial speech was not at issue.[8] The Court also held that the ordinance was not unduly vague because a reasonable businessman, such as the petitioner, could understand its meaning.[9] The ordinance was enforceable against the challenging party as to activities clearly banned and thus was not open to facial challenge.[10] Finally, the Court set aside enforcement issues until post-enforcement challenge and refused to hypothesize about possible infringement of the petitioner's constitutional rights in the absence of evidence of such infringement.[11]

The *Hoffman Estates* decision significantly limits the applicability of two constitutional doctrines, overbreadth[12] and vagueness,[13] in cases where businessmen challenge a statute's validity on First and Fourteenth Amendment grounds. Businessmen are assumed to have greater awareness and understanding of legislation than are ordinary citizens.[14] The Court narrowly interpreted the overbreadth and vagueness doctrines by finding no First Amendment protection for the communicative aspect of retail marketing.

This Comment addresses the changes in overbreadth and vagueness doctrine stemming from the decision and examines implications of these changes for lawmakers and for the business community. In Part I, the Comment considers the constitutional and legislative background of the *Hoffman Estates* opinion. It highlights developments in the overbreadth and vagueness doctrines and examines the scope and history of drug paraphernalia legislation. Part II focuses on the treatment of the overbreadth

---

[7] *Id.* at 505. Although the ordinance nominally imposed only civil penalties, the village conceded that it was "quasi-criminal," and the Court tested it with corresponding strictness. *Id.* at 499 n.16. The ordinance has in fact precluded the legal sale of drug paraphernalia in Hoffman Estates; there was at least one vendor of drug paraphernalia, Flipside, in Hoffman Estates before the passage of the ordinance, and there have been no license applications under the ordinance. Telephone interview with Richard N. Williams, Village Attorney, Hoffman Estates (Nov. 2, 1982).

[8] 455 U.S. at 496–97.

[9] *Id.* at 500–01.

[10] *Id.* at 499–500.

[11] *Id.* at 503–04.

[12] *See infra* text accompanying notes 17–30.

[13] *See infra* text accompanying notes 31–32.

[14] *See Hoffman Estates*, 455 U.S. at 498.

and vagueness doctrines within the opinion. Part III discusses several implications of the decision for policymakers, including a heightened standard of review for businessmen facing economic regulations based on a "reasonable businessman" standard, and an exemption from First Amendment protections of commercial marketing and display practices. The conclusion notes that the *Hoffman Estates* opinion will severely limit the applicability of the overbeadth and vagueness doctrines on facial challenge.

I. THE BACKGROUND OF *Hoffman Estates*

Retailers in the drug accessories trade[15] have challenged drug paraphernalia statutes on a variety of constitutional grounds.[16] Serious legal battles have centered on the constitutional doctrines of overbreadth and vagueness. The following sections examine the bases of overbreadth and vagueness challenges to legislative action and explore the development of drug paraphernalia regulations.

---

[15]Sales of drug paraphernalia are reported to be as much as three billion dollars a year. HOUSE SELECT COMM. ON NARCOTICS ABUSE AND CONTROL, 96TH CONG., 2D SESS., DRUG PARAPHERNALIA 3 (Comm. Print 1980) [hereinafter cited as DRUG PARAPHERNALIA REPORT].

[16]*See* Florida Businessmen for Free Enter. v. Florida, 673 F.2d 1213, 1220 (11th Cir. 1982) (Fourteenth Amendment: potential for arbitrary enforcement); Kansas Retail Trade Coop. v. Stephan, 522 F. Supp. 632, 637–38 (D. Kan. 1981) (Fourteenth Amendment: denial of due process—vagueness and lack of warning); Atkins v. Clements, 529 F. Supp. 735, 744 (N.D. Tex. 1981) (Fourteenth Amendment: equal protection); Lady Ann's Oddities, Inc. v. Macy, 519 F. Supp. 1140, 1145 (W.D. Okla. 1981) (First and Fourteenth Amendments: overbreadth); Franza v. Carey, 518 F. Supp. 324, 330 n.10 (S.D.N.Y. 1981) (Bill of Attainder: selective enforcement); New Eng. Trade Accessories Ass'n v. Tierney, 528 F. Supp. 404, 412 (D. Me. 1981) (Double Jeopardy: violating both state and local laws); Tobacco Rd. v. City of Novi, 490 F. Supp. 537, 544–45 (E.D. Mich. 1980) (First Amendment: restriction of free speech); New Eng. Accessories Trade Ass'n v. Browne, 502 F. Supp. 1245, 1254 (D. Conn. 1980) (Fourteenth Amendment: potential for discriminatory treatment), *vacated and remanded*, 679 F.2d 873 (2d Cir. 1981); World Imports Inc. v. Woodbridge Township, 493 F. Supp. 428, 433 (D.N.J. 1980) (Fourteenth Amendment: lack of legitimate state interest); Mid-Atlantic Accessories Trade Ass'n v. Maryland, 500 F. Supp. 834, 848 (D. Md. 1980) (Fifth and Fourteenth Amendments: deprivation of private property); *id.* at 849–50 (First and Ninth Amendments: right to privacy); Delaware Accessories Trade Ass'n v. Gebelein, 497 F. Supp. 289, 296 (D. Del. 1980) (Fourth and Fourteenth Amendments: illegal search and seizure); Nova Records v. Sendak, 504 F. Supp. 938, 943 (S.D. Ind. 1980) (Eighth Amendment: cruel and unusual punishment); Record Revolution No. 6, Inc. v. City of Parma, 492 F. Supp. 1157, 1164 n.2 (N.D. Ohio) (Supremacy Clause), *aff'd*, 638 F.2d 916 (6th Cir. 1980), *vacated and remanded*, 451 U.S. 1013 (1981); Bambu Sales v. Gibson, 474 F. Supp. 1297, 1299 (D.N.J. 1979) (Commerce Clause: restriction on interstate commerce); *id.* (Contract Clause: interference with contractual obligations).

A. *The Overbreadth Challenge*

The overbreadth doctrine gives unharmed litigants standing to challenge statutes that unnecessarily restrict constitutional rights.[17] The Supreme Court developed this doctrine in two opinions, *Thornhill v. Alabama*[18] and *Broadrick v. Oklahoma*.[19]

Thornhill involved a challenge by union picketers against an Alabama statute banning picketing of businesses. The Court conducted a two-step analysis in determining whether the anti-picketing statute was unconstitutionally overbroad. First, the Court found that picketing constitutes "speech" protected by the First Amendment. Such a preliminary finding was critical because a statute cannot be overbroad with respect to activities not protected by the Constitution. Second, the Court considered whether the Alabama statute unduly infringed the protected rights of speech and association. The Court defined an overly broad statute as one "which does not aim specifically at evils within the allowable area of State control but, on the contrary, sweeps within its ambit other activities that in ordinary circumstances constitute an exercise of freedom of speech or of the press."[20] The Court held that abridgement of constitutional freedoms could be justified where a clear and substantial danger would otherwise result. Holding that picketing did not pose such a danger,[21] the Court struck down the statute as void on its face.

Generally courts evaluate a statute by considering whether the provision impairs the rights of the complaining party in light of the attending circumstances. This traditional standing rule prohibits petitioners from invoking rights of third parties in individual claims. The *Thornhill* court created an important exception to this rule by holding that where the petitioner claims that a law limits constitutionally protected First Amendment freedoms, the court should apply a broad third-party standing rule and evaluate the statute on its face, as applied to all persons, and not limit itself to the facts of the case.[22] Although the petitioner may not have shown that his own rights were violated, this third-party standing rule allows him to challenge the regu-

---

[17]For a general discussion of the overbreadth doctrine and its application in specific cases, see L. TRIBE, AMERICAN CONSTITUTIONAL LAW §§ 12-24 to -26 (1978).

[18]310 U.S. 88 (1940).

[19]413 U.S. 601 (1973).

[20]310 U.S. at 97.

[21]*Id.* at 104–05.

[22]*Id.* at 98.

latory statutes on behalf of all those who may be unjustly harmed. The Supreme Court found this broader rule necessary to prevent statutory restrictions on protected speech or expression.[23]

In *Broadrick v. Oklahoma*,[24] the Court limited its analysis in *Thornhill* by distinguishing between statutes regulating pure speech and those regulating conduct. The challenged statute prohibited civil service personnel from participating in political activities, parties, or campaigns, beyond voting in public elections.[25]

The Court concluded that because an overbreadth challenge to a statute is such a significant departure from traditional review procedures, the doctrine should be restrictively interpreted and applied.[26] Specifically, "where conduct and not merely speech is involved . . . the overbreadth of a statute must not only be real, but substantial as well, judged in relation to the statute's plainly legitimate sweep."[27]

The effect of *Broadrick* has been to distinguish between the overbreadth analysis applied to cases challenging restrictions on "speech" and the analysis applied to cases challenging restrictions on "conduct." In "speech" cases, the Court will apply the broader standing rule described in *Thornhill*, and will allow petitioners to assert that third parties' constitutionally protected activity is inhibited by an overly broad statute.[28] This reflects the belief that speech limitations represent direct censorship of ideas and pose a clear and grave threat to First Amendment rights.[29] But in "conduct" cases, where the belief is that statutory restrictions pose a more peripheral challenge to First Amendment freedoms, the petitioner is required to show that the overbreadth is "real" and "substantial" in relation to the legitimate scope of the provision before being permitted to invoke the broader standing rule.[30]

Overbreadth challenges to statutes thus involve three major

---

[23]*Id.* at 97.

[24]413 U.S. 601 (1973).

[25]For the full text of the statute, see *id.* at 603 n.1.

[26]*Id.* at 613–15. An overbroad statute is remedied by narrowing the scope of the statute to eliminate the law's encroachment upon constitutionally protected activity, or by striking down the law completely where such a saving construction cannot be fashioned. *Id.* at 613.

[27]*Id.* at 615.

[28]*Id.* at 612–13.

[29]*Id.*

[30]*Id.* at 615–16.

issues. First, the statute must infringe speech or conduct that is constitutionally protected. Second, if the statute restricts protected speech, a challenger will have standing to assert that third parties' rights are infringed by the law. Finally, if the statute restricts protected conduct, a challenger must show that the law substantially restricts protected rights, relative to the legitimate scope of the regulation.

B. *The Vagueness Challenge*

A statute also may be challenged on its face on the ground that it is unduly vague. The Supreme Court detailed a two-prong test for vagueness in *Grayned v. City of Rockford*.[31] The Court explained that

> vague laws offend several important values. First, because we assume that man is free to steer between lawful and unlawful conduct, we insist that laws give the person of ordinary intelligence a reasonable opportunity to know what is prohibited, so that he may act accordingly. Vague laws trap the innocent by not providing fair warning. Second, if arbitrary and discriminatory enforcement is to be prevented, laws must provide explicit standards for those who apply them. A vague law impermissibly delegates basic policy matters to policemen, judges, and juries for resolution on an *ad hoc* and subjective basis, with the attendant dangers of arbitrary and discriminatory application.[32]

The first test for vagueness is whether ordinary citizens can understand the intended scope of the regulation. It is essential that the regulation give "fair notice to those to whom [it] is directed."[33] If the law does not define clearly the extent of prohibited activity, citizens' due process rights may be infringed; individuals can be charged with violating a law with which they believe they had complied.

A statute also can be held vague if it does not contain detailed standards to prevent arbitrary or discriminatory enforcement by those who are to apply it.[34] Due process requires fair and equal application of laws to all citizens. Where a statute does not

---

[31]408 U.S. 104 (1972).

[32]*Id.* at 108.

[33]*Id.* at 112 (citing American Communications Ass'n v. Douds, 339 U.S. 382, 412 (1950)).

[34]*Id.* at 109.

sufficiently guide officials in determining the precise nature of prohibited conduct, discretionary enforcement power is left in the hands of the police.[35] Such a statute is vague because it does not ensure all citizens equal treatment under the law. This test considers whether the statutory definition of the prohibited activity sufficiently guides officials charged with carrying out the provisions.

C. *The Development of Drug Paraphernalia Laws*

Lawmakers have had difficulty defining the term "drug paraphernalia,"[36] because items used as drug accessories have lawful as well as unlawful uses.[37] Most laws restricting the sale of drug paraphernalia were challenged soon after passage,[38] and many of the early provisions were held unconstitutional by federal courts.[39] In May 1979, the White House requested that the Justice Department draft legislation to aid states and communities in their regulatory efforts.[40] The Drug Enforcement Administration, a division of the Justice Department, responded with the

---

[35]*Id.*

[36]Drug paraphernalia is a loose term describing objects used in growing, harvesting, manufacturing, refining, preparing, and ingesting illicit drugs.

[37]For example, glass water pipes can be used as tobacco smoking accessories, and "roach clips," items designed to hold burning marijuana cigarettes or other items that are too small to be held by hand, can be used as electrical wiring implements. World Imports, Inc. v. Woodbridge Township, 493 F. Supp. 428, 430 (D.N.J. 1980). Tobacco smokers may prefer colored rolling paper, just as marijuana users often do. Courts, however, have taken judicial notice that, in some cases, illegitimate uses dominate legitimate uses. For example, one judge noted that only the "naive" could believe that syringes and eyedroppers sold by vendors of drug paraphernalia were used to feed small injured birds. *Id.* at 431.

[38]*See supra* note 16 and accompanying text.

[39]Federal courts overwhelmingly rejected early attempts at legislation. Only two regulations received court approval. Gasser v. Morgan, 498 F. Supp. 1154 (N.D. Ala. 1980); Tobacco Rd. v. City of Novi, 490 F. Supp. 537 (E.D. Mich. 1980). Nine courts struck down drug paraphernalia regulations. High Ol' Times, Inc. v. Busbee, 515 F. Supp. 176 (N.D. Ga. 1980), *rev'd*, 673 F.2d 1225 (11th Cir. 1982); Knoedler v. Roxbury Township, 485 F. Supp. 990 (D.N.J. 1980); Magnani v. City of Ames, 493 F. Supp. 1003 (S.D. Iowa 1980); Music Stop, Inc. v. City of Ferndale, 488 F. Supp. 390 (E.D. Mich. 1980); Indiana Chapter, Nat'l Org. for the Reform of Marijuana Laws v. Sendak, No. TH-75-142-C (S.D. Ind. Feb. 4, 1980) (available on LEXIS, Genfed library, Dist file), *vacated*, 631 F.2d 734 (7th Cir. 1980); Bambu Sales v. Gibson, 474 F. Supp. 1297 (D.N.J. 1979); Record Museum v. Lawrence Township, 481 F. Supp. 768 (D.N.J. 1979); Housworth v. Glisson, 485 F. Supp. 29 (N.D. Ga. 1978), *aff'd*, 614 F.2d 1295 (5th Cir. 1980); Riddle v. Clack, No. CA-3-77-0525-D (N.D. Tex. Aug. 25, 1977) (available on LEXIS, Genfed library, Dist file).

[40]*Drug Paraphernalia: Hearings Before the House Select Comm. on Narcotics Abuse and Control*, 96th Cong., 1st Sess. 29–31 (1979) (testimony of Lee I. Dogoloff, Assoc. Director for Drug Policy, Domestic Policy Staff, the White House).

Model Drug Paraphernalia Act (MDPA),[41] an amendment to the widely enacted Uniform Controlled Substances Act.[42]

Despite widespread adoption of the carefully drafted MDPA, drug paraphernalia regulations continued to face judicial scrutiny. Challengers urged that the drug paraphernalia statutes were overbroad in that they banned advertising,[43] overly restricted items with innocent uses,[44] and infringed commercial speech by prohibiting display of merchandise.[45] Challengers also challenged the statutes on vagueness grounds.[46] They claimed that popular statutory definitions of proscribed items, such as objects "intended for use," "designed for use," "marketed for use," "used with," and "adapted for use with" illegal cannabis or drugs, are vague since most "drug related" items have non-drug related uses.[47] Challengers also charged that the statutes did not contain sufficient enforcement guidelines and thus permitted police officers to enforce the law discriminatorily and to harass individuals with particular lifestyles.[48]

The critical issue underlying this judicial morass was the stringency with which vagueness is to be tested. Generally, courts that struck down drug paraphernalia legislation applied a stricter standard of review, and a looser standard of standing, than did courts that upheld regulatory statutes. In the former cases, judges considered a broad range of applications of the statutes

---

[41]The text of the Act, with the Drug Enforcement Administration's comments, is reprinted in DRUG PARAPHERNALIA REPORT, *supra* note 15, at 29–38 app. E.

[42]UNIF. CONTROLLED SUBSTANCES ACT, 9 U.L.A. 187 (1970); *see* DRUG PARAPHERNALIA REPORT, *supra* note 15, at 29 app. E (Prefatory Note to MDPA).

The MDPA is divided into four articles. Article I defines drug paraphernalia as "items used, intended for use, or designed for use" with illegal drugs. MDPA art. I, *reprinted in* DRUG PARAPHERNALIA REPORT, *supra* note 15, at 30–31 app. E. The model statute lists fourteen "factors" that enforcement officials "should" consider when determining whether an item falls within the statute's proscribed limits. These factors include the proximity of the object to controlled substances, instructions provided with the object concerning its use, advertising concerning use, and the manner in which the object is displayed for sale. *Id.* at 31 app. E. Article II outlaws possession, manufacture or delivery, delivery to a minor, and advertising to promote sale, of objects designed for use as drug paraphernalia. *Id.* art. II, at 32 app. E. Article III authorizes civil forfeiture of drug accessories, *id.* art. III, at 32 app. E, and Article IV contains a severability provision to save remaining portions of the statute if some sections are held invalid, *id.* art. IV, at 33 app. E.

[43]*E.g.*, Kansas Retail Trade Coop. v. Stephan, 522 F. Supp. 632, 642 (D. Kan. 1981).

[44]*E.g.*, Gasser v. Morgan, 498 F. Supp. 1154, 1167 (N.D. Ala. 1980).

[45]*E.g.*, Tobacco Rd. v. City of Novi, 490 F. Supp. 537, 544–45 (E.D. Mich. 1980).

[46]*E.g.*, Information Management Servs. v. Borough of Pleasant Hills, 512 F. Supp. 1066 (W.D. Pa. 1981).

[47]*See supra* note 37 and accompanying text.

[48]*E.g.*, Flipside, Hoffman Estates, Inc. v. Village of Hoffman Estates, 639 F.2d 373, 384 (7th Cir. 1981), *rev'd*, 455 U.S. 489 (1982).

and found some that disqualified the provision.[49] In the latter cases, courts confined the inquiry to the facts of the case, and upheld laws with potential ambiguities if they in some way applied to the complaining party.[50]

II. THE *Hoffman Estates* DECISION

In an opinion by Justice Marshall in *Hoffman Estates*,[51] the Supreme Court answered many of the questions raised by drug paraphernalia statutes. Although drug paraphernalia statutes are the narrow subject of *Hoffman Estates*, the case is primarily a procedural ruling with applications to all statutes, particularly those regulating businesses. The decision carefully outlines the procedures and review standards for overbreadth and vagueness challenges. The Court emphasized the narrow scope of the vagueness review: a statute that applies to the challenger's conduct is not vague. Although a challenger can assert third parties' rights in an overbreadth challenge to a statute that restricts First Amendment rights, the broad standing rule is not extended to the vagueness analysis. The Court also ruled that consideration of discriminatory enforcement issues should be postponed, absent a showing of actual discrimination, until post-enforcement challenge.

A. *Overbreadth*

The Court held that neither regulation of marketing nor regulation of display of drug-related literature infringes *noncommercial* speech. Such provisions do not ban the sale of literature, but merely restrict the *manner* of such sale.[52] Challengers of the

---

[49]*See, e.g.*, Record Revolution No. 6 v. City of Parma, 638 F.2d 916, 931 (6th Cir. 1980), *vacated and remanded*, 451 U.S. 1013 (1981); Information Management Servs. v. Borough of Pleasant Hills, 512 F. Supp. 1066, 1072–74 (W.D. Pa. 1981); Back Door Records v. City of Jacksonville, 515 F. Supp. 857, 858–59 (E.D. Ark. 1981); Knoedler v. Roxbury Township, 485 F. Supp. 990, 992–93 (D.N.J. 1980); Record Museum v. Lawrence Township, 481 F. Supp. 768, 774 (D.N.J. 1979).

[50]*See, e.g.*, New Eng. Accessories Trade Ass'n v. City of Nashua, 679 F.2d 1, 7 (1st Cir. 1982); Levas & Levas v. Village of Antioch, 684 F.2d 446 (7th Cir. 1982); Brache v. County of Westchester, 658 F.2d 47, 51 (2d Cir. 1981), *cert. denied*, 455 U.S. 1005 (1982); Tobacco Rd. v. City of Novi, 490 F. Supp. 537, 546 (E.D. Mich. 1980).

[51]Village of Hoffman Estates v. Flipside, Hoffman Estates, Inc., 455 U.S. 489 (1982). Justice White filed a concurring opinion; Justice Stevens took no part in the consideration or decision of the case.

[52]455 U.S. at 496.

Hoffman Estates ordinance therefore had to show specifically how their own commercial speech was infringed.

Holding that overbreadth principles do not extend to commercial speech,[53] the Court did not grant the petitioner broad standing rights to raise possible violations of the commercial speech rights of third parties.[54] The Court found that the only commercial speech involved was the display and marketing of items in the manner preferred by the retailer.[55] The *Hoffman Estates* ordinance proscribed the display of drug paraphernalia in proximity to literature encouraging drug use.[56] The Supreme Court held that such a restriction does not "appreciably" limit speech,[57] except possibly speech promoting illegal drug use. Because government can ban outright any activity or speech proposing illegal transactions, a community can regulate such speech by imposing marketing restrictions.[58]

B. *Vagueness*

The Court held that a facial challenge to a statute may be made only if the statute is vague in all of its applications.[59] Businessmen face a heightened vagueness standard. An economic regulation is subject to a less strict standard of review because of the limited scope of such regulation and businessmen's heightened level of knowledge about relevant legislation.[60]

The Supreme Court interpreted the Hoffman Estates ordi-

---

[53]*Id.* at 497.

[54]*Id.* at 496–97.

[55]*Id.* at 496. For example, the retailer displayed the magazine *High Times* and books such as *The Marijuana Growers Guide* close to pipes and colored rolling papers. *Id.* at 502.

[56]*Id.* at 492–93 n.3.

[57]Because the *Hoffman Estates* ordinance involved a licensing provision, the Court did not consider the constitutionality of statutes that ban totally the sale and advertisement of drug paraphernalia. The MDPA and statutes patterned after that Act form a significant body of legislation imposing outright bans on the sale of drug accessories. It is clear from the Supreme Court's discussion of the overbreadth doctrine that a plaintiff challenging a commercial regulation will not be accorded the right to assert third parties' claims against the scope of the statute. Nevertheless, a plaintiff can assert third parties' rights when challenging statutory bans on noncommercial speech. Such a claim could arise in a First Amendment challenge to a law banning the sale of literature containing drug paraphernalia advertisements.

[58]455 U.S. at 496.

[59]*Id.* at 497.

[60]*Id.* at 498.

nance holding "designed for use" to indicate the manufacturer's intended use for an item, as evidenced by the item's objective features.[61] The Supreme Court's validation of the "designed for use" standard ended a long-standing dispute among lower courts.[62] The Court employed the definitions urged by lawmakers: a manufacturer "designs" an item for use, a seller "markets" it, and a buyer "uses" it.[63] The Court also found the "marketed for use" standard sufficient to withstand the vagueness challenge. The Court applied this standard to retailers, whose due process rights were deemed protected because "marketing" requires intentional display with the purpose of promoting a particular type of sale and because the ordinance sufficiently alerted businessmen of ordinary intelligence of the consequences of their conduct.[64]

C. *Post-enforcement Challenges*

The *Hoffman Estates* opinion included an admonition to courts that regulatory standards are legislative matters within the discretion of community leaders.[65] The judge's view of the effectiveness of the legislation is to play no part in the decision.[66]

The overall effect of the *Hoffman Estates* decision is to strengthen the presumptive validity of economic regulations by limiting pre-enforcement overbreadth and vagueness challenges. The closing paragraphs of the opinion, however, warn law enforcement agencies that further challenges to drug paraphernalia statutes are likely to be heard in the absence of precautions against arbitrary enforcement.[67] In the drug paraphernalia setting, the battle now shifts to post-enforcement challenge.

---

[61]*Id.* at 500–02.

[62]The Court's saving interpretation of the standard was urged in the village's brief, Appellant's Brief at 9–14, and in an amicus brief submitted by the attorneys general of 21 states, Brief of Ark., Colo., Conn., Del., Fla., Idaho, Ind., Kan., La., Me., Md., Neb., Nev., N.J., N.M., N.C., Okla., Pa., Tex., Utah, and Wash. as Amici Curiae at 9–12. The attorneys general urged that the "designed for use" threshold was unambiguous. They noted that "design" includes objective characteristics and an intentional use. Brief of Amici Curiae, *supra*, at 9–10 (quoting Delaware Accessories Trade Ass'n v. Gebelein, 497 F. Supp. 289, 297 (D. Del. 1980)).

[63]455 U.S. at 501–02.

[64]*Id.* at 502–03.

[65]*Id.* at 504–05.

[66]*Id.* at 505.

[67]*Id.* at 503–04.

III. IMPLICATIONS FOR POLICYMAKERS

The *Hoffman Estates* decision significantly affects regulations far outside the drug paraphernalia arena. On overbreadth and vagueness challenges, the ruling effectively insulates statutory regulations from judicial review by narrowing the scope of inquiry on both challenges. The effect of the decision is threefold: first, it establishes specific review standards for businessmen facing economic regulations; second, it exempts commercial marketing and display practices from First Amendment overbreadth protection; and third, it holds that for a facial challenge to succeed on vagueness grounds, the challenger must show that there are *no* valid applications of the statute. The vagueness holding, perhaps the most important one in the opinion, has broad application to all statutes. These limitations on pre-enforcement challenges strengthen the presumptive validity of economic legislation and postpone constitutional challenges until the post-enforcement stage.

Hoffman Estates establishes a high standard of legal awareness and compliance for the business community. Economic regulations are no longer to be judged by the "reasonable person" standard applicable to general legislation.[68] Courts will assume that businessmen plan their activities carefully, are aware of legal developments, and organize their activities in light of such knowledge. Thus, when a businessman challenges an economic regulation as vague, the court will examine the statute by considering whether the "reasonable businessman," not the "reasonable man," can understand the statute.

By varying the test for vagueness with the nature of the enactment, the Supreme Court effectively prevents businessmen from feigning ignorance of uses by hiding behind the broad and general "reasonable person" standard. This problem was particularly evident in the drug paraphernalia business, where numerous challengers, who were proprietors of head shops, feigned ignorance of the true uses for the items they sold.[69] *Hoffman Estates* also clarifies the distinction between commercial and noncommercial speech. This distinction is the critical and controlling issue in an overbreadth challenge to a regulation, because a petitioner who convinces the court that First Amend-

---

[68] *Id.* at 498.

[69] *See supra* note 37.

ment freedoms are at issue can invoke the broad standing rule reserved for such cases.

Any business regulation, such as a zoning or licensing provision, restricts businessmen's freedom to conduct their businesses as they desire. The overbreadth doctrine applies, however, only to those freedoms protected by the Constitution, typically the First Amendment guarantee of free speech and communication of ideas. Two aspects of business activity are essentially communicative: advertising, and marketing and display of merchandise. Significantly, the Supreme Court held marketing and display to be outside the protected area. The Court found that manner of display involved only "attenuated" commercial speech interests, not communicative issues.[70] This seems to imply that any regulation short of a ban on display cannot reach protected speech, because it burdens only the manner, not the existence, of a sale. In its effort to uphold community regulations of commercial activity, the Court stripped a significant volume of business communication from First Amendment protection.[71]

It is not clear how far a community may extend regulations before infringing protected conduct. For example, the MDPA totally prohibits the possession, manufacture, delivery, or advertising of drug paraphernalia items.[72] After *Hoffman Estates*, a community apparently may ban "advertising," an inherently communicative act, if the advertising proposes an illegal transaction. But the question remains whether a community will be able to ban the sale of an entire magazine that contains a single drug paraphernalia advertisement.

IV. CONCLUSION

The Supreme Court's *Hoffman Estates* ruling has broad implications for all statutes. The Court responded to the frustration of communities over their inability to control business activity within their boundaries. As far as facial overbreadth and

---

[70]455 U.S. at 496.

[71]The Court did note that even *were* the display of merchandise considered speech, it could be regulated because it encouraged an illegal activity. *Id.* The damage to First Amendment expression, however, was wrought in the initial finding that marketing and display of merchandise was not a communicative act.

[72]MDPA art. II, *reprinted in* DRUG PARAPHERNALIA REPORT, *supra* note 15, at 32 app. E.

vagueness challenges are concerned, legislative enactments have been elevated to near-absolute constitutional validity. While this decision should not be viewed as an absolute license for legislatures to restrict commercial conduct, it does alleviate the trouble legislatures sometimes have had in writing constitutional laws that would be upheld by the courts. Legislators need be less concerned with ambiguities inherent in drafting regulations and more concerned with developing effective, nondiscriminatory enforcement procedures.

The Court in *Hoffman Estates* outlines an extremely narrow vagueness construction, introduces differentiated vagueness standards for business and non-business plaintiffs, and gives commercial speech a restrictive definition. With the Court's renewed emphasis on post-enforcement challenge, individuals' rights to challenge the validity of legislation *before* prosecution has given way to a strengthened legislative hand. Challengers may have to await prosecution before raising their concerns about an unconstitutional economic regulation.

COMMENT

BANNING "ACTUARIALLY SOUND" DISCRIMINATION: THE PROPOSED NONDISCRIMINATION IN INSURANCE ACT

STEPHEN R. KAUFMAN*

During the last two decades, as Congress systematically outlawed discrimination in areas such as employment, housing, and credit,[1] the field of insurance remained untouched. Now civil rights policy may be extended to the insurance industry through the Nondiscrimination in Insurance Act,[2] a bill currently pending before Congress.[3] The bill prohibits all insurers from discriminating on the basis of race, color, religion, sex, or national origin in the terms, conditions, rates, benefits, or requirements of any insurance, annuity, or pension contract.[4] Specifically, the direct or indirect use of any actuarial table based on a forbidden criterion would be prohibited as a means of distinguishing among insureds.[5] The bill creates a private cause of action to enforce its provisions.[6] However, this method of enforcement may be utilized only after state remedies have been exhausted.[7]

Since ethnicity is, practically speaking, no longer used as an explicit risk classification factor by the insurance industry,[8] sex

---

*B.A., Allegheny College, 1981; member, Class of 1984, Harvard Law School.

[1]Title VII of the Civil Rights Act of 1964, 42 U.S.C. §§ 2000e–2000e-17 (1976 & Supp. V 1981) (employment discrimination); Title VIII of the Civil Rights Act of 1968, 42 U.S.C. §§ 3601–3619 (1976 & Supp. V 1981) (discrimination in housing); Equal Credit Opportunity Act, 15 U.S.C §§ 1691–1691f (1976 & Supp. V 1981) (discrimination in consumer credit and finance).

[2]H.R. 100, 98th Cong., 1st Sess. (1983) (introduced by Rep. John D. Dingell (D-Mich.), Jan. 3, 1983); *see also* S. 372, 98th Cong., 1st Sess. (1983) (introduced by Sen. Mark O. Hatfield (R-Or.), Feb. 1, 1983). The Senate bill is entitled the Fair Insurance Practices Act.

[3]129 CONG. REC. S829 (daily ed. Feb. 1, 1983) (statement of Sen. Packwood).

[4]H.R. 100, § 2(a), (b).

[5]*Id.* § 4(b)(1), 4(c)(1).

[6]*Id.* § 6. The United States Attorney General also may bring a civil action under § 8.

[7]*Id.* § 5.

[8]*Nondiscrimination in Insurance Act of 1981: Hearings on H.R. 100 Before the Subcomm. on Commerce, Transportation, and Tourism of the House Comm. on Energy and Commerce*, 97th Cong., 1st Sess. 172 (1981) (statement of James J. Sheridan, Comm'r of Nat'l Ass'n of Indep. Insurers) [hereinafter cited as *1981 House Hearings*]; *id.* at 195 (statement of Andre Maisonpierre, Alliance of Am. Insurers); Bailey, Hutchison & Narber, *The Regulatory Challenge to Life Insurance Classification*, 25 DRAKE L. REV. 779, 793 & n.54 (1976). The Senate Report on S. 2204, 97th Cong., 2d Sess. (1982), the predecessor to the current Senate bill, stated at its beginning that discrimination on the basis of race does exist in the insurance field but then devoted the rest of

discrimination is the primary focus of the bill. Although discrimination against women in the availability, terms, and coverage of insurance policies is a serious problem,[9] this Comment addresses the most controversial aspect of the bill, its prohibition on the use of sex as a factor in the calculation of rates and benefits.

Women currently pay higher rates for health and disability insurance but lower rates for life insurance.[10] In annuities, they either pay higher rates or receive lower benefits.[11] Young men are generally charged higher premiums for automobile insurance.[12] The pending bill mandates the use of merged-gender tables to calculate rates and benefits. While rates may be increased or decreased to comply with the bill, benefits under existing contracts may not be decreased.[13]

Current differences in rates and benefits in life insurance and annuities stem from the use of sex-based tables, which show that women, on the average, live longer than men.[14] While the validity of these tables has been recently questioned,[15] the weight of authority still supports their accuracy.[16] In any case,

---

the report entirely to the issue of sex discrimination. S. REP. No. 671, 97th Cong., 2d Sess. (1982).

[9]*See 1981 House Hearings*, *supra* note 8, at 40–41 (statement of Catherine East, Nat'l Org. of Women); *id.* at 57 (statement of Mary W. Gray, Women's Equity Action League); *id.* at 169 (statement of Judy Schub, Nat'l Fed'n of Business & Professional Women Clubs); S. REP. No. 671, 97th Cong., 2d Sess. 2–3 (1982). These forms of discrimination could be corrected by a far less sweeping bill.

[10]*1981 House Hearings*, *supra* note 8, at 168–69 (statement of Judy Schub, Nat'l Fed'n of Business & Professional Women Clubs); *Nondiscrimination in Insurance Act: Hearings on S. 2477 Before the Subcomm. on Antitrust, Monopoly and Business Rights of the Senate Comm. on the Judiciary*, 96th Cong., 2d Sess. 12–13 (1980) (statement of Dr. Quincalee Brown, Am. Ass'n of Univ. Women) [hereinafter cited as *1980 Senate Hearings*].

[11]*1980 Senate Hearings*, *supra* note 10, at 12–13 (statement of Dr. Quincalee Brown, Am. Ass'n of Univ. Women).

[12]*1981 House Hearings*, *supra* note 8, at 179–80 (statement of Diana Lee, Nat'l Ass'n of Indep. Insurers). Several states have already banned sex discrimination in automobile insurance rates. *See, e.g.*, MASS. GEN. LAWS ANN. ch. 175, § 113B, ch. 175E, § 4 (West Supp. 1982).

[13]H.R. 100, § 4(c)(2). This section also subjects rate modifications necessary for compliance with the bill to approval by the relevant state agency regulating insurance.

[14]*See* City of Los Angeles, Dep't of Water & Power v. Manhart, 435 U.S. 702, 704 (1978).

[15]*See* Brilmayer, Hekeler, Laycock & Sullivan, *Sex Discrimination in Employer-Sponsored Insurance Plans: A Legal and Demographic Analysis*, 47 U. CHI. L. REV. 505, 559 (1980) (sex is a "spurious, weak and unstable predictor of mortality") [hereinafter cited as Brilmayer].

[16]*See 1981 House Hearings*, *supra* note 8, at 72 (statement of Barbara J. Lautzenheiser, Am. Council of Life Ins.); S. HUEBNER & K. BLACK, LIFE INSURANCE 303 (10th ed. 1982); Benston, *The Economics of Gender Discrimination in Employee Fringe Benefits:* Manhart *Revisited*, 49 U. CHI. L. REV. 489 (1982); Gold, *Of Giving and Taking:*

proponents of the ban against sex discrimination in insurance do not rest their opinion on alleged actuarial invalidity; rather, they believe that the use of sex-based tables is wrong as a matter of "simple justice."[17] For purposes of this Comment, the statistical accuracy of such tables is assumed, and the bill is examined solely on the merits of eliminating sex discrimination as a matter of social policy and economic efficiency.

The bill would expand the idea of discrimination in insurance far beyond the definitions of discrimination found in current state statutes. The McCarran-Ferguson Act[18] grants to the states primary responsibility for regulating the business of insurance. Scholars who have recently surveyed state laws found no state that prohibits the use of sex as a factor in determining rates for individual life or health insurance or for individual annuity policies.[19] States do prohibit "unfair discrimination" in rate-setting under their Unfair Trade Practices Acts.[20] However, the use of sex as a factor has not been regarded as "unfair discrimination," since the validity of sex-based actuarial tables has been questioned only recently.[21] One commentator characterizes the state cases challenging the use of sex as a classification factor in insurance as "rather meager" and largely unsuccessful.[22]

Hearings were held on the Nondiscrimination in Insurance Act in both the Ninety-sixth and Ninety-seventh Congresses.[23]

Applications and Implications of City of Los Angeles, Department of Water & Power v. Manhart, 65 VA. L. REV. 663, 699 (1979); *see also* Kimball, *Reverse Sex Discrimination:* Manhart, 1979 AM. B. FOUND. RESEARCH J. 85, 108–09 (discussing how mortality tables are developed).

[17]S. REP. No. 671, 97th Cong., 2d Sess. 4 n.2 (1982).

[18]15 U.S.C. §§ 1011–1015 (1976).

[19]Bailey, Hutchison & Narber, *supra* note 8, at 800; Key, *Sex-Based Pension Plans in Perspective:* City of Los Angeles, Department of Water and Power v. Manhart, 2 HARV. WOMEN'S L.J. 1, 39–40 (1979).

[20]Bailey, Hutchison & Narber, *supra* note 8, at 782; *see, e.g.*, N.J. STAT. ANN. § 17B:30-12(c), (d) (West Supp. 1982).

[21]*See supra* note 15 and accompanying text.

[22]Key, *supra* note 19, at 40. She provides a good summary of these cases, finding that most of the suits have attacked availability and term differences based on sex rather than challenging rate differentials. *Id.* at 41–43. Finding state action is the primary problem, according to Key, since most of the actions are based on the Equal Protection Clause, U.S. CONST. amend. XIV, § 1, the Civil Rights Act of 1871, 42 U.S.C. § 1983 (1976), and state Equal Rights Amendments. Key indicates that suits based on the Civil Rights Act of 1871, 42 U.S.C. § 1985(3), which prohibits private conspiracies, have not met with any greater success. Key, *supra* note 19, at 40, 43 & n.230.

[23]*1981 Housing Hearings*, *supra* note 8; *1980 Senate Hearings*, *supra* note 10; *Nondiscrimination in Insurance: Hearings on H.R. 100 Before the Subcomm. on Consumer Protection and Finance of the House Interstate and Foreign Commerce Comm.*, 96th Cong., 2d Sess. (1980). Hearings held by the Senate Committee on Commerce, Science and Transportation on July 15, 1982, have not yet been published.

Both houses are expected to act expeditiously on the bill this year.[24] The likelihood of congressional action during this Congress makes scrutiny of the bill's implications important. Section I of this Comment examines the argument that sex discrimination in insurance should be outlawed as a matter of social policy. Section II considers possible economic effects of the bill, both short-term and long-term. Analysis of both sides of the controversy will enable a reasoned judgment on the merits of the bill.

I. CONSIDERATIONS OF SOCIAL POLICY

The question presented by the bill is whether discrimination that is not necessarily arbitrary or irrational, but that is based on "objective and accepted" group distinctions,[25] may still be harmful enough to justify elimination. Sex discrimination in insurance, in the form of charging men and women different rates and offering them different benefits, does not violate one's sense of justice as immediately as do other forms of sex discrimination, such as compensating a woman less than a man for equal work. The distinction probably lies in the nature of the discrimination. Whereas the latter form of discrimination may stem from an evil intent or from reliance on arbitrary and unjustified stereotypes, actuarially-based discrimination appears to be firmly grounded in objective reality. Thus, the premise of the bill is that all race and sex discrimination in insurance is wrong, whether or not it is invidious.

Essentially, this debate involves two distinct concepts of fairness. On the one hand, supporters of the bill believe that grouping by sex for insurance purposes is repugnant to concepts of individualism and fairness, as well as objectionable as an original

---

[24]Hearings were held by the House Committee on Energy and Commerce on February 22 & 24, 1983. Senator Packwood said when the bill was introduced,

> Last year's bill, S. 2204, was favorably reported from the Senate Commerce Committee which I chair. As chairman of that committee I intend to move expeditiously to hold hearings on the bill and to again report the bill. I believe that the bill will be considered by the full Senate this year and I predict it will become law.

129 CONG. REC. S829 (daily ed. Feb. 1, 1983) (statement of Sen. Packwood). In a report on the CBS Evening News of February 22, 1983, correspondent Leslie Stahl stated that the bill was "expected to pass" and that President Reagan had not yet taken a position.

[25]Justice Blackmun, concurring in City of Los Angeles, Dep't of Water & Power v. Manhart, 435 U.S. 702, 724 (1978) (citations omitted), observed: "I might have thought, too, that there is nothing arbitrary, irrational, or 'discriminatory' about recognizing the objective and accepted disparity in female-male life expectancies in computing rates for retirement plans."

matter. On the other hand, the insurance industry contends that the individuals are currently being treated equitably.

A Supreme Court decision five years ago lends strong support to the arguments of those in favor of the bill. In *City of Los Angeles, Department of Water & Power v. Manhart*, the Court held that requiring men and women to make unequal contributions to an employer-operated pension fund violated Title VII of the Civil Rights Act of 1964.[26] The Court, emphasizing the importance of treating the insured as an individual rather than as a member of a class, explicitly prohibited an employer from using sex as a predictor of individual risk.[27] The Second and Ninth Circuits have found the individual rights arguments made in *Manhart* compelling enough to extend its holding to pension plans operated by private insurers through the employer.[28] Moreover, these courts interpreted *Manhart* as mandating equal pension benefits as well as equal payments.[29] The Sixth Circuit, however, has refused to extend *Manhart* along either avenue, thus reaching a conflicting result.[30]

Although the strong individual rights arguments presented in *Manhart* have been followed by the Second and Ninth Circuits, these cases were decided under Title VII of the Civil Rights Act of 1964,[31] which deals only with employment. The Nondiscrimination in Insurance Act would extend the ban on sex discrimination far beyond its current boundaries to the entire insurance market. Nevertheless, sex discrimination by private insurers is as offensive to concepts of individual equality as is sex discrimination by employers.

The decision to group by sex results in the further deleterious consequences of stigmatization and fragmentation of society. One ever-present danger in attaching discriminatory group labels to individuals is the formation or perpetuation of group stigmas and stereotypes. The *Manhart* majority recognized the

---

[26]435 U.S. 702, 717 (1978).

[27]*Id.* at 710.

[28]Spirt v. Teachers' Ins. & Annuity Ass'n, 691 F.2d 1054, 1062–63 (2d Cir. 1982), *petition for cert. filed*, 51 U.S.L.W. 3394 (U.S. Nov. 23, 1982) (No. 82-791); Norris v. Arizona Governing Comm., 671 F.2d 330, 334 (9th Cir. 1982), *cert. granted*, 51 U.S.L.W. 3287 (U.S. Oct. 12, 1982) (No. 82-52). *Norris* also held that voluntary discriminatory pension plans as well as involuntary plans are forbidden. *Id.* at 333.

[29]*Spirt*, 691 F.2d at 1061; *Norris*, 671 F.2d at 334; *accord, e.g.*, EEOC v. Colby College, 589 F.2d 1139, 1144 (1st Cir. 1978).

[30]Peters v. Wayne State Univ., 691 F.2d 235, 238 (6th Cir. 1982), *petition for cert. filed*, 51 U.S.L.W. 3394 (U.S. Nov. 23, 1982) (No. 82-794).

[31]42 U.S.C. §§ 2000e–2000e-17 (1976 & Supp. V 1981).

harm that can emanate from this process: "Practices that classify employees in terms of religion, race or sex tend to preserve traditional assumptions about groups rather than thoughtful scrutiny of individuals."[32] In pension plans, for instance, requiring women to take home a lower monthly paycheck because greater contributions have been exacted, the practice outlawed in *Manhart*, or paying women lower monthly benefits upon retirement, reinforces the stereotype of a female employee as not being "worth" as much as a male employee. Such practices are offensive to the ethic of equal pay for equal work. Similarly, higher rates for women in health insurance may strengthen the notion of women as the "weaker sex." Stereotypes can be harmful to males, too. The exorbitant premiums paid by youthful male automobile drivers attach negative connotations to the entire group, when in fact most young males are safe and prudent drivers.[33]

More broadly, current practices foster an unhealthy view of society. Supporters of sex-based tables are concerned that a merged-gender table will cause one sexual group to "subsidize" the other.[34] But some individuals always subsidize the insurance of others; this is the very essence of insurance.[35] It is difficult to understand why it is so important that men not subsidize women, or vice versa. The conception of society immanent in such a view is one of fragmentation, of men versus women. The laws should instead encourage the ideal of a cohesive society composed of human beings who happen to be either male or female.[36] The current presumption that sex is an acceptable classification should be reversed. The Nondiscrimination in Insurance Act is a positive step in this direction.

---

[32]435 U.S. at 709.

[33]*1981 House Hearings*, *supra* note 8, at 63 (statement of Mary W. Gray, Women's Equity Action League).

[34]Kimball, *supra* note 16, at 106–08. *But cf.* Brilmayer, *supra* note 15, at 508–11 (criticizing this attitude).

[35]*Cf. Manhart*, 435 U.S. at 710.

[36]The idea that society should not be fragmented into blacks versus whites provided part of the reason for discarding the use of race as a risk classification. Indeed, supporters of the ban on sexual classification have analogized the use of sex as a factor to the use of race and have challenged their opponents to name a meaningful distinction. *See* Brilmayer, *supra* note 15, at 536–39. Distinguishing race and sex is difficult because blacks do die earlier than whites, S. REP. No. 671, 97th Cong., 2d Sess. 7 (1982), so that allowing race-based tables to be drawn presumably would permit more accurate predictions of longevity to be made. Arguments that race discrimination is "worse" seem extremely tenuous and dubious. Even one outspoken critic of merged-gender tables admits that no significant distinction can be made. Benston, *supra* note 16, at 512–13.

Yet opponents of merged-gender tables insist that people are being treated fairly as individuals under the present system. First, they assert that individuals receive equitable treatment if the cost of insurance to each person is directly related to the degree of risk that that person adds to the risk pool.[37] Second, they contend that equality among individuals in insurance is achieved when the present value of the expected benefit from an insurance contract is the same for everyone who has made the same payment.[38] For example, suppose a man and a woman have both created life annuities worth $100,000 through equal contributions. Assume the periodic payments begin upon retirement and the woman is statistically expected to live for ten years after retirement, while the man is expected to live only eight. At retirement the man will begin receiving $12,500 per year for the rest of his life, while the woman will begin receiving $10,000 per year. However, the present value of the future stream of income is equal, since the woman is expected to live an extra two years.

Both of these arguments are flawed from the standpoint of social policy. In the first place, the insurance industry seeks to justify assessing individual risk on the basis of group membership rather than on individual characteristics. The sex mortality differentials, for example, may reflect significantly different behavior patterns between men and women—such as smoking, drinking, military service, automobile driving frequency, and career pursuit—rather than any genetic factor.[39] The individual's contribution to the risk pool should be based on specific and direct factors rather than on an indirect criterion such as gender.[40]

Secondly, these arguments are flawed because they *first* make the decision to group by sex and *then* address the issue of equity

[37]*See* S. HUEBNER & K. BLACK, *supra* note 16, at 442; Benston, *supra* note 16, at 497.

[38]*See* Benston, *supra* note 16, at 503; Kimball, *supra* note 16, at 97–103.

[39]Lewis & Lewis, *The Potential Impact of Sexual Equality on Health*, 297 NEW ENG. J. MED. 863, 865 (1977).

[40]*See* Benjamin R. Schenck, former Superintendent of Insurance, Statement Before the New York Assembly Standing Comm. on Insurance, at 9 (Mar. 6, 1974) (quoted in Comment, *Gender Classifications in the Insurance Industry*, 75 COLUM. L. REV. 1381, 1391 n.50 (1975)):

> [I] do not believe we know enough, in most areas, to assert that the statistical differences which may exist are caused by sex, rather than merely reflecting the different sex roles created by society. If the latter is true, a good insurance system . . . would urge insurance companies to look beyond sex—to the true causes of statistical differences—and to make insurance price distinctions based on these true causes rather than based on sex.

among individuals. The justification for this initial decision is that a statistical difference in longevity exists between the sexes. But before one can even speak in terms of male and female life expectancies, one must decide to group by sex.[41] And, while there are several possible ways to classify any one individual, it is perhaps no more than habit that makes sexual classification appear acceptable.[42] This type of initial classification should be prohibited, because it is unfair to the individual and results in the problems of stigmatization and fragmentation of society.

II. Economic Implications of the Bill

A. *Temporary Economic Effects*

Much of the political debate surrounding the bill has focused on the cost impact of a merged-gender rate upon men and women as groups. For instance, it is projected that women in the aggregate will pay $360 million more for life insurance and $700 million more for automobile insurance.[43] However, women would pay approximately $106 million less for health insurance, excluding maternity coverage.[44] The largest benefit to women would come in pensions and annuities, where the bill mandates that benefits due after the effective date may be increased but not decreased to achieve equivalence between payments to men and women.[45] This controversial "topping up" provision would cost insurers between $2 billion and $5.5 billion.[46] Because men

---

[41]Brilmayer, *supra* note 15, at 511–12.

[42]*See Manhart*, 435 U.S. at 710.

[43]*1981 House Hearings*, *supra* note 8, at 276 (statement of the Am. Academy of Actuaries).

[44]*Id.*

[45]H.R. 100, § 4(c)(2). "Topping up" may be required by the Equal Pay Act, 29 U.S.C. § 206(d) (1976). Key, *supra* note 19, at 28 n.158.

[46]S. Rep. No. 671, 97th Cong., 2d Sess. 18 (1982). The $2 billion estimate was made by the American Council on Life Insurance. The $5.5 billion estimate, made by the American Academy of Actuaries, is likely to be revised downward soon. *Id.* Also, these figures assume that *Manhart* and its progeny do not demand equal benefits in employer pension, annuity, and retirement plans. *Id.* This assumption is extremely dubious given the recent cases of Spirt v. Teachers' Ins. & Annuity Ass'n, 691 F.2d 1054 (2d Cir. 1982), *petition for cert. filed*, 51 U.S.L.W. 3394 (U.S. Nov. 23, 1982) (No. 82-791), and Norris v. Arizona Governing Comm., 671 F.2d 330 (9th Cir. 1982), *cert. granted*, 51 U.S.L.W. 3287 (U.S. Oct. 12, 1982) (No. 82-52). The Supreme Court has already heard oral arguments on *Norris* on March 28, 1983. Arizona Governing Comm. v. Norris, 51 U.S.L.W. 3718 (U.S. Mar. 28, 1983) (No. 82-52); *see also Arguments Before the Court: Pensions and Retirement--Employer Liability for Insurers' Sex-based Annuity Tables*, 51 U.S.L.W. 3713 (1983). If the Supreme Court follows *Spirt* and *Norris* by directing that equal benefits be paid in the workplace, the bill will have much less effect.

may benefit under some of the pension or annuity plans, such as the joint and survival option plan or the early retirement plan,[47] the American Academy of Actuaries claims that it is unable to determine the additional cost to each sex.[48] However, it seems clear that the larger share of the reallocation would go to women and probably would more than compensate them for their increased costs for life and automobile insurance.[49]

Much of the "cost" of the bill is thus simply an income transfer between men and women. The "topping up" provision would saddle the insurance industry with a significant short-term cost, but this would eventually be passed on to consumers in the form of slightly higher rates for pensions and annuities.[50]

The magnitude of the income transfer may surprise one at first, but it should be remembered that it is an unavoidable result of requiring rates to be paid by people as individuals, not as men or women. The "topping up" provision appears to be an unfortunate necessity. It would be both unfair and politically unfeasible to reduce the annuity or pension payments of a person already in retirement who justifiably expects to continue receiving the same periodic payment. The public will eventually bear the cost of "topping up" as well as the administrative costs, which are estimated at $1.345 billion,[51] but these initial costs will be spread over time and among millions of people.

B. *Permanent Market Effects*

The Nondiscrimination in Insurance Act will create new incentives for insurers, which ultimately may result in permanent

---

[47]*See* Gold, *supra* note 16, at 693–96; *see also 1981 House Hearings*, *supra* note 8, at 277 (statement of Am. Academy of Actuaries).

[48]*1981 House Hearings*, *supra* note 8, at 277 (statement of Am. Academy of Actuaries).

[49]*See id.* at 170 (statement of Judy Schub, Nat'l Fed'n of Business & Professional Women Clubs). Indeed, the mere fact that so many women's organizations support the bill suggests the truth of this assertion. *See* S. REP. No. 671, 97th Cong., 2d Sess. 2 n.1 (1982) (representative list of organizations supporting the bill).

[50]Bernstein & Williams, *Title VII and the Problem of Sex Classifications in Pension Programs*, 74 COLUM. L. REV. 1203, 1224 (1974); *see also 1981 House Hearings*, *supra* note 8, at 72 (statement of Barbara J. Lautzenheiser, Am. Council of Life Ins.) (public, not insurance industry, will bear cost of legislation); S. REP. No. 671, 97th Cong., 2d Sess. 20 (1982) ("topping up" provision may result in a minimal increase in costs of annuity, pension, or retirement plans for individual consumers).

[51]*1981 House Hearings*, *supra* note 8, at 275 (statement of the Am. Academy of Actuaries). This estimate is based on an effective date of the bill midway between the 90-day period now provided in the bill and a "low-impact" two to three year period. The American Academy of Actuaries claims that the 90-day implementation period in § 11 of H.R. 100 would simply be impossible to meet.

market reallocations. These incentives may offset each other or be subverted by appropriate regulation or litigation.

The market effects stem from the fact that every group of insureds inevitably has a certain ratio of males to females. Because differences in longevity between gender groups do exist, the sexual composition of the pool directly affects the cost of the pool to the insurer. For instance, an insurer with a male-dominated pool can expect to pay out more in life insurance benefits in a typical year than an insurer with a female-dominated pool. Under the current system, sexual mix is automatically taken into account by charging different rates to men and women. However, with merged-gender tables, consideration of the sexual composition of the pool is necessary in order to prevent underfunding or overfunding and thus to maintain an actuarially sound business.[52]

Consequently, the Supreme Court in *Manhart*, although outlawing unequal contributions to the employer-operated pension plans based on sex-distinct tables, specifically permitted insurers to consider the composition of the employer's work force in determining the cost of a pension plan.[53] The Senate Committee on Commerce, Science and Transportation similarly recognized the importance of sexual composition of the group:

> The Committee is not questioning the fact that differences in rates for males and females have been observed in the past, nor that these differences are considered in estimating costs for the future . . . . The Committee has no quarrel with the actuary who insists upon his right to develop experience results by whatever group characteristics seem to have a substantive bearing on his results. The insurance industry, however, would have the obligation of seeing that the application of actuarial factors will not be used to establish differentials in benefits or rates between the sexes.[54]

Even if conscious consideration of the sexual ratio had been prohibited by the bill, companies could have recognized it through their experience ratings, as the Committee seems to

---

[52]*See* Rutherglen, *Sexual Equality in Fringe Benefit Plans*, 65 VA. L. REV. 243–44 & n.201 (1979); *The Supreme Court, 1977 Term*, 92 HARV. L. REV. 57, 311 (1978). Bernstein & Williams believe that neglecting to consider sexual composition in a group plan is not harmful because the estimates are crude and can be adjusted later. They were the only authority found who believed sexual composition to be essentially irrelevant. Bernstein & Williams, *supra* note 49, at 1228–30.

[53]435 U.S. at 718.

[54]S. REP. No. 671, 97th Cong., 2d Sess. 22 (1982).

realize.[55] Experience ratings measure the longevity, for example, of the actual group insured by the company. This technique is frequently employed by large companies who insure enough individuals to make the results statistically valid.[56]

Allowing insurers to consider sexual composition in estimating their costs but forcing them to charge rates based on a merged-gender table creates certain unwanted incentives. In life insurance, for instance, the insurer can reduce his costs by attaining the highest female to male ratio possible, since females represent lower risks than males. Because a company's merged-gender rate will be directly related to its costs, companies with a higher female to male ratio can offer lower merged-gender rates and thus defeat the competition. In a similar vein, since the merged-gender rate falls somewhere between the current male and female rates, men's rates will be relatively underpriced and women's rates relatively overpriced. Thus, insurance companies will, on the average, reap more profit from each sale to a woman. For these reasons, insurers will have a strong incentive to favor women over men in the life insurance market. Obviously, incentives to favor men will exist in the health insurance and annuities markets for precisely the same reasons.

Observers of the insurance market believe that insurers will take steps motivated by these incentives. For example, insurers may "do everything possible in their marketing, health requirements and otherwise," to avoid selling health insurance to women because of their greater loss potential.[57] One can imagine internal bonuses paid to agents who sell to the highest proportion of overpriced insureds.[58] Thus, the incentives created by the bill may make it more difficult for the underpriced group to obtain insurance.

While discrimination in availability violates the bill,[59] it may

---

[55]*Id.*; *see also* Note, *Sex Discrimination in Employee Fringe Benefits*, 17 WM. & MARY L. REV. 109, 133–34 (1975).

[56]Kimball, *supra* note 16, at 133–34.

[57]Comment, *Gender Classifications in the Insurance Industry*, 75 COLUM. L. REV. 1381, 1391–92 (1975) (citing Benjamin R. Schenck, former Superintendent of Ins., Statement Before the New York Assembly Standing Comm. On Insurance, at 8 (Mar. 6, 1974)).

[58]However, if internal commissions and bonuses can be prevented, these salesmen will have the same incentive to sell to all. Indeed, Huebner and Black note that the distribution of risks within a class is always skewed toward the higher risk end of the class because of the willingness of salesmen and home office underwriters to relax standards slightly to make a sale. S. HUEBNER & K. BLACK, *supra* note 16, at 444–45.

[59]H.R. 100, § 4(a)(1) makes unlawful differences in availability or even delay based

be difficult to monitor. Thus, more stringent government controls will have to be imposed on the marketing, underwriting classification, and rating functions of the industry in order to ensure that underpriced coverage is equally available.[60] Increased regulation of the industry, as well as litigation under the private cause of action created by the bill,[61] will impose a substantial cost upon society with uncertain effectiveness. However, strong regulation and successful litigation from the outset may reduce this cost. Nonetheless, the problem of insurers favoring certain customers may still exist regardless of regulation. For instance, insurance companies could probably initiate contacts, door to door, with only one gender group without violating the bill.

However, selection by the public may counterbalance the effect of discrimination by insurance companies in favor of better risks. Thus, people for whom the price of an insurance product is disproportionately high might buy less insurance coverage or none at all; those for whom prices are disproportionately low might buy more insurance than they otherwise would have.[62] In general, women will be more attracted to annuities than they currently are, while life insurance will become more attractive to men.

Whether revenues actually are decreased or not will depend on the interplay between this adverse selection factor and the predicted tendency of insurance companies to engage in more strenuous efforts to sell to the relatively overpriced consumer. In other words, the consumer who has the incentive to refrain from purchasing insurance will be precisely the one whom the insurance companies will try most strongly to attract. Thus, it is possible that the two primary market distortions caused by the bill will effectively cancel each other out. Whether they do or not, or more realistically, to what extent one compensates for the other, depends on economic factors such as the elasticity of demand for insurance and the degree of competition in the insurance market.[63] Greater elasticity would augment the adverse selection effect, while greater competition may increase

---

on a forbidden factor. Section 4(b)(3) outlaws any statement or advertisement by insurers indicating a preference or limitation on the basis of an illegal factor.

[60] *1981 House Hearings*, *supra* note 8, at 198 (statement of Andre Maisonpierre, Alliance of Am. Insurers); *id.* at 277 (statement of Am. Academy of Actuaries).

[61] H.R. 100, § 6.

[62] *1981 House Hearings*, *supra* note 8, at 73 (statement of Barbara J. Lautzenheiser, Am. Council of Life Ins.).

[63] One commentator has suggested that competition "is dulled because it is enormously

the incentive of insurers to use marketing devices to sell primarily to persons who are relatively better risks on a cost basis.

A further complication resulting from the incentive for insurers to discriminate against underpriced risks exists in the pension fund context. Numerous commentators have noted that requiring employers to sponsor pension funds that pay equal benefits to men and women upon receipt of equal contributions creates a disincentive to hire females.[64] In a system that is employer-operated, or in which the employer makes matching contributions, the employer can increase profits by substituting men for women.

Although it is certainly difficult to determine whether the disincentive would be sufficient to have a perceptible effect on the hiring of women,[65] future Supreme Court decisions under Title VII may make the effect of the bill on this phenomenon negligible. While *Manhart* outlawed the practice of demanding unequal contributions from men and women in employer-operated pension funds,[66] the Court may extend that decision to a ban on unequal benefits in the entire employer-employee context.[67] If employers were required by Title VII to collect equal contributions and pay equal benefits, the disincentive to hire females would still exist but it would be due to Title VII's restrictions, not the bill's.[68] Eliminating the discrepancy that might otherwise exist between the regulation of insurance under Title VII and the regulation of the private insurance market under the bill is a positive argument for passing the bill.[69]

---

difficult for consumers to compare the terms, or even the prices, of annuity and life insurance contracts." Brilmayer, *supra* note 15, at 531 n.125.

[64]Benston, *supra* note 16, at 510; Bernstein & Williams, *Sex Discrimination in Pensions:* Manhart's *Holding v.* Manhart's *Dictum*, 78 COLUM. L. REV. 1241, 1243 (1978); Kimball, *supra* note 16, at 133–34; Note, *Sex Discrimination and Sex-Based Mortality Tables*, 53 B.U.L. REV. 624, 703 (1973).

[65]At least one commentator believes that it would not be significant. *See* Key, *supra* note 19, at 18–20.

[66]435 U.S. 702 (1978).

[67]*See supra* note 45.

[68]This is the reason Bernstein & Williams object strenuously to the dictum in *Manhart* that the holding does not "call into question the insurance industry practice of considering the composition of an employer's work force in determining the probable cost of a retirement or death benefit plan." 435 U.S. at 718. Bernstein & Williams argue that this dictum should be ignored since it would create a disincentive to hire females. Bernstein & Williams, *supra* note 65, at 1242. However, as noted earlier, forcing employers and insurers to ignore the sexual composition of the work force can be subverted by the use of experience ratings. *See supra* notes 54 & 55 and accompanying text. The disincentive stems from the equal contributions-equal benefits mandate, not recognition of the sexual mix of the work force.

[69]*See* Laycock & Sullivan, *Sex Discrimination as "Actuarial Equality": A Rejoinder to Kimball*, 1981 AM. B. FOUND. RESEARCH J. 221, 228 (1981).

If sex is prohibited as a risk classification, insurance companies may seek other factors as replacements. Four requirements should be fulfilled in the development of new factors. The variable must be susceptible to accurate measurement, perceived relationships must be expected to extend into the time when benefits are paid, the variable should not be susceptible to moral hazard or self-selection, and the costs of assessment should not exceed the benefits, in the form of more accurate risk classifications, that derive from it.[70] The concept of moral hazard refers to the notion that personal factors that are not immutable, such as smoking habits, are easily lied about by the insured. The costs of monitoring such activity are likely to be prohibitive.[71] Thus, at least one commentator has concluded that at present only age and sex are capable of predicting longevity with substantial accuracy at a reasonable cost.[72]

Other problems cast doubt upon the likelihood of insurance companies quickly finding new unobjectionable risk predictors. The temporary losses faced by the insurance industry may absorb funds that would otherwise be used for research and development.[73] The use of factors such as smoking and drinking habits may impinge on the privacy of the insured.[74] Factors such as occupation may include a built-in sex bias. Finally, increased reliance on place of residence could harm minorities and the poor.[75]

However, industry representatives claim that market competition constantly spurs companies to develop new factors.[76] This suggests that the bill may actually provide a catalyst for the development of new methods of risk classification. The benefits derived from obtaining new and accurate risk predictors become more significant when the insurance industry is prohibited from using sex as a classification factor. Therefore, since greater benefits would be derived from a new variable, insurance companies should be willing to spend more to develop them. Thus, the bill may function in a positive way.

---

[70]Benston, *supra* note 16, at 499–500.

[71]*See* Kimball, *supra* note 16, at 106.

[72]*Id.* at 120.

[73]*1981 House Hearings*, *supra* note 8, at 278 (statement of the Am. Academy of Actuaries).

[74]*See id.* at 124 (statement of Wright H. Andrews, Jr., Nat'l Assoc. of Life Ins. Cos.).

[75]*See id.* at 31 (statement of Althea T.L. Simmons, NAACP).

[76]*Id.* at 189 (statement of Diana Lee, Nat'l Assoc. of Indep. Insurers); *id.* at 282 (statement of Jay C. Ripps, Am. Academy of Actuaries).

CONCLUSION

The battle over the Nondiscrimination in Insurance Act pits the merits of eliminating discrimination in insurance as a positive social goal against the short-term economic losses and the danger posed by possibly permanent market misallocations. The desirability of prohibiting different rates and benefits for men and women touches basic notions of fairness and equality. Current practices reinforce negative stereotypes of both sexes. Even worse, sex-based rates and benefits are based on an unhealthy conception of a society clearly fragmented into males and females for economic purposes.

The bill would cause some short-term losses to insurers that will later be passed on to the public. These should be reduced by extending the ninety-day implementation period to six months or a year. More permanent economic reallocations are also likely to occur, although precise effects are difficult to calculate. Because these distortions are more intractable, Congress should shift its focus from temporary losses and income transfers, which admittedly are more politically visible issues,[77] to the more serious issues of the bill's encouragement of discrimination in availability and its creation of an adverse selection effect. I have argued that these factors work in opposite directions and thus may cancel each other out. Also, if *Manhart* is extended to equal benefits and to private insurers operating through an employer, the bill will have little incremental effect on the possible disincentive to hire females created by the equal contribution-equal benefit situation. Developing other factors to replace sex will not be easy, but the bill provides the proper incentive to encourage insurers to invest in such research.

Although possible permanent market distortions should be the subject of further study, they are not likely to be nearly as disastrous as the insurance industry predicts. Passage of the Nondiscrimination in Insurance Act during the Ninety-eighth Congress would mark an important step in the fight for sexual equality.

---

[77] *See id.* at 19–20 (statement of Rep. Scheuer); *id.* at 63 (statement of Rep. Lent).

BOOK REVIEWS

REGULATION AND ITS REFORM. By *Stephen G. Breyer.* Cambridge, Mass.: Harvard University Press, 1982. Pp. viii, 472, appendices, notes, index. $25.00 cloth.

*Review by Douglas H. Ginsburg**

In *Regulation and Its Reform*, Stephen Breyer attempts to develop a normative theory of regulation, one that "considers *justifications* for regulation, not causes" (p. 9). He begins accordingly, offering a typology of regulatory regimes and an understanding of the difficulties inherent in the administration of each type of regime. His goal is to provide an analysis that can suggest, with respect to any particular objective, the regulatory regime best suited to its achievement.

Breyer begins his analysis with a "simple axiom for creating and implementing any program," namely that one should first determine clearly the program's objectives, examine the alternative methods of achieving those objectives, and then choose the "best method" for doing so (p. 5). All the programmatic objectives Breyer has in mind can be brought within the rubric of correcting market failures. In other words, he starts with a strong presumption in favor of unregulated markets, and then makes the non-controversial observation that departures from reliance on the market should be as well-tailored as possible to the correction of specific defects in the market outcomes.

I. TREES

In Chapter 1, Breyer surveys what he describes as the "typical justifications for regulation" (p. 15). These are control of monopoly power; control of rents or "excess profits"; compensation for externalities generated in the production of goods; inadequate production in the private market of information relevant to consumption choices; and avoiding "excessive" competition, which he rejects as an "empty box" (p. 29).

Breyer gives short shrift to several other possible justifications for regulatory systems on the oddly irrelevant and factually dubious ground that they are less often advanced in the United

*Professor of Law, Harvard Law School; B.S., Cornell University, 1970; J.D., Harvard Law School, 1973.

States than elsewhere (p. 32). The first is unequal bargaining power, which, as he acknowledges, underlies the antitrust exemption for labor unions and agricultural cooperatives. This justification is also frequently advanced to support proposals purporting to protect consumers in standardized contractual relationships such as those with landlords and utilities. The second justification is "rationalization," another name for industry-wide planning. Breyer notes that in the 1960's the Federal Power Commission unsuccessfully tried to encourage coordination and planning in operations among electric generating and transmitting facilities (p. 33). That is indeed an obscure example, especially since similar ideas are again so much in vogue. Rationalization is now being advocated for application to industries that are viewed either as potential sources of economic growth and exports, or as victims of coordinated competition from the Japanese, under such names as "indicative planning" and, most recently, "industrial policy."[1] The third in this list of "other" predicates for regulation is the problem of "moral hazard," as seen in the increasing governmental regulation of health care and health insurance (p. 33).

With respect to paternalism, a fourth justification for regulation independent of market failure, Breyer questions whether such a "mistrust of consumer rationality is consistent with the notions of freedom of choice that underlie the free market," but at the same time acknowledges that "it plays an important role in many governmental decisions" (p. 34). It has always seemed interesting to this writer that paternalism is not an acceptable justification for governmental action in our constitutional system and yet appears to serve in fact as an important explanation for many government regulations.

American traditions of individualism, self-sufficiency, personal freedom, and consumer sovereignty are flatly inconsistent with pure paternalism, which is the notion that the choices of even a fully informed consumer should be overridden by the state on the ground that the choice is bad for the person in question. Yet we see regulators and legislators churning out, with increasing frequency, restrictions on individual choice that cannot plausibly be explained except upon the view that "the poor fools don't know what is good for them." A government would be unlikely to admit, however, on grounds of electoral

[1] *See, e.g.*, I. MAGAZINER & R. REICH, MINDING AMERICA'S BUSINESS (1982).

politics alone, that it holds this view of its citizens, much less claim that protecting fools from themselves is a legitimate role for the police power of the state. Thus, laws requiring motorcyclists to wear helmets are struck down by one court as purely paternalistic, but upheld by another on the ground that the fools who ride motorcycles and get their heads smashed might become wards of the state or might consume state-provided medical resources, and thus the state must protect the public from this burden. If this is but the nose of a camel, the tent of individual liberty is probably not large enough to accommodate the rest of the beast.

Finally, Breyer notes that regulation is sometimes justified in terms of "scarcity," but instead of noting that everything is scarce in this scarcest of all possible worlds, he makes the marvelously even-handed observation that a scarcity or a shortage calling for regulation may arise either from a sudden failure of supply, which might work an unacceptably severe hardship on some portion of the public, or from regulation itself, "as when natural gas must be allocated because of rent control or when an agency awards licenses to enter an industry" (p. 34).

Having surveyed these various rationales for regulation, Breyer devotes the next six chapters to the six types of "classical regulation" in his typology. The first type is cost-of-service ratemaking, which may be familiar to the nonspecialist as public utility regulation. Local utilities that supply electricity, gas, water, and telephone service, if they are not municipally owned, are regulated in such a way as to allow them to earn a reasonable return on their investment—ideally no more and no less. A variety of problems arise in determining what the regulated company's investment is, what constitutes a reasonable rate of return, how best to structure the company's rates for various classes of consumers to enable it to cover its fixed costs, and how to deal with the departures from efficient levels of investment and productivity that regulation is likely to induce. Breyer surveys these problems in an absolutely masterful twenty-three pages. For one who has tried to teach public utility regulation to law students, every page is humbling.

The second classical mode of regulation Breyer examines is "historically based price regulation," by which he means economy-wide price controls based on prices during a base period, rather than on each individual firm's costs, as in cost-of-service ratemaking (p. 60). As he observes, historically based price con-

trol is a more practical technique than cost-of-service ratemaking to regulate the prices of a large number of firms with disparate costs. Breyer's thesis is that the inequities and anomalies of an historically based price control system mount rapidly with time, however, and therefore that the system increasingly tends, due to the superimposition of waivers, exceptions, and pass-through provisions for cost increases, to resemble cost-of-service ratemaking. For illustration, he draws liberally on World War II experience under the Office of Price Administration, with occasional references to the briefer dalliances with price freezes and price controls during the Korean War and the later stages of the Vietnam War.

It is not entirely clear why Breyer considers historically based price regulation to be one of the six classical modes of regulation. The system is analytically related to cost-of-service regulation and, as Breyer notes, tends to evolve toward it. As Breyer acknowledges, however, historically based price regulation has been attempted only for relatively brief emergency—indeed wartime—periods. The author in fact never returns to a discussion of this type of regulation, except to note, in a later section concerning the use of taxes as an alternative to classical regulation, that oil price regulation instituted in the 1970's to control rents and windfall profits was a variation of historically based pricing (p. 165).

Breyer's third "classical" mode of regulation is allocation under a public interest standard. This is the means—one dares not call it a system—by which television and radio broadcast licenses are handed out, and by which the Federal Communications Commission (FCC), like a lottery, turns shopkeepers into millionaires. Before deregulation, regulators also awarded airline routes and exceptional allocations of natural gas, after determining which of many applicants for the valuable right, by privately exploiting it, would best serve the "public interest." Breyer uses the FCC experience as his text. He tells the oft-told horror story that traces the FCC's devolution from its public interest mandate to its many incoherent and conflicting standards, through its interminable proceedings, to its incomprehensible, inconsistent, and, at least for a time, corrupt decisions.

To get the flavor of the FCC's problem, consider the following not atypical dilemma. A commissioner is asked to choose among nine applicants for a television broadcast license, seven of which are indistinguishable in most respects and incommensurable in

others, another of which is unique in having various ethnic minorities among its noncontrolling shareholders, and the last of which publishes a newspaper that is editorially consistent with the views of the commissioner. It must be difficult for that commissioner to decide whether to select a winning applicant by throwing dice, throwing a bone to minorities, or throwing in the towel by voting for the newspaper because it endorsed the president who appointed him.

The fourth classical regulatory mode Breyer examines is standard-setting. Again he examines what one can only hope is, like the FCC's award of licenses, a "worst case" example of regulatory behavior. This is, of course, the National Highway Traffic Safety Administration (NHTSA), which sets standards for automotive safety and fuel economy. One can feel some sympathy for an agency charged in the late 1960's with introducing a healthy and successful automobile industry, as it then was, to the facts of life in Washington. The industry was realistically hoping to sell more than ten million cars per year to consumers who could have bought funny-looking little foreign cars if they had wanted more fuel efficiency, or not-so-funny-looking, not-so-little foreign cars if they had been more concerned about safety. Not surprisingly, the industry did not see the need for its regulation—it was doing exactly what its customers wanted it to do. Unfortunately, the resulting intransigence on the part of the industry, which had vastly superior access to technical information, was later matched by the paranoid regulatory style of Joan Claybrook, a Nader protege, who became the NHTSA Administrator under President Carter.

Breyer uses this not entirely typical context to illustrate how standard-setting evolves toward a negotiating process in which industry and government positions and strategies are very much influenced by: (1) the parties' differential access to information, (2) the difficulty that the agency would have in enforcing a regulation against an unwilling firm, (3) the anticompetitive effects that the regulation will almost certainly have, and (4) the ever-present prospect of judicial review, which makes the negotiating relationship even more adversarial, formalized, and consequently prolonged (pp. 109–18).

Breyer suggests that these four factors "help explain why the standard-setting process inevitably deviates from the policy planner's cost/benefit model" (p. 118). This assertion cannot be denied, but it should be kept in perspective: deviations from

models are a matter of degree. Standard-setting is rarely undertaken in such an adversarial context, by such an inept agency, and against such a strongly protesting and unified industry. There are tens of thousands, perhaps hundreds of thousands, of standards that regulatory agencies either develop or borrow from industry without equivalent difficulties. From local building code standards for electrical wiring and household plumbing pipe, to the Federal Aviation Administration's standards for the certification of aircraft, one cannot find a standard-setting process as troubled as that of the NHTSA in dealing with automotive safety and fuel economy. That is not to say that the NHTSA's experience may not provide a good basis for illustrating the potential problems inherent in standard setting. Indeed, if anything, Breyer might have given more emphasis to the potential for the anticompetitive use of the standard-setting process by firms that could benefit from it.[2]

Contrary to Breyer's assertion, the NHTSA experience does not very clearly "explain why the process resembles negotiation, why standards are so closely tied to precedent, why the process is time-consuming, and why standards once set . . . prove relatively immune to revision" (p. 118). The standard-setting process does not generally resemble negotiation so much as it reflects industry-agency cooperation, which is itself often detrimental to the public; nor is it closely tied to precedent, nor time-consuming, because judicial review is not probable in most standard-setting contexts. While it is true that standards, once set, tend to resist revision, that is probably due less to the process that produces it than to the fact that the firms protected by an existing standard are in a better position to defend it than are firms whose new technology is excluded from the market by the standard to oppose it. In general, standard-setting thus tends to act as a contract between risk-averse incumbent regulators and revenue-producing incumbent technologies to exclude new technologies and new regulators for as long as possible.

The next chapter in this sextet is on historically based allocation. Breyer uses the example of oil allocation by the Federal Energy Office (FEO) in response to the Arab oil embargo in late 1973. He presents the system as being, like historically based price regulation, inherantly unstable. Breyer notes that "[i]t

---

[2]*See* American Soc'y of Mechanical Eng'rs, Inc. v. Hydrolevel Corp., 102 S. Ct. 1935 (1982).

soon evolves toward either standard-setting or public interest allocation," and develops the problems characteristic of those regimes (p. 120).

The last "classical" regulatory technique discussed is individualized screening. There Breyer examines the Food and Drug Administration's (FDA) food additive screening system. Breyer uses the FDA case to describe a variety of recurring and difficult problems in the regulation of scientific subject matter. These include the development of a test for risk; the problematic desire, if not necessity, to rely upon experts while at the same time appreciating the uncertainty under which the experts themselves are laboring; and the difficulty of even positing cost/benefit calculations where there may be lives or disabilities at stake on one or both sides of the balance. For example, an agency that must decide whether to ban the general use of saccharin in order to avoid the risk that it is a carcinogen must also consider the effects of such a ban on weight gain, a contributor to death from heart attacks and strokes. To lawyers and others who are interested in regulation but are not specialists in the food and drug area, this chapter is a useful addition to the literature.

Breyer concludes this extremely sensitive and discriminating chapter with the following observation:

> Many of the problems of individualized screening grow out of the fact that neither science nor professional training nor common sense is likely to provide administrable, accurate tests for the regulator to use . . . .
>
> As the regulator becomes ever more interested in fine-tuning the screening process (and as he seeks to uncover ever smaller risks), he faces serious problems in measuring risk and assessing benefits. He cannot rely fully upon expert committees to substitute their judgment for his. He is likely to make scientifically irrational distinctions between, say, old and new substances. The upshot has been a variety of suggestions to use less restrictive devices. For example, clear disclosure in the form of labelling may be preferable to banning a substance where the risks are not serious or where benefits are potentially great . . . . But one can determine the comparative advantages of major alternatives to classical regulation only after considering the advantages and defects of these alternatives (p. 155).

Thus ends Breyer's long survey of the six modes of classical regulation. And thus begins his brief survey of what he considers the principal alternatives—which can be used as either substi-

tutes for or supplements to classical regulation. These alternatives turn out to be reliance upon (1) antitrust laws, primarily in order to prevent such predatory pricing as might arise from excessive competition; (2) disclosure requirements; (3) taxes (which can be used (a) to transfer income, for instance, from those who control a suddenly scarce resource, like oil, to those who must purchase it, or (b) to shape behavior, such as polluting, while at the same time perhaps providing more flexibility than would a fixed standard such as a maximum effluent discharge regulation); (4) creation of marketable property rights, such as a saleable right to discharge an effluent, to broadcast on a radio frequency, or to land at an airport; and (5) liability rules, which can be altered judicially or legislatively to give individual citizens or classes the ability to use court processes to protect themselves from harms in the production and consumption of goods. Breyer also deals with (6) bargaining as an alternative, or an input, to the regulatory process, particularly with respect to occupational safety and health standards. The basic idea is that bargained-for standards might be more useful than rigid regulatory standards because they could be tailored to particular work settings in a manner that federal agency rulemaking, or even adjudication, virtually precludes.

Finally, Breyer raises the possibility of (7) nationalization, which could "in principle" be used to deal with problems of natural monopoly or of significant externalities, but which Breyer gives only passing consideration, apparently because of "the relative unpopularity in the United States of nationalization as a solution to regulatory problems" (p. 183). Thus, he does not examine any of the principal examples of state enterprise in the United States, such as the United States Postal Service, the Tennessee Valley Authority, the Federal Reserve Banks, the Federal Deposit Insurance Corporation, or the many municipally owned local utilities.

On the basis of his three surveys—market failure justifications for regulation; classical types of regulation; and, more briefly, alternatives to classical regulation—Breyer concludes Part I of his book with "General Guidelines for Policy-Makers." These guidelines, which he connects up, more or less, to the preceding chapters, can be summarized as follows:

(1) Regulation should not be used unless it is clear that a situation involves serious market defects for which regulation offers a cure;

(2) Regulators should aim at the worst cases and strive for simplicity, because "increased efforts to fine-tune regulation often will not yield improved results" (p. 185); and

(3) Finally, "classical regulation ought to be looked upon as a weapon of last resort. The problems accompanying classical regulation would seem sufficiently serious to warrant adopting a 'least restrictive alternative' approach to regulation" (p. 185). By this, Breyer intends to preserve as much of the unregulated marketplace as possible, overlooking minor market defects and considering "incentive-based intervention, such as taxes or marketable rights, or disclosure regulation, bargaining, or other less restrictive forms of interventions before turning to classical regulation itself" (p. 185).

Breyer acknowledges that the preceding chapters do not prove the validity of his approach. Rather, by bringing out the problems inherent in classical regulation and celebrating the virtues of a competitive marketplace, they support the reasonableness of a preference hierarchy that runs from unregulated markets to incentive-oriented intervention, and only as a last resort, to thoroughgoing substitution of regulation for the market—at least to this reviewer, who has studied regulation long enough to doubt whether, taken as a whole, it offers a net benefit to the public.

II. Better Trees

Part II of Breyer's book is entitled "Appropriate Solutions," and it is a curious business indeed. Here Breyer introduces his "mismatch thesis," which is that:

> regulatory failure sometimes means a failure to correctly match the tool to the problem at hand. Classical regulation may represent the wrong governmental response to the perceived market defect. If so, regulatory reform would consist not of procedural or structural changes within the agency, but rather of a detailed examination of a particular program with an eye toward radical substantive change—a change in the use of classical regulation itself. Recently, such major changes in regulatory programs—in airlines, trucking, and natural gas—indicate the gradual acceptance of this point of view (p. 191).

Breyer then uses a table to suggest tentative "matches" between typical market failures and possible solutions, including

both classical regulation and the alternatives he proposes (p. 192). For example, natural monopoly is matched with both cost-of-service ratemaking and nationalization—in the author's lexicon, with one "classical" and one "alternative" response. The control of rents or excess profits is matched with both taxes and deregulation, both alternatives to classical regulation.

After explicating the chart, Breyer offers three maxims to help identify areas in which mismatches are likely to be formed. First, classical regulation should not be used to control excessive competition. This conclusion is based upon the next two chapters, which deal with airline and truck regulation. They suggest that deregulation and reliance upon the antitrust laws are more suitable approaches than cost-of-service ratemaking in structurally competitive industries. Second, either taxes or deregulation is preferable to classical regulation for purposes of rent control. This conclusion is illustrated later in a chapter that details the sorry history of natural gas price regulation, again mismatched with a variation on cost-of-service regulation. Finally, externalities are more effectively dealt with by taxes, marketable rights, and perhaps bargaining, than by classical regulation alone. Here, standard-setting used in various pollution control efforts provides an example of a "partial mismatch."

Breyer concludes Part II with a chapter on what he calls a "possible match," namely natural monopoly regulation of long-distance telephone rates. Breyer's point here is that although local telephone service is a natural monopoly and should be regulated as such, it is not at all clear whether long-distance telephone service should be brought within the same regulatory ambit, subjected to a less intrusive form of regulation, or deregulated entirely. The argument rests upon a complex set of assumed conditions that is substantially altered by the recent consent decree under which AT&T's local operating companies will become independent of any long-distance carriers.[3] It should be said, however, that Breyer's description of the process by which long-distance telephone service became somewhat competitive before the consent decree is another gem of succinct explication, without oversimplification, of a complex situation.

---

[3]*See* United States v. American Tel. & Tel. Co., 1982-2 Trade Cases (CCH) ¶ 64,979 (D.D.C. 1982).

The curiosity of Part II is more in what it fails to do than what it does. Part I surveyed six regulatory techniques—with illustrations drawn from public utility regulation, wartime price controls, FCC station licensing, NHTSA standard-setting, FEO gas allocation, and FDA food additive screening—and then briefly surveyed several so-called alternatives to classical regulation.

In Part II Breyer gives us four case studies—regulation of the airline, trucking, natural gas, and long-distance telecommunications industries. He shows that three of the six classical techniques explained in Part I have been misapplied in these areas and explains why three of the alternative techniques (antitrust, taxation, and marketable property rights) should have been used in their place. The other three classical regulatory techniques surveyed in Part I do not show up in the case studies, either as techniques that were used or even as techniques that should have been used. Similarly, the other four alternatives to classical regulation (disclosure, liability rules, bargaining, and nationalization) are neither the subjects of any case studies, nor are they suggested as possible solutions in any of the case studies we are given to illustrate mismatches, partial mismatches, and possible matches between ends and means. This disjunction can be traced, I think, to the arbitrary way in which Breyer has implicitly defined regulation, or at least the regulatory universe, and has divided it between "classical" and alternative forms of regulation.

III. Forests

Breyer specifically acknowledges, under the heading of "Limitations" in the Introduction, that "no serious effort is made to define 'regulation' . . . . Efforts to distinguish intellectually between . . . government 'regulatory' action and the entire realm of governmental activity are difficult and the subject of controversy. This book does not enter or resolve that controversy" (p. 7).

Breyer does not find it necessary to define what he means by "regulation" because he does not say anything about "regulation" as a whole. Instead, he has identified six types of regulation, which he has arbitrarily denominated "classical," and contrasted them with another, arguably less intrusive, set of alternative approaches to achieving governmental objectives. Some or all of these alternatives are "regulatory" by any defi-

nition. Some, such as taxation, may be more "intrusive" than a classical regime like standard setting, depending upon the amount of the tax, the stringency of the alternative standard, and one's notion of intrusiveness. Breyer even admits that at least one of his alternatives, nationalization, is more, rather than less, intrusive than classical regulation by his own measure of intrusiveness.

The question of whether a particular approach to solving a problem is deemed to be "regulation" is of little moment, of course. More important is the criterion by which a particular approach is included or excluded from the study. Perhaps this is less important to Breyer because his treatment of the various regulatory responses that he does include in the study is explicitly normative. In his words:

> [T]he approach of this book is normative. It considers *justifications* for regulation, not causes. It discusses the problems of regulation from the point of view of a conscientious regulator who is trying to do his job well. This approach differs from efforts to produce general causal theories explaining the origin of regulation or its effects (p. 9).

If Breyer were more concerned with the origins of "regulation," perhaps he would feel compelled to define the term. In fact, however, he is concerned with the proper uses of regulation, and more specifically with the proper uses of his arbitrary collection of six "classical" modes of regulation, and the various alternatives to them that he surveys. His exclusively normative focus thus frees him from the difficult task of justifying the way in which he has defined the scope of his inquiry. For example, he does not have to answer the question of why he includes taxes and marketable rights as alternatives to standard setting within the scope of his inquiry into pollution control, but omits any serious discussion of liability rules and excludes any mention of subsidies or government procurement policy. Yet, procurement also can be used to induce technological change, and it is more incentive-oriented than standard-setting. Both procurement and subsidies can be used to foster technological improvements, which is part of the challenge of controlling pollution. In sum, this is not a book about the means by which governmental objectives could be achieved, but rather about the means that Breyer is interested in or believes are politically feasible. It is probably not in fact a matter of Breyer's interest because he does not even seem to be interested in some of the

alternatives that he mentions. He gives no sustained discussion of disclosure, liability rules, bargaining, or nationalization, which are four of the seven alternatives to classical regulation that he mentions.

The real limitation of this book, however, is precisely that it is purely normative. Breyer eschews any attempt at a positive analysis of either the origins or the persistence of regulation, however it is defined. "Rather, the book's premise is that the comparative merits of alternative methods should be analyzed in terms of the other factors given here, and *then* historical, political, procedural, or (administrative) structural considerations should be introduced to see if, and how, they change one's judgment about the 'best' method" (p. 9) The introduction of these other factors is simply beyond the scope of this book.

To restate the matter somewhat tendentiously, Breyer has decided to analyze the characteristics, strengths, weaknesses, and inherent problems of a half dozen regulatory techniques, and more or less to consider another half dozen alternatives, in order to link each with the types of problems to which it is suited, disregarding the historical, political, and administrative contexts within which real world regulatory regimes are proposed, adopted, and applied to real world problems.

Who would want to know the results of this exercise? No one, I fear, except a hypothetical character Breyer creates for the occasion, the "conscientious regulator who is trying to do his job well" (p. 9). This is perhaps an example of the perversity by which a supply is thought by some to create its own demand. This conscientious regulator presumably lives and regulates, in an ahistorical, apolitical, procedural and institutional vacuum. This hypothetical regulator has no past, no political party, and no purpose except to "do his job well." He is the personification of the regulation itself, which is proposed by no one, comes from nowhere, and purports to do—what? To regulate? To what end? For whose benefit?

Again, Breyer produces what he seems to think will be an ahistorical, apolitical guide for his equally neutered regulator. This guide is what a "consensus of fairly knowledgable outside observers would consider a reasonable human goal, given our society and our economic and governmental systems" (p. 8). What can this possibly mean? What are these outside observers fairly knowledgable about? Reasonable human goals? Our society and our economic and governmental systems? If they

know about our society and our economic and governmental systems, then they can answer precisely the questions that the positive analysts have been asking but Breyer seeks to avoid, namely: Where did these regulations come from? Whose interests are they supposed to serve? As a conscientious regulator, what goals—whose goals—should I be pursuing?

It is no answer to say "reasonable human goals." If even a "conscientious regulator" wants to know whether to allocate broadcast station licenses by using dice, to help minorities, or to further his political party, surely he cannot decide by asking which one is a uniquely reasonable human goal. I am reminded of a student at Harvard College who, while taking the introductory course in economics was first exposed to Regulation Q, the limitation on the interest that banks could pay for deposits, and inferred that the regulation was intended to limit saving and spur consumption. He had, in other words, inferred its intent from its obvious and necessary effect, which is one mode of positive analysis, the "by their fruits shall ye know them" school.

Breyer's sleight of hand is to assume that a group of fairly knowledgable outside observers would reach a consensus about the purpose of a given regulatory statute. Legislators in fact may have a variety of different reasons for favoring a particular law—money, votes, principles, or ignorance. The affected industry would typically also be divided: those firms that perceive that they will benefit from the regulation will, of course, favor it. It is unlikely in the extreme that there is a unique answer to the conscientious regulator's question of what to do when faced with a particular regulatory choice.

IV. Branches

Part III of the book is entitled "Practical Reform." Its two chapters are appended to, but are not really a part of, the book. The first chapter details the process and the strategy by which the author, as counsel to the Senate Judiciary Committee, orchestrated its hearings on airline deregulation and made the issue of airline regulatory reform "politically visible and live" (p. 339).

This chapter is awkwardly pitched; the author seems manipulative, if not cynical, in discussing the strategy he used in the airline deregulation hearings, for example in his emphasis on

the importance of the order in which witnesses were called. The hearing day is described as being "structured to provoke debate and drama" (p. 335). "The net dramatic effect" of Senator Kennedy's interrogation of the CAB chairman "was to help erase the impression of a scientific, accurate set of CAB pricing rules and suggest a confusing, complex system that served no apparent public purpose" (p. 336). Yet, this reviewer knows from discussion with the author that no such manipulative bravado was intended by Breyer at all. On the contrary, Breyer was trying to make the point that arguments on the merits, specifically arguments derived from the analysis contained in the book, can be persuasive. His point was that the political process can ingest the mismatch analysis, and that as applied to a particular case, it can be powerful enough to elicit legislative change, at least when persuasively put forth.

The absence of a positive analysis is again quite limiting. Make no mistake: airline deregulation did not come about *solely* because of the force of Breyer's analysis. But he omits the politics of the matter. It is therefore hard to know whether the analysis would take one very far in arguing for the correction of some other regulatory "mismatch." The airline industry, for example, contains a small number of firms. It is not, in the electoral scheme of things, a significant employer. It sells a consumer service primarily to middle-class and business consumers. As a consequence, most legislators would have few airline employees, but many airline patrons, among their constituents. It gradually became clear that deregulation would benefit passengers, and thus most legislators. Why was this not equally clear when regulation was first imposed?

Contrast banking deregulation. There are 15,000 commercial banks, ranging from the tiniest newly formed institution with perhaps $100,000 in capital to the largest holding companies, with assets of over $100 billion; every business and virtually every consumer (and therefore almost every constituent) is a bank customer. There are a variety of banking trade organizations representing, among others, small banks, large banks, bank holding companies, and banks involved in securities activities, each with different interests. Surely the political economy of airline deregulation and of banking deregulation will differ, regardless of whether the mismatch analysis applies with equal or even greater force to some particular banking regulation. Without attention to a positive theory of regulation and dere-

gulation, however, one can only guess whether Breyer's mismatch thesis can be given practical force outside of the context in which it obviously originated, namely, airline deregulation.

The second chapter in Part III, the last chapter in the book, surveys the major "generic" approaches to regulatory reform. These include proposals for the hiring of "better people"; procedural changes; structural changes in both internal agency management and in congressional, presidential, and judicial supervision of agency performance; and finally, proposals designed to encourage substantive change toward less restrictive alternatives, such as competitive impact statements. Breyer provides an adequate survey of various such reform proposals made by others, but it is reasonably clear that his heart is not in them. He acknowledges the contributions that these proposals can make to his concept of regulatory reform, the correction of mismatches and the substitution of less for more intrusive regulatory regimes, but they are quite beside his main point.

V. ACORNS

Breyer's main point is quite beside the real value of his book. The value of the book is in its case studies and its detailed explanations of how individual regulatory regimes have operated, of the problems they have encountered, and of the ways in which they have failed. As in the airline hearings he designed, it is clear that Breyer's talent is the lawyer's mastery, and virtually unanswerable presentation, of the detailed facts that make a case. Breyer's various micro-analyses, and his command of the facts, make convincing cases for the wrong-headedness of several of the regulatory regimes he has studied. *Of course* airline regulation was misconceived—as a matter of economic efficiency; so too were regulation of motor carriers and natural gas producers—as matters of economic efficiency. The economic case is relatively simple, and has been made before;[4] Breyer merely makes it in a perhaps more lawyer-like fashion.

Breyer's grander mismatch thesis is ultimately unsatisfying. If regulatory regimes consistently betray a mismatch of announced (or "reasonably human") ends and chosen means, it does not seem sufficient to point out what a band of fairly

[4]*See, e.g.*, A. KAHN, THE ECONOMICS OF REGULATION: PRINCIPLES AND INSTITUTIONS (1971).

knowledgable outside observers would notice, and to call the effort law reform. At some point, one must try to understand *why* the mismatches recur. Perhaps legislators and regulators do not usually care about efficiency, or, at any rate, have other concerns of greater importance to them. Either one harbors a no-learning-by-doing model of human behavior and political institutions, or one must conclude at some point that the process giving rise to these regulatory "mismatches" does not consider them mismatches at all. Then one begins to craft a scientific—that is, positive—study of the origins and workings of regulation, which would advance knowledge rather than advance only one's opinions.

THE HOUSE AND FOREIGN POLICY: THE IRONY OF CONGRESSIONAL REFORM. By *Charles W. Whalen, Jr.* Chapel Hill, N.C.: University of North Carolina Press. 1982. Pp. x, 207, appendix, notes. $9.95 paper, $18.95 cloth.

*Review by S.E. Billet**

The value of Charles Whalen's most recent book, *The House and Foreign Policy*, lies in its joining of Whalen's practical political experience with a systematic academic analysis of the congressional reforms of the 1970's. Most present and former members of the House and Senate fail to make this connection in their writings.[1] The members characteristically advance a particular partisan position, lack rigorous objectivity, or rely on purely anecdotal and unsystematic methodology. Whalen has avoided these pitfalls and has made a worthy contribution to the body of knowledge on the Congress and the impact of congressional reforms on the formulation of United States foreign policy.

Whalen devotes much of the first four chapters to chronicling the conditions that led to the adoption of the reforms made in

*Legislative Assistant to Representative Barbara A. Mikulski (D-Md.). B.A., East Carolina University, 1974; M.A., East Carolina University, 1979; Doctoral Candidate, Department of Government and International Studies, University of Notre Dame.

[1]Notable exceptions include Representative Richard Bolling (D-Mo.) and Representative Paul Simon (D-Ill.); *see* R. BOLLING, POWER IN THE HOUSE (1968); P. SIMON, THE ONCE AND FUTURE DEMOCRATS (1982). reviewed in this issue at 681.

the House in the 1970's and to detailing those reforms. Whalen describes the House before the reforms as having a secretive internal organization whose activities were controlled mainly by its leadership cadre. The House leadership set the agendas, and the committee chairs jealously guarded their prerogatives to hold closed hearings and to control the appointment of subcommittee chairs. These internal politics probably were sufficient to encourage a reform movement among the membership of the House. But a further impetus for reform was provided by the concurrent tragedy of Vietnam and the executive's abuse of its war-making powers. With these developments grew the frustrating realization that the nation's elected representatives had lost control of the nation's foreign policy. A growing number of House members began to view reform as a vehicle to enable the House to reassert its constitutional right to influence the making of foreign policy.

According to Whalen, the reforms that the House enacted had three interdependent objectives: "decentralization of power within the House; greater openness in House procedures; and strengthened House capacity to deal with the executive branch" (p. 26). Many of the reforms were directed at the House committee system and at the use of seniority to distribute power within that system. By restricting the number of subcommittee chairs that could be held by any one member, and by guaranteeing that subcommittees would be provided sufficient staff resources, the reform movement accomplished at least two steps toward decentralization of power within the House. First, by expanding the number of members that shared the leadership powers of the House, it distributed the control over legislative activity more evenly among members of the majority party. Second, by creating multiple power loci, the reform movement set up a subcommittee system whose leadership had broad discretion to take action in many areas without centralized control by the House leadership. This trend toward individualism was encouraged by the inability of the national party organizations, by their virtual nonexistence, to develop a unitary policy in the absence of control by a centralized House leadership.

As part of this series of reforms, House hearings were opened to the public and House voting procedures were revised, thereby making it possible and necessary for members to develop individual positions in a vastly expanded policy arena. These reforms, in combination with a decline in the need to follow

dictates from the House leadership, exposed House members to a multitude of alternative policy positions. In the area of foreign policy, representatives were now expected by their constituents to take positions on issues like the war in Southeast Asia. Members responded by taking public stands that were far afield of their committee responsibilities. Many members lacking expertise in foreign and defense affairs attempted to educate themselves in these areas—at least to the extent necessary to take informed positions, to make periodic floor statements, or to promote occasional amendments to foreign affairs legislation.

To aid representatives in assuming positions on this expanded policy spectrum, "think-tanks" were created within the House, such as the Members of Congress for Peace through Law and the Congressional Clearinghouse on the Future. These groups possess sufficient resources to enable members without a foreign affairs or armed services committee assignment to address foreign policy issues with some competence. Private foundations and interest groups, ranging from the Heritage Foundation to Common Cause, also have proven to be a valuable source of information and assistance on foreign policy matters, but these resources often are underutilized by members and their staffs. Nevertheless, the result of this increase in educational resources is that House members have become better able to examine the executive's foreign policy, to make informed criticism, and to offer useful alternatives.

Spurred by abuses of executive power during the Vietnam war, Congress sought further reforms by attempting to reassert its foreign policy role through the passage of the War Powers Resolution in 1973.[2] This legislation was intended to serve as a vehicle to control the executive's war-making prerogatives and to revitalize Congress' role in the decisionmaking process.

The author concludes, however, that despite their role in making the House a more democratic institution, the reforms have had an overall adverse effect on the nation's foreign policy and on the ability of the House to aid constructively in its development. Whalen's review of some of the more horrible errors of foreign policy paints a picture of the House's bungling on foreign policy issues that tends to evoke much incredulity

---

[2] 50 U.S.C. §§ 1541–1548 (1976); *see* Note, *Policing Executive Adventurism: Congressional Oversight of Military and Paramilitary Operations*, 19 HARV. J. ON LEGIS. 327 (1982).

and embarrassment. The author recounts the transgressions of the well intentioned reformers on issues ranging from support for the World Bank to the arms embargo of Turkey.

Whalen contends that, with the exception of the human rights area, attempts to redirect United States foreign policy through the floor amendment process largely have been designed to "impose limitations on a foreign policy agenda already established and submitted to the Congress by the executive branch" (p. 127). Moreover, Whalen sees the reforms as exacerbating the most negative results of the Mayhew thesis: that members of Congress vote and act in accordance with their conception of self-interest, which in most cases means that they vote or posture in such a way as to ensure their continuation in office.[3] House members look to foreign policy issues not as chances to advance national interests, but either as opportunities for the advancement of their electoral prospects or as potential pitfalls that could have a dramatic impact on their chances for reelection.

While suggesting that most of the democratic aspects of the House reforms should remain intact, the author reviews remedies proposed by others for the problems that the House reforms have created. In a laundry list approach, Whalen considers some of the many proposals for reform, but rejects most (pp. 152–62). He cites two serious flaws of these recommendations for a "reform of the reforms," specifically that they disregard members' self-interests and do not attack the mechanisms that create potential campaign problems (pp. 162–63).

Whalen then proceeds to offer his own suggestions for reform. His proposals include two rule changes designed to restrict the introduction of noncommittee amendments and to increase the use of the modified open rule that forbids the introduction of an amendment on the floor unless a rule permits its consideration.

Whalen's suggested reforms, however, do not address the preoccupation of individual House members with matters of reelection and power enhancement. This reviewer believes that some form of partisan reform, aimed at mitigating the explicit electoral connection that has developed between the constituent and the representative, should be considered. The creation, for example, of strong central party organs, with sufficient campaign war chests to command their candidates' loyalty, might

[3]DAVID R. MAYHEW, CONGRESS: THE ELECTORAL CONNECTION 5–6, 49–65 (1974).

bring about a more unitary foreign policy by developing well-articulated positions for party candidates to endorse.

The difficulties of reversing the public's expectations that representatives should take positions—even when those positions run counter to perceived national interests—are exacerbated by the current realities of the international economic system. In the post-war era, the Western nations and Japan have attempted to promote the goal of a liberal world trading order by seeking to reduce and eventually to eliminate barriers to international trade. The high level of interdependence among trading partners, however, heightens each trading partner's sensitivity to the economic conditions in other countries. When a country's economy suffers from unemployment, recession, or inflation, beggar-thy-neighbor economic policies that seek to blame a foreign competitor and to export the problem offer an easy and an often politically expedient way of dealing with domestic economic difficulties. While these policies seem politically wise, they are extremely disruptive to the development of a stable world economic order and to the formation of a rational and effective foreign policy.

Although the reforms that Dr. Whalen suggests might eliminate the more narrowly motivated initiatives of a particular interest group, the constituencies of the broader industrial sectors, such as the steel, textile, automobile, and leather goods sectors, could exert tremendous pressure for economic protectionism by seeking the introduction of noncommittee amendments to further their economic interests in the face of a hostile committee or House leadership. Whalen's recommendations seem to be directed at a Congress which is static and unaffected by the increasing pressure in the area of international trade. Economic protectionism already has assumed a major position in the campaign rhetoric of at least one of the leading candidates for the 1984 Democratic presidential nomination, Walter Mondale, and pressure has been exerted for years by the automobile and the steel industries for the United States to "get tough" on Japan for flooding the United States with automobiles and for "dumping" surplus steel in United States markets in violation of existing agreements.

Whalen also should account for the conflict between the Congress and the Executive Branch, and for the growth of executive control over foreign policy since the end of the Second World War. While there may have been some brief reversals or inter-

ruptions in that trend, many writers seem to agree that the President is more free than ever to commit the United States to foreign military encounters.[4] Much of this continued growth in presidential war-making power is a necessary response to the need for secrecy, the pace of world events, and the inability of Congress to act quickly. The War Powers Resolution has been one of the casualties of this growth in executive power. Yet Congress itself is at least as responsible as the Executive Branch for the resolution's decreasing utility. Congress has chosen to insist on the observance of the WPR's provisions only on a case-by-case basis. When actions taken by the President have proven to be popular with the public, such as in some stages of the United States' involvement in Southeast Asia, Congress has not insisted on adherence to the Resolution.

This precedent of selective enforcement by Congress and selective observance by the President has rendered the War Powers Resolution ineffective, leading one to wonder if the House will be properly equipped to confront the President if necessary. Certainly, the Congress now has greater staff and technical resources at its disposal to assemble the necessary information to confront the President, but one only can speculate as to how effective the Congress would be in the absence of a workable War Powers Resolution. Senator Thomas Eagleton (D-Mo.), one of the original cosponsors of the Resolution, is among those who now find the Resolution ineffective. To his credit, Eagleton foresaw the problems when his original measure was amended to its present form, and voted against the final version of the War Powers Resolution.[5]

The turmoil in El Salvador provides a good illustration of the current congressional inability to influence the President's conduct of foreign policy. The International Security and Development Cooperation Act of 1981[6] contains provisions that require the President to certify that El Salvador has made significant progress in five specific areas before additional military aid is sent to the regime in that country.[7] The certification

---

[4]Works that deal with the expansion of presidential power in the post-World War II period include: E.S. CORWIN, THE PRESIDENT, OFFICE AND POWERS, 1787–1957 (1957); E.C. HARGROVE, THE POWER OF THE MODERN PRESIDENCY (1974); R. NEUSTADT, PRESIDENTIAL POWER (1980); and A.M. SCHLESINGER, JR., THE IMPERIAL PRESIDENCY (1973).

[5]119 CONG. REC. 25,079 — 86 (1973).

[6]Pub. L. No. 97-113, 95 Stat. 1519 (1981).

[7]*Id.* § 728, 95 Stat. 1555; *see* note under 22 U.S.C. § 2370 (Supp. IV 1981).

provision calls for a twice annual review by the Executive Branch.[8] Yet despite gross violations of basic human rights, the nonexistence of a land reform program, and the lack of progress in prosecuting those people who are responsible for the deaths of American nuns and trade unionists, the Reagan administration continues to certify El Salvador as eligible for the receipt of military supplies.

While Dr. Whalen's suggested revisions of House rules are meritorious, they do not deal with the problems facing the present Congress. The Congress needs to enhance its ability to confront the Executive Branch in a meaningful way. The reforms recommended by Whalen seem designed to deal with the problems of previous Congresses, and fail to take account of the continued growth of presidential power in foreign policy, which threatens to overwhelm the ability of Congress to deal with foreign policy in a responsible way.

---

[8] *Id.* § 728(b), 95 Stat. 1555.

RECENT PUBLICATIONS

ECONOMIC EFFECTS OF SOCIAL SECURITY. By *Henry J. Aaron*. Washington, D.C.: The Brookings Institution, 1982. Pp. xii, 84. $12.95 cloth, $5.95 paper.

For decades, economists have tried to reach agreement on the effects of social security on personal savings and on the supply of labor. Henry Aaron, in *Economic Effects of Social Security*, argues that they have failed: "[E]conomists may be found who argue that social security has decreased savings, increased savings, or had no perceptible effect. Similarly, economists may be found who hold that social security has increased labor supply, reduced it, left it unchanged, or caused offsetting changes in labor supply by workers of different ages" (p. 2). Aaron, an expert on social security who has been intimately involved in recent efforts to modify the system, believes that because of this basic confusion, proposals to modify the system should not be judged by their predicted economic effects (p. 2), but rather by decisionmakers' "perceptions of equity, adequacy of benefits, fairness of taxes, and similar qualitative considerations" (p. 2).

First, Aaron discusses concern of policymakers that social security taxes create disincentives for private saving, thereby retarding growth of the nation's capital stock and thus reducing output and the real wage level. This fear, however, is based on two assumptions that Aaron argues are unsubstantiated. First, the effects of social security on saving cannot be readily determined because such a determination requires an estimate of social security wealth (the per cent value of expected future benefits), but there is too great a range of plausible estimates to warrant any conclusion. Second, the precise relationship between changes in national savings and the level of national capital formation is unknown. He notes that in a perfectly integrated international capital market, a substantial proportion of any increment in national savings will finance foreign capital formation. Moreover, Aaron argues that the most serious impediment to increased national savings is government dissaving, in the form of the federal deficit. Aaron asserts, therefore, that the best approach for increasing national savings is to reduce the federal deficit.

Aaron then analyzes the fear of policymakers that social security tax decreases the labor supply at given wage levels,

thereby decreasing aggregate output and increasing the price level. Aaron emphasizes that the social security system's overall effects on labor supply are not yet known. He recognizes that the system does play an important role in determining the age at which workers retire, but he contends that policymakers must look "not only on the effects of social security, but also on those of private pensions, general labor policy, and controllable labor market conditions" (p. 66).

Aaron concludes that economists must overcome the complexities of the system, the deficiencies of the existing data, and the theoretical difficulties of developing a model before they can reach a consensus on the effects of social security on savings and labor supply. Until a deeper understanding is achieved, analysts' findings will rest on oversimplified assumptions, including ones concerning the length of individuals' planning horizons and the quantity of information individuals rely on in making their economic decisions.

Aaron asserts that the inadequacy of data on the effects of social security stems from the lack of maturity of the system, the numerous modifications and additions that have been adopted since the system was created in 1935, and the uncertainty concerning prospective modifications. Although he believes that increased observation of the social security system will improve the estimates of its effects on the supply of labor, Aaron argues that the estimates of the system's effect on savings will remain inadequate. He notes that precise estimates of the system's effect on savings would require the development of a model incorporating the effects of "the diverse behavior of many people in quite different stages of life and economic circumstances, with different levels of education, and from different cultural backgrounds" (p. 10).

By providing a critical survey of the technical inquiries into the economic effects of social security, Henry Aaron's book can guide both the knowledgeable policymaker and the reader lacking an economic theory or statistics background. The book hopefully will encourage informed public discourse on this important issue.

Douglas Lister

THE PLIGHT OF THE THRIFT INSTITUTIONS. By *Andrew S. Carron.* Washington, D.C.: The Brookings Institution, 1982. Pp. viii, 89, appendix, index. $10.95 cloth, $4.95 paper.

In view of the rapid changes in the relevant areas of economics and law, any book about the current state of financial institutions runs the risk of being outdated before it can be circulated. Although recent legislation has caused some of Andrew S. Carron's suggestions in *The Plight of the Thrift Institutions* to become obsolete, the book contains analysis that continues to be useful.

The difficulties of the thrift industry—which includes savings and loans, mutual savings banks, and credit unions—have resulted from governmental regulation that required thrifts to violate two basic principles of financial management. First, by investing almost wholly in residential mortgage loans, thrifts violate the principle that assets should be diversified to reduce market risk. Second, by funding these long-term, primarily fixed-rate, mortgage assets through short-term savings account liabilities, thrifts violate the principle that assets and liabilities should be matched in risk and maturity.

Financial difficulties did not occur during times of low, steady interest rates, but problems have developed in recent years. While thrifts' mortgage asset income is locked into the old, low interest rates, thrifts are losing their low-cost source of funds as millions of people convert low-interest savings deposit accounts into high-interest instruments, such as money market funds. The result is staggering: in 1981, the market-value net worth of thrifts was a *negative* $44 billion (p. 19 table 1-4). The threat to the thrifts also represents a grave danger to the troubled housing market, a huge exposure to loss in the state and federal thrift institution deposit insurance systems, and a decline in public confidence in financial institutions as a whole.

Carron resists the temptation, however, to dwell on the crisis aspects of the thrifts' situation, noting that there is no danger of wholesale failure due to the thrifts' illiquidity in the immediate future (p. 19). Instead, he emphasizes the chronic problems of the industry and the need for a coherent transition policy that will ease thrifts into the modern financial world and leave them with the flexibility they need to adapt to a rapidly changing environment. He relies in part on an economic model of the

industry (pp. 27–38, app.), which is a laudable effort to provide an objective basis for testing policy decisions. Because the model's projections extend only through 1983, however, they are of little utility.

Carron suggests that the goals of any program of governmental response should include: first, aid to specific ailing thrift institutions, and, second, modification of the powers and structure of the thrift industry in order to restore practical, efficient intermediation to the market for home mortgages (p. 50).

Since Carron wrote this book, the Garn-St. Germain Depository Institutions Act of 1982 (DIA) has been enacted in an effort to meet these goals.[1] Carron testified at the committee hearings for the bill,[2] and his analysis in the book is useful in predicting whether the current effort to assist the thrift industry will be successful.

Aid to failing institutions. Carron endorses a program of direct cash grants to failing thrifts as "potentially the most efficient and effective" solution to the plight of the thrift institution (p. 51). Although the Reagan administration has ruled out a direct cash grant program (p. 51), the DIA includes provisions that vastly increase the powers of the Federal Deposit Insurance Corporation (FDIC) and the Federal Savings and Loan Insurance Corporation (FSLIC) to provide aid to failing institutions. These powers include making loans to, purchasing assets and liabilities of, and making contributions to failing institutions.[3] The DIA also provides a novel type of assistance to thrifts in the form of "capital instruments," debt-like equity securities issued by the thrift to the FSLIC, on terms very favorable to the thrift. These instruments are designed to increase the net worth of the thrift while providing temporary funds to meet its liquidity needs.[4]

This program of indirect aid suffers from the flaw of hiding the extent of governmental subsidies to thrifts. Nevertheless, the program carries out Carron's major recommendations in this area (pp. 86–87), and probably will help to avoid liquidity crises in individual institutions during the transition of the industry.

---

[1]Pub. L. No. 97-320, 96 Stat. 1469 (1982).

[2]S. REP. NO. 536, 97th Cong., 2d Sess. 14 - 15, *reprinted in* 1982 U.S. CODE CONG. & AD. NEWS 3054, 3068 - 69.

[3]DIA §§ 111, 122 (to be codified at 12 U.S.C. §§ 1823(c), 1729(f), respectively).

[4]*Id.* §§ 201 - 203 (to be codified in scattered sections of 12 U.S.C.); *see* S. REP. NO. 536, 97th Cong., 2d Sess., 9 - 13, *reprinted in* 1982 U.S. CODE CONG. & AD. NEWS at 3063 - 66.

The DIA also permits the FDIC and the FSLIC to facilitate mergers of failing institutions with stronger institutions by sweetening the deal with guarantees, loans, or other assistance to the acquirer.[5] There have been many such mergers recently: the FSLIC arranged sixty-one mergers in 1981, and 105 in the first seven months of 1982.[6] As Carron points out, however, the merger trend is self-limited because there are relatively few institutions with the financial strength to absorb troubled thrifts and banks (p. 70).

The significance of the enhanced merger power lies instead in the DIA's explicit permitting of interstate mergers to save a failing institution.[7] This is a marked departure from prior law, which had maintained the vast majority of banks and thrifts as single-state institutions. This may signal the beginning of a movement to interstate banking,[8] long sought by many observers.[9]

Carron criticizes the introduction of such massive changes in the structure of financial markets to solve an essentially temporary problem (p. 71). However, because one of the central problems of the thrift industry has been regulation-enforced inability to compete with nationwide financial institutions, it seems appropriate to relax state restrictions where the result will meet the goals of both aiding troubled thrifts and aiding the deregulation and normalization of this segment of the financial market.

Restructuring the thrift industry. In order to bring thrifts into a more competitive position, the DIA allows them to offer services and to make investments like those of their competitors. Thrifts now may offer checking accounts to businesses, may place a large part of their assets in nonresidential secured loans, and may enjoy increased participation in the commercial loan market.[10] Although these new powers may be useful in the long term to help thrifts balance the risk and maturity of their assets with those of their liabilities, they have few short-term advan-

---

[5]DIA § 123(a) (to be codified at 12 U.S.C. § 1730a(m)).

[6]S. REP. NO. 536, 97th Cong., 2d Sess. 4, *reprinted in* 1982 U.S. CODE CONG. & AD. NEWS at 3058.

[7]DIA § 123(a) (to be codified at 12 U.S.C. § 1730a(m)).

[8]There is a three-year sunset provision for this amendment. DIA § 141 (to be codified in scattered sections of 12 U.S.C.).

[9]*See, e.g.*, U.S. DEP'T OF THE TREASURY, THE REPORT OF THE PRESIDENT: GEOGRAPHIC RESTRICTIONS ON COMMERCIAL BANKING IN THE UNITED STATES (1981).

[10]DIA §§ 312, 321 - 330 (to be codified at 12 U.S.C. §§ 1464(b), 1464(c), and 3503).

tages. Thrifts will have neither the expertise nor the financial strength in the near future to embark on new programs of investment in areas in which banks are entrenched competitors (p. 66).

The DIA, in an unprecedented circumvention of limitations on interest rates to be paid on consumers' deposit accounts, authorizes a new federally insured market-rate account for both thrifts and banks that competes directly with money market funds.[11] Carron criticizes this development, concluding that it is unlikely to help thrifts (pp. 60–61). In the short run, this measure may in fact do substantial harm to the liquidity position of thrifts. First, although the market-rate account is designed to attract funds to the cash-short thrifts, the funds are expensive relative to funds from savings and checking accounts. In addition, the cost of federal insurance combined with the relatively small size of market-rate accounts puts them at a cost disadvantage in relation to money market funds.

Second, it is not clear that the market-rate accounts will attract and keep either new investors or current patrons of the money market funds. Market-rate accounts initially attracted investors from money market funds by offering an interest rate up to three percentage points higher than the average for money market funds. The gap has now narrowed to less than one percentage point.[12] In addition, market-rate accounts carry more restrictions on the number of transactions the consumer can make. The accounts offer the advantage of insurance, but it is questionable whether consumers will see the accounts as significantly more secure than funds such as those offered by Merrill Lynch or Dreyfus. Thrifts and banks also offer the convenience of serving other financial needs (p. 65), but the new accounts most likely would attract the current customers of the thrift or bank. Thus, customers may simply transfer their low-yield deposit account funds to high-yield market-rate accounts, causing the institution to incur greater costs.

For these reasons, although the new powers of thrifts represent a leap forward in removing restrictive regulation—a major cause of the thrifts' problems—these abilities may be practically unusable or harmful to thrifts' liquidity positions. This is unavoidable: if thrifts are to be brought into the modern financial

---

[11]*Id.* § 327 (to be codified at 12 U.S.C. § 3503).

[12]Bennett, *Money Funds: Rate War Over*, N.Y. Times, Mar. 16, 1983, at D1, col. 3.

era, sacrifices will have to be made both by the government, which is largely responsible for the present predicament, and by the owners of "nonviable" thrifts, who have taken a business gamble and lost (p. 88). Ultimately, however, taxpayers must bear the greatest burden, that of protecting depositors from loss through federal insurance.

Kathryn E. Hoff-Patrinos

SUGGESTED STATE LEGISLATION 1983. Edited by *The Committee on Suggested State Legislation.* Lexington, Ky.: The Council of State Governments, 1982. Pp. x, 335, cumulative index. $15.00 paper.

Suggested State Legislation is a collection of topical and concrete legislative proposals designed by and for state officials. The proposals are compiled by the Committee on Suggested State Legislation, an organ of the Council of State Governments. The Council is organized to assist state officials as a research and information service, a lobbying agent in Congress; and a forum for regional cooperation. The Committee is composed of 150 state officials appointed by their respective governments. The Committee solicits legislative proposals from state officials and legislators, federal agencies, and private and public interest groups.

The proposals are intended to be a source of "both legislative ideas and drafting assistance for state government officials" (p. ix). The compilation substantially achieves the latter goal: the proposals are in complete statutory form, containing titles and definitions as well as the substantive provisions, and are divided into internally cross-referenced and consistent sections. Unfortunately, the collection has several omissions that reduce, but do not eliminate, its value as a source of new legislative ideas. The proposals contain sparse, if any, accompanying commentary. Each statute is introduced by only one or two brief paragraphs that merely summarize the provisions rather than provide guidance on the nature of the problem or on how the language should be interpreted to implement properly the intent of the proposal. Moreover, although each statute does note the state, agency, or group that proposed or enacted it, the Com-

mittee makes further research difficult by failing to provide specific citations to any existing proposals or enactments.

The 1983 edition contains over forty proposals. Most of the proposals concern topics that are timely and useful, although some, such as the Insect Sting Emergency Treatment Act and the Mold and Die Retention Act, deal with topics of very limited interest. Of particular importance are proposals for the economic regulation of electronic banking, the restriction of child pornography, the prevention of private paramilitary training oriented toward violence, and the protection of migrant farm workers against unsafe working conditions and abusive employers. This volume also includes a cumulative index by subject to all proposals that have been published since the first issue in 1941.

This compilation serves two important functions. First, it highlights policy problems that are particularly suited to legislative solutions. Second, it provides a detailed basis on which to develop and compare specific solutions. *Suggested State Legislation* thus should be welcome on the desks of not only state legislators and their staffs, but of all persons who are closely involved in the development or implementation of new state legislation.

Thomas McCord

THE NEW DEAL LAWYERS. By *Peter H. Irons*. Princeton, N.J.: Princeton University Press, 1982. Pp. 351, notes, index. $19.50 cloth.

Franklin D. Roosevelt and the New Deal Congress of 1933 to 1937, in implementing an unprecedented program of industrial and agricultural reform, gave broad discretion to regulate the national economy to new agencies such as the National Reconstruction Administration (NRA), the Agricultural Adjustment Administration (AAA), and the National Labor Relations Board (NLRB). These agencies were authorized to regulate by methods previously not attempted or used only by the states. Peter H. Irons, in *The New Deal Lawyers*, presents a rich account of how these agencies were defended by the New Deal lawyers against constitutional challenges before a federal judiciary dominated at all levels by conservative Republican appointees hostile to the New Deal. These young lawyers were responsible for

developing the particular arguments and theories that shaped the constitutional battle that changed the future role of the government in economic affairs.

Irons, relying on litigation files in the National Archives and interviews with New Deal lawyers, including Paul Freund, Thomas Corcoran, Charles Fahy, and Milton Katz, explores the variety of conflicts that faced the lawyers as they developed litigation strategies to test the constitutionality of the New Deal programs. Irons describes the unique style of each agency, characterizing the lawyers as "legal politicians" (NRA), "legal reformers" (AAA), and "legal craftsmen" (NLRB) (pp. 5–6).

To a great extent, the style of each agency reflected the approach of the General Counsel for each group of lawyers. Donald Richberg, who headed the legal politicians of the NRA, had extensive access to the White House, and used his relationship with Roosevelt to assume broad administrative responsibilities. Jerome Frank, who led the legal reformers of the AAA, was an important contributor to the legal realist movement, which advocated the use of administrative agencies to regulate the national economy. Finally, Charles Fahy, who led the legal craftsmen of the NLRB, was a methodical lawyer who viewed social and political reform as the province of politicians rather than lawyers. The agencies utilized these different approaches in facing similar conflicts with the Department of Justice, the administrators of their respective agencies, and the federal judiciary.

The legal politicians of the NRA and the legal reformers of the AAA were not successful in defending the constitutionality of their agencies because of two sets of conflicts. Conflicts developed between the lawyers and their respective agencies over the appropriate role of the lawyers in policymaking. Intense jurisdictional battles also raged with the Department of Justice over whether the lawyers of the Department or those at the agencies should control the litigation of the cases.

These conflicts severely hampered the ability of the NRA lawyers to develop and implement an effective litigation strategy. The NRA instituted actions against small manufacturers for violations of trade codes negotiated between the NRA and industry. These cases were ill-suited to test the agency's power to regulate under the Interstate Commerce Clause because the focus on specific types of small violators, such as gasoline stations and meat merchants, offended both the liberals' bias

against concentration and the conservatives' hostility to regulation. The result was a unanimous rejection of the legislation by the Supreme Court in *Schechter Poultry Corp. v. United States*.[1]

At the AAA, the legal reformers almost immediately ran into conflict with the conservative leadership of the agency over the treatment of tenant farmers in the AAA program to reduce crop acreage in the South. The AAA's failure to formulate effective litigation strategies resulted in a divided Supreme Court holding the AAA unconstitutional in *United States v. Butler*.[2]

The NLRB's legal craftsmen fared much better. This was due in part, according to Irons, to the NLRB lawyers learning from the mistakes of the NRA and AAA (p. 240). Moreover, the Wagner Act,[3] establishing the NLRB, was drafted with the experience of the NRA in mind (p. 227). The NLRB avoided conflict with the Department of Justice because the Act gave it the power to litigate its cases independently. The Act also carefully delineated administrative procedures to be followed, using the Federal Trade Commission as a model. The eventual success of the NLRB, according to Irons, was also due to the carefully developed litigation strategy designed by Fahy and his legal craftsmen (p. 243). The lawyers litigated the cases most clearly involving interstate commerce and anticipated the cases that would offer the most difficulty. In addition, the NLRB benefited from the growing public outcry over the *Schechter* and *Butler* decisions and Roosevelt's threatened Supreme Court-packing scheme. The agency's success was reflected in an NLRB victory upholding its constitutionality in *NLRB v. Jones & Laughlin*.[4]

Irons also recognizes and explores some of the problems stemming from the varied personalities within each agency. He notes, for example, the dilemma facing liberal young lawyers, influenced by decentralists such as Brandeis and Frankfurter, in the position of defending the constitutionality of the NRA and AAA trade codes that effectively allowed industries to restrain competition and fix prices.

In the past, most of the attention paid to the New Deal era has focused on the Supreme Court and on the role of political

---

[1]295 U.S. 495 (1935).
[2]297 U.S. 1 (1936).
[3]National Labor Relations (Wagner) Act of 1935, Pub. L. No. 74-198, ch. 372, 49 Stat. 449 (currently codified at 29 U.S.C. §§ 151 — 167 (1979).
[4]301 U.S. 1 (1937).

events during the early New Deal. Yet, as Irons clearly illustrates, the shaping of the constitutional debate was largely the result of various litigation strategies and styles used by the agency lawyers. By focusing on the strategies and styles of the agencies' lawyers, *The New Deal Lawyers* makes an important contribution to our understanding of how constitutional battles are fought and how government lawyers helped to shape the regulatory state which exists today.

Charles A. Blanchard

THE ONCE AND FUTURE DEMOCRATS: STRATEGIES FOR CHANGE. By *Paul Simon*. New York, N.Y.: The Continuum Publishing Company, 1982. Pp. ix, 169, notes, index. $12.95 cloth.

Writing before the 1982 mid-term elections, Paul Simon, presently a fifth-term Democratic Congressman from Illinois, argues in *The Once and Future Democrats* that the future of his party can be bright. Simon does not base his optimism upon the nation's negative reaction to high unemployment, an influential factor in the 1982 elections. Instead, he believes that the Democrats must present the voters with a comprehensive alternative to the Reagan program. Although Simon does not fully address some of the problems inherent in his fiscal and political strategy, his book provides a comprehensive Democratic alternative.

Simon premises this Democratic alternative, however, on a balanced budget, a basic tenet of Republican fiscal conservatism. He emphasizes that the third largest federal expenditure is for interest on the national debt, and that deficit spending is inflationary.

Simon's first domestic proposal is a jobs program, which is particularly noteworthy because he has recently introduced it in the House.[1] The proposal calls for a project-oriented program like the Works Progress Administration (WPA), which Simon believes can provide up to three million jobs for the chronically unemployed. The program would provide work for those who have been unemployed for at least thirty days and who have a

[1]H.R. 777, 98th Cong., 1st Sess. (1983).

total family income of less than $15,000. The government would pay the minimum wage and would require those whom it employs to work a four-day week and to devote the fifth day to finding other employment. Simon estimates that the proposal would cost fifteen billion dollars if it were initiated immediately on a full-scale basis, but suggests that the costs should be deferred by gradually phasing in the program.

A second set of proposals concerns health care. One proposal endorses a significant expansion of Medicare coverage. Another proposal sets forth a catastrophic medical expenses program under which the federal government would reimburse a family for ninety per cent of all medical expenses in excess of twenty per cent of the family's annual income. Simon estimates that implementation of both proposals would cost twenty-two billion dollars. He suggests that the government should fund these proposals either by a gradual phase-in, or by a "slight tax increase" (p. 62).

A third area that Simon addresses is industrial modernization. His first proposal is to provide additional tax deductions for corporate expenditures on research. Simon's second proposal is to create an agency similar to the New Deal's Reconstruction Finance Corporation. This joint government-business agency would have loan-making powers that would be used to encourage growth in targeted industries and to aid in the transition of industries eliminated by foreign competition or by domestic economic conditions.

A fourth set of proposals focuses on the nation's energy needs. Simon's proposals urge the federal government to develop an emergency energy program, to promote a major national conservation drive, to increase waste conversion programs, to form a quasi-government corporation to bid for foreign oil, to fund coal-acid rain research, to provide incentives to utilities to promote solar energy, and to increase investigation into fusion research.

In his fifth set of proposals, Simon suggests ways of achieving savings in the defense budget. The nation can achieve defense savings, he asserts, because its real defense needs are in conventional areas, improvements in which are less costly than the Reagan nuclear defense proposals. He therefore proposes a shift toward more easily operated, less expensive conventional weapons. Simon also proposes a peace time draft, which he states "is coming, regardless of who promises what" (p. 136), due to

a predicted twenty-five per cent decrease in the number of people of recruitment age within the next decade. Simon's proposal avoids the Vietnam draft's socially inequitable deferments by requiring all Americans at age eighteen, or following high school graduation, to choose between one year of service in either the military or a community program, such as VISTA. Substantial educational benefits after service would provide an incentive to choose the military.

Although Simon suggests that these defense savings would offset the costs of his proposed domestic programs, expenditures both for his new proposals and for existing federal programs probably would cause deficits as large as the Reagan administration's. The only difference would be that more money would be spent on social programs and less on defense. Such deficits are inconsistent with Simon's initial premise of a balanced budget, and lead to the suspicion that he implicitly endorses significant tax increases. A clearer statement about the need for future tax increases would improve the book's overall fiscal consistency.

Simon's other proposals do provide a comprehensive solution to many of the nation's problems. His proposed energy guidelines would respond to the inevitable consumption of current oil surpluses. His Reconstruction Finance Corporation concept, with its emphasis on long-term planning, would enhance prospects for a sustainable recovery. Simon's jobs proposal, with its reliance on project creation, would avoid the tendency of programs such as CETA to substitute their workers for those who otherwise would have been employed. In addition, his deductions for research and development would stimulate productivity more effectively than the Reagan tax cuts because only the deductions would directly target businesses.

Simon's proposed programs, however, directly conflict with the Reagan philosophy of reducing intervention by the federal government. Simon recognizes this noninterventionist trend, and suggests a political strategy of emphasizing the benefits of his proposals to the voters. But this strategy of simply communicating the positive aspects of the Democratic proposals does not address a significant factor: the noninterventionist claim of "getting government off the people's backs" has strong populist appeal. Simon should challenge this populist claim by pointing to the Reagan administration's antitrust and environmental policies. He might argue that such policies actually get

government "off the backs" of large corporations and the affluent at the expense of middle and lower class individuals.

Although a challenge to the noninterventionist populist claim could improve this book's political strategy, and a clearer statement about the need for tax increases to balance the budget could rectify its fiscal inconsistency, *The Once and Future Democrats*, nonetheless, is a success in two respects. It provides a comprehensive Democratic alternative to the Reagan administration's programs and a thoughtful, nonpartisan overview of possible future campaign issues and legislation. As such, Simon's book is a valuable reference for both the politician and the political scientist.

Seth Mones

AMERICAN POLITICS AND PUBLIC POLICY: SEVEN CASE STUDIES. Edited by *Allan P. Sindler.* Washington, D.C.: Congressional Quarterly Press, 1982. Pp. ix, 256. $8.95 paper.

The issue analyses in *American Politics and Public Policy: Seven Case Studies* are presented in an effort to define and to critique the interaction between political interests and the making of public policy. According to the editor, Allan Sindler, the accounts "demonstrate the American political system's strengths and weaknesses and capacity to formulate and carry out public policy" (p. vi).

The book is primarily designed as a supplementary text for a traditional American government and public policy course. Sindler organizes the book into seven sections: Presidency, Congress, bureaucracy, courts, political parties, domestic policy, and defense policy. The essays fall into two basic categories that are distinct from this imposed organization. Four of the essays focus on the nitty-gritty detail of how Congress passes bills or agencies develop regulations. These studies deal with the passage of the 1981 Budget Resolution and Reconciliation Act, the efforts by the Food and Drug Administration to require druggists to provide information leaflets for prescription drugs, the initiatives to enact oil price decontrols and a windfall profits tax, and the development of the MX missile. The remaining three studies are more "issue-based," focusing on arguments

concerning a particular issue, or presenting a historical overview of a particular policy. These include a piece by Sindler analyzing the advantages and disadvantages of the electoral college and direct national election systems, a study tracing the changes in public policy on abortion as reflected in court decisions and legislative enactments, and an analysis of the American political party as it has evolved from a "political machine" structure to an "issue-oriented" organization.

Three of the case studies are of particular interest. Bruce Oppenheimer's "Resolving the Oil Pricing Issue: A Key to National Energy Policy," one of the four process-oriented studies, is one of the best contributions. Oppenheimer traces four major presidential and congressional efforts to develop a "coherent" oil pricing policy. Through his analysis of the legislation introduced from 1974 to 1980, Oppenheimer conveys to the reader an excellent sense of the tortuous passage of bills as they suffer from successive attacks and modifications in subcommittees and committees, on the House and Senate floors, and in conference committees, and finally, from the risk of disapproval in the Oval Office. His essay reflects a sensitivity to the role that individual members of Congress play in aiding or hindering a bill, and the effect that public opinion and a sense of urgency have on the ultimate success of any public policy.

Steven Smith's essay on the 1981 budget battles, "Budget Battles of 1981: The Role of the Majority Party Leadership," is similar to Oppenheimer's piece in its emphasis on the "hardball" realities of congressional processes. Smith presents a concise account of the chronological events in the 1981 budget process, exploring the public and political pressures existing at each stage. Smith focuses on why the Democratic majority party leadership was unable to counteract effectively the Reagan "steamroller" and to pass its own budget resolution and reconciliation act. Smith's analysis benefits from his experience on the Democratic Policy and Steering Committee, but lacks a sufficient understanding of the types of pressures placed on individual legislators. There is insufficient analysis of the effect of constituent pressures, media exposure, and personal ideology, which caused members of Congress to react in distinct, "atomized" modes throughout the budget battles. Nevertheless, Smith presents a useful and integrated overall picture of the budget process, including an analysis of the critical role the budget process presently plays in shaping legislation.

Austin Sarat's piece focusing on abortion policy, "Abortion and the Courts: Uncertain Boundaries of Law and Politics," one of the three "issue-based" essays, is valuable in its analysis of the proper judicial role in shaping public policy. Sarat views the courts as integral, valid players in the political process who, through the results of their decisions, "rearrange the stage on which political conflict is played out" (p. 121). Sarat's comprehensive review of the changes in abortion policy is designed to demonstrate how players in political conflicts legitimately move from one branch and level of government to another in the search for the most advantageous resolution. Sarat traces the development of abortion policy from the pre-1973 efforts to liberalize state abortion laws to the effect of the Supreme Court's 1973 decision legalizing abortion. He then follows the subsequent efforts of state and national legislatures to restrict the reach of the new constitutional right of abortion and concludes with the recent Supreme Court decisions that uphold such restrictions. Sarat's study serves as an effective vehicle for his thesis that it is appropriate for the courts to take an active role in the political process, as well as a useful compendium of the changes that have occurred in abortion policy in recent decades.

The collection achieves its primary goal of providing students with a sense of the interplay between political actors and institutions and the making of public policy. Most of the essays capture the dynamic interdependence of political and financial pressures, personal interests, and political party goals. Students perhaps could receive a similar sense of the dynamics of the political process through reading a compilation of well-written *Washington Post* stories or the *Congressional Quarterly Weekly Report* articles on a particular policy issue. This collection is useful, however, in its translation of daily-coverage stories into case studies. First, the effort legitimizes, as a valid object of academic study, the ongoing personal and political dynamics that do shape public policy. Second, as accounts written after the fact, the case studies often draw on candid interviews with participants that enhance analysis of past events. In addition, each case study increases accessibility to a particular policy issue by compiling the relevant background information in one piece of writing.

Appreciation of *American Politics and Public Policy*, however, does not have to be restricted to students. Although the

case studies will not provide many new insights on the "process" of politics to those who have worked in Congress or the federal bureaucracy, they are concise and readable accounts of the politics underlying a number of recent, interesting policy issues. As "story-telling," some of the essays are captivating. Thus, for the student seeking an initial exposure to the way the American political process shapes public policy, or for the hardened "politico" seeking an interesting overview of a particular policy issue, *American Politics and Public Policy* can provide worthwhile and enjoyable reading.

Chai Feldblum