// HARVARD JOURNAL

on

LEGISLATION

VOLUME 18 WINTER 1981 NUMBER 1

ARTICLE

RESTRAINING THE REGULATORS: LEGAL PERSPECTIVES ON A REGULATORY-BUDGET FOR FEDERAL AGENCIES
Lance D. Wood, Elliott P. Laws & Barry Breen 1

NOTES

PROTECTING PRIVATE EMPLOYEES' FREEDOM OF POLITICAL SPEECH
Mark T. Carroll 35

THE SOFT DRINK INTERBRAND COMPETITION ACT OF 1980: ANTITRUST LOSES ITS FIZZ
Leonard R. Stein 91

RESTRICTIVE MEDICAL MALPRACTICE COMPENSATION SCHEMES: A CONSTITUTIONAL "QUID PRO QUO" ANALYSIS TO SAFEGUARD INDIVIDUAL LIBERTIES
Howard Alan Learner 143

FEDERAL PROTECTION OF EMPLOYMENT RECORD PRIVACY
Thomas E. Reinert, Jr. 207

BOOK REVIEW

Hough, SOVIET LEADERSHIP IN TRANSITION
Harold J. Berman 253

RECENT PUBLICATIONS 259

BOOKS RECEIVED 263

Published three times during the academic year (Winter, Spring, and Summer) by the Harvard Legislative Research Bureau, Langdell Hall, Harvard Law School, Cambridge, Mass. 02138. ISSN 0017-808x.
Third person pronouns, wherever used, refer to both men and women.
Subscriptions per year: United States, $7.50 (single copy, $4.00); foreign, $9.00 (single copy, $4.50). Subscriptions are automatically renewed unless a request for discontinuance is received.
Back issues prior to the current volume are available from Fred B. Rothman & Co., 10368 W. Centennial Road, Littleton, Colorado 80123, at $6.50, check with order. For prices of complete back volumes or of sets, please inquire of Fred B. Rothman & Co.
Copyright © 1981 by the Harvard Legislative Research Bureau

HARVARD JOURNAL ON LEGISLATION

President
MICHAEL J. O'MALLEY

Article Editor
JODY E. FORCHHEIMER

Note Editor
PETER A. GLICKLICH

Managing Editor
DONALD J. KUNZ

Research Editor
ELIZABETH S. STONG

Congress Editor
GEORGE H. WHITE

Executive Editor
CARLTON M. SMITH

Book Review Editor
DEBORAH G. SHULEVITZ

Bureau Director
JEFFREY G. BERMAN

Asst. Bureau Director
RUTH A. BOURQUIN

Board of Editors
MITCHELL L. BERG
SUSAN J. DALEY
THOMAS G. HAMERLINCK
ELLIOT K. GORDON
RICHARD M. GRAY
DAVE C. MAHAFFEY
CHERYL A. NAKAMURA
BRADLEY G. BEHRMAN

Staff Assistant
LINDA J. XENOS

Published three times during the academic year in conjunction with the Harvard Legislative Research Bureau

CHRISTINE A. ALEXANDER
MARGARET A. ALLEN-CRAWFORD
MICHAEL J. ASTRUE
BRUCE S. BENNETT
WAYNE R. BIDSTRUP
LESLIE A. BLACKMON
SHIRLEY R. BOYD
ANDREW P. BRIDGES
SETH J. CHANDLER
PATRICIA J. CHRIST
MARK S. COHEN
ROBERT E. CULBERTSON, JR.
PAUL S. DARMITZEL
JESS H. DRABKIN
MICHAEL R. EASTMAN
ROBERT W. EMERSON
MATTHEW E. EPSTEIN
PATRICK F. FISHER
BARBARA J. GIERYMSKI
MATTHEW J. GOLDMAN
WILLIAM J. GUZICK
ROBERT P. HANEY
KATHRYN E. HOFF-PATRINOS
ROBERT T. HOWARD
WILLIAM R. JENKINS
DANIEL P. JOHNSON
MARY E. KANE
BRIAN D. KILB
LAUREL J. LENFESTEY
RICHARD LUSTIGER
MICHAEL E. LYTTON
ROBERT D. MANFRED, JR.
BRIAN J. MARTIN
VICTORIA T. MELENDEZ
STEVEN S. MICHAELS
BARBARA C. MOSES
CLAIRE C. OSBORN
SHARON K. OXBOROUGH
LARRY R. PARKINSON
LAURENCE F. PULGRAM
BRENDAN J. RADIGAN
STAN REEVES
MARIA A. SARATH
RONALD SCHEINBERG
GEORGE G. SIZE
JEFFERY SMISEK
JEFFREY A. STEIN
DAVID A. SUPER
BRENT D. TAYLOR
DON THOMPSON
PAMELA T. TROUT
MARY L. WOOLSEY
DANIEL J. ZUBKOFF

ARTICLE

RESTRAINING THE REGULATORS: LEGAL PERSPECTIVES ON A REGULATORY BUDGET FOR FEDERAL AGENCIES*

LANCE D. WOOD,** ELLIOTT P. LAWS,†
AND BARRY BREEN‡

The cry for regulatory reform presently resounds loud enough to be likely to bring about significant political action. One potential form such action could take is the development of regulatory budget legislation. Senate Bill 51 and House Bill 76, both introduced in the ninety-sixth Congress, were identical regulatory budget bills calling for the gradual imposition of limits on the costs of compliance that federal regulatory agencies could impose on the non-federal sector. This novel means of indirectly limiting federal regulators has been the source of much political discussion but little or no legal analysis.

In this Article, Messrs. Wood, Laws, and Breen provide a framework for legal analysis by describing some of the possible features of a regulatory budget and some of the conceptual problems involved in developing one. They present a series of theoretical options for splitting regulatory budget responsibility between the legislative and the executive branches, and suggest a design which they assert will foreclose a number of potential constitutional problems. Finally, the authors warn of serious potential conflicts between a regulatory budget and the Congressional Budget and Impoundment Control Act of 1974, and recommend that Congress solve such problems in any proposed legislation that may be designed to produce a regulatory budget.

Introduction

Federal regulatory agencies[1] have come under attack. Their critics have accused them of catering to special interests, stifling

---

* Opinions expressed in this article are those held by the authors individually. This article does not purport to reflect the views of any governmental department or agency.

** J.D., *cum laude*, University of Michigan, 1973; LL.M, with Highest Honors, George Washington University, 1977. Assistant Chief Counsel for Environmental Programs, United States Army Corps of Engineers.

† B.A., St. John's University, 1977; J.D., Georgetown University Law Center, 1980. Assistant District Attorney for New York County.

‡ A.B., with Honors, Princeton University, 1978. Member, Class of 1981, Harvard Law School.

1 For purposes of this Article, the term "regulatory agencies" generally includes those executive or independent administrative agencies responsible for making and enforcing rules to govern economic activities of the non-federal sectors. For example, regulatory agencies include the Occupational Safety and Health Administration, the Environmental Protection Agency, and the Securities and Exchange Commission. On the other hand, the term "regulatory agencies" does *not* include the federal courts, the Congress, or the Executive Office of the President. In the course of implementing

competition, and mandating wasteful or unnecessary expenditures.[2] In fact, "regulatory reform" has become a major concern of politicians and academicians alike.[3]

There is a large federal regulatory bureaucracy.[4] That bureaucracy imposes very significant economic costs on the American economy. But how significant are these costs? There is no simple answer. Costs can be measured simply in terms of the level of net outlays of the regulatory agencies, or in terms of the direct and indirect economic burdens agencies impose on the sectors they regulate. Measured by net outlays, some twelve regulatory agencies together spent over $8.53 billion in fiscal year 1979 alone.[5] Estimates of the burdens imposed on

---

regulatory reform, one possible approach in defining the scope of a specific proposal would be to list the agencies to be covered. *See, e.g.*, The Regulatory Reform Act of 1977, S. 600, 95th Cong., 1st Sess. (1977), for a list of major federal agencies which might be considered "regulatory."

2 *See, e.g.*, R. NOLL, REFORMING REGULATION: AN EVALUATION OF THE ASH COUNCIL PROPOSALS 100 (1971); *The Federal Paperwork Burden, Hearings Before the Subcomm. on Government Regulations of the Senate Select Comm. on Small Business*, 92d Cong., 2d Sess., and 93d Cong., 1st Sess. 461 (1972-1973); *Commission on Federal Paperwork, Hearings Before a Subcomm. of the House Comm. on Gov't Operations on H.R. 16424*, 93d Cong., 2d Sess. 34 (1974); S. PELTZMAN, REGULATION OF PHARMACEUTICAL INNOVATION (1974); 3 SENATE COMM. ON GOVERNMENTAL AFFAIRS, 95TH CONG., 1ST SESS., STUDY ON FEDERAL REGULATION: PUBLIC PARTICIPATION IN REGULATORY AGENCY PROCEEDINGS 16-19 (Comm. Print 1977); U.S. GENERAL ACCOUNTING OFFICE, LOWER AIRLINE COSTS PER PASSENGER ARE POSSIBLE IN THE UNITED STATES AND COULD RESULT IN LOWER FARES ii-iii (1977).

3 *Cf.* AMERICAN ENTERPRISE INSTITUTE, THE CANDIDATES 1980: WHERE THEY STAND 66-67 (1980) (Ronald Reagan); B. MITNICK, THE POLITICAL ECONOMY OF REGULATION (1980).

4 The Congressional Budget Office estimated that in 1976 there were at least 84,000 employees working for regulatory agencies. U.S. CONGRESSIONAL BUDGET OFFICE, THE NUMBER OF FEDERAL EMPLOYEES ENGAGED IN REGULATORY ACTIVITIES at v (1976).

5 U.S. DEP'T. OF THE TREASURY, DOC. NO. 3278, TREASURY COMBINED STATEMENT OF RECEIPTS, EXPENDITURES AND BALANCES OF THE UNITED STATES GOVERNMENT FOR THE FISCAL YEAR ENDED SEPTEMBER 30, 1977 (1979). The twelve regulatory agencies include the Occupational Safety and Health Administration, the Civil Aeronautics Board, the Commodity Futures Trading Commission, the Securities and Exchange Commission, the Consumer Product Safety Commission, the Federal Communications Commission, the Federal Maritime Commission, the Federal Trade Commission, the Interstate Commerce Commission, the Nuclear Regulatory Commission, the Environmental Protection Agency, and the Federal Aviation Administration.

However, even the seemingly straightforward measure of tallying net outlays has ambiguities. Is it appropriate to include all of an agency's net outlays, or only those most directly related to regulatory programs? Some agencies, for example, sponsor public awareness advertising campaigns, or fund research. What about regulatory expenses of agencies not ordinarily thought of as regulatory? The Army Corps of Engineers, for example, is well known for its civil works projects and for its military preparedness role, but it also regulates harbors and navigable rivers. Questions such as these make even the "mechanical" tallying of agency expenditures dependent upon

the non-federal sector, however, range from $10 billion to $130 billion annually.[6] These latter costs are the ones most commentators consider to be the greater evil.

This is not to say that all regulation is bad or unnecessary. Regulation is a tricky business: we want some of it, but not too much. For example, we want protection from exposure to thalidomide, but we do not want restrictions on access to safe and effective cures for cancer. In some theoretical sense, the optimal level of regulation is that amount just sufficient to balance marginal social costs and benefits, at the desired level of national economic activity. The practical calculus, however, is performed much more crudely.[7]

At present, the federal government has no mechanisms for consciously limiting the burdens which its regulatory agencies impose, or for setting priorities among them. Presently, a typical federal regulatory agency can write regulations or promulgate rules with only a partial understanding of the level of economic burdens they impose, and with little or no regard of similar costs imposed on the same groups of regulated industries by other federal agencies. Yet resources are finite. Therefore, unless the federal government has some efficient way of controlling the burdens it imposes through its regulatory bureaucracy, massive economic dislocation, and all the human misery it entails, ultimately may result.

Past Presidents have attempted to control the structure and effects of federal regulation by both institutional and informal means.[8] Within the executive branch, there are now three well-established institutional mechanisms designed to monitor implicit federal regulatory costs:

---

nice definitions and interpretations. Therefore, the total cited in the text should be treated as an approximation.

6 1 SENATE COMM. ON GOVERMENTAL AFFAIRS, 95TH CONG., 1ST SESS., STUDY ON FEDERAL REGULATION: THE REGULATORY APPOINTMENTS PROCESS, at III (Comm. Print 1977).

Regulation-imposed costs include reporting costs (e.g., additional costs of filling out forms and filing reports), compliance costs (e.g., costs required to purchase and install new capital equipment), and deadweight economic losses (e.g., implicit costs due to mandated anti-competitive or inefficient practices). Whatever the final incidence of these costs, significant costs probably are borne in the short run by each actor in the non-federal sector: private businesses, state and local governments, non-profit organizations, and individuals as consumers or wage earners.

7 *See* note 28 *infra*.

8 *See generally* DeMuth, *Constraining Regulatory Costs, Part I: The White House Review Programs*, REGULATION, Jan.-Feb., 1980, at 13.

(1) The Regulatory Council — composed of Cabinet-level officials — which sets policy on major issues relevant to multiple regulatory agencies;[9]

(2) The Council on Wage and Price Stability and the Regulatory Analysis Review Group — in the Executive Office of the President — which selectively reviews proposed regulations;[10] and

(3) Executive Order No. 12,044,[11] which requires each covered agency[12] to analyze the benefits, costs, and alternatives associated with its major regulations.

These mechanisms, however, tend to provide only a narrow, case-by-case review of regulatory initiatives, in which the perspective of the proponent agency dominates.

For example, section 3 of Executive Order No. 12,044 requires every executive branch agency (i.e., excepting the independent regulatory agencies) to prepare a "regulatory analysis" for any proposed regulations identified as likely to have major economic consequences.[13]

The order prescribes the following guidelines for identifying those regulations which require regulatory analyses:

> (a) Criteria. Agency heads shall establish criteria for determining which regulations require regulatory analyses. The criteria established shall:
>
> (1) ensure that regulatory analyses are performed for all regulations which will result in (a) an annual effect on the economy of $100 million or more; or (b) a major increase in costs or prices for individual industries, levels of government or geographic regions; and
>
> (2) provide that in the agency head's discretion, reg-

---

9 *See* UNITED STATES GOVERNMENT MANUAL 1980-81, at 724 (1980).

10 *See id.* at 108-09; Kelman, *Occupational Safety and Health Administration* in THE POLITICS OF REGULATION 256 (J.Q. Wilson ed. 1980).

11 Exec. Order No. 12,044, 3 C.F.R. 154 (1978) extended to April 30, 1981 by Exec. Order No. 12,221, 45 Fed. Reg. 44,249 (1980) [hereinafter cited as Exec. Order No. 12,044].

12 Covered agencies include both those under formal jurisdiction of the executive departments and those commonly known as the independent agencies. However, the President's unilateral authority over the independent regulatory agencies is narrowly circumscribed. Therefore, where executive orders are issued by the President, their applicability to independent regulatory commissions generally is intended to be advisory only. Executive Order No. 12,044 does not apply to independent agencies. Exec. Order No. 12,044, *supra,* note 11, at § 6(a)(5).

13 *Id.*

> ulatory analysis may be completed on any proposed regulation.[14]

Agencies subject to the order must establish procedures for any regulations which require a regulatory analysis:

> (b) Procedures. Agency heads shall establish procedures for developing the regulatory analysis and obtaining public comment.
>
> (1) Each regulatory analysis shall contain a succinct statement of the problem; a description of the major alternative ways of dealing with the problem that were considered by the agency; an analysis of the economic consequences of each of these alternatives and a detailed explanation of the reasons for choosing one alternative over the others.
>
> (2) Agencies shall include in their public notice of proposed rules an explanation of the regulatory approach that has been selected or is favored and a short description of the other alternatives considered. A statement of how the public may obtain a copy of the draft regulatory analysis shall also be included.
>
> (3) Agencies shall prepare a final regulatory analysis to be made available when the final regulations are published.[15]

Because it applies only to very substantial regulatory agency action and because it leaves the analysis to be performed by the agency itself, Executive Order No. 12,044 is but a very modest check on the promulgation of federal regulations.

Relatedly, the President and the Executive Office of the President have been able to exert some restraining influence through their informal powers of control, including their ability to influence budgets, through the President's power to dismiss his appointees, through the Executive Office's ability to review and comment on proposed regulations, and through the Office's position, which allows it to mediate or arbitrate disagreements among the various federal agencies or among federal agencies and non-federal interest groups.[16] As an example of the latter,

---

14 *Id.*

15 *Id.*

16 *See also* text following note 50 and preceding note 51 *infra*. President Ford drew attention to the magnitude of federal regulation, perhaps with the intention of alerting Congress or the public to comment upon or to act to check agency action. *See, e.g.*, Exec. Order No. 11,821, 39 Fed. Reg. 41,501 (1974) (expired) *as modified by* Exec.

consider the strict regulations that were initially proposed to limit cotton-dust air pollution in the textile industry. Those regulations were toned down after negotiations among a number of federal agencies, the industry, and workers' organizations.[17] Many of those negotiations were coordinated and mediated by the President's Office of Management and Budget (OMB).[18] Of course, the OMB must be circumspect when it attempts to influence the substantive content of proposed regulations; otherwise, it could be criticized or subjected to lawsuits alleging illegal interference with rulemaking authority vested in other agency heads by statute.[19]

Such existing controls do not provide at least three functional restraints:

> (1) a set of priorities for imposing costs, both within and among federal regulatory programs;
> (2) a mechanism to force consideration of alternative uses for those resources for the benefit of society; and
> (3) a measure of and a limit to the total and cumulative burden that may be imposed on each affected industry or sector.[20]

A regulatory budget, along with related procedures, could be developed to provide these restraints.

This Article describes what a regulatory budget is, investigates which mix of legislative and executive functions in a regulatory budgetary process would most surely pass constitutional muster, and considers how a regulatory budget might interact with existing legislation, specifically, the Congressional Budget and Impoundment Control Act of 1974.

Order No. 11,949, 42 Fed. Reg. 1017 (1977) (expired), which required economic impact statements when major new regulations were proposed.

17 *See* W. GELLHORN, C. BYSE, & P. STRAUSS, ADMINISTRATIVE LAW 136-37 (7th ed. 1979).

18 *See id.*

19 The legal standards that govern the manner and timing by which a President or his or her appointees may attempt to persuade an executive branch official to exercise his or her discretion are currently the subject of significant debate in scholarly legal publications and in the courts. *See, e.g.*, Bruff, *Presidential Power and Administrative Rulemaking*, 88 YALE L.J. 451 (1979); In re Permanent Surface Mining Regulation Litigation, No. 79-1144 (D.D.C., filed May 16, 1980). *See also* United States Department of Justice Memorandum Opinion written in connection with the *In re Permanent Surface Mining* case by Larry A. Hammond, Acting Assistant Attorney General, Office of Legal Counsel (January 17, 1979) (on file at the HARVARD JOURNAL OF LEGISLATION.)

20 *See Regulatory Budgeting and the Need for Cost-Effectiveness in the Regulatory Process, Hearing Before the Joint Economic Comm.*, 96th Cong., 1st Sess. 3 (1979) (statement of George C. Eads, Member of the Council of Economic Advisors).

I. WHAT IS A REGULATORY BUDGET?

A regulatory budget would set ceilings on the costs that regulatory agencies could impose on the non-federal sector by means of their power to promulgate and enforce regulations, and would coordinate those ceilings. A regulatory budget usefully can be analogized to the federal government's existing fiscal budget. Each was criticized as an untested innovation when first proposed.[21] The fiscal budget provides accountability by equalizing federal expenditures and revenues (including receipts from debt issues). Similarly, proponents claim that a regulatory budget would provide accountability of the regulatory agencies by equalizing expenditures mandated by federal regulations and available national resources.[22]

A regulatory budget could be designed and implemented in any of several possible forms. One form would establish an overall ceiling on the "total costs" that all federal regulatory agencies could impose on the non-federal sector every year. A related form would establish an overall ceiling for all agencies and individual ceilings for each covered department and agency. A third form would establish an overall ceiling, a ceiling for each covered department and agency, and a ceiling on particular programs, with special attention to overlaps among the different departments and agencies. Any of these forms might also break down overall ceilings into component categories of reporting costs, compliance costs, and deadweight economic losses.[23]

Certain federal agencies or certain types of regulations logically may be excluded from the regulatory budget. For example, regulations governing internal government operations, such as personnel matters, procurement, and agency organizational procedure and practice should not be covered because they do not involve direct regulation of the non-federal sector.[24] Moreover,

---

21 *Cf. id.* at 6-8.

22 Interview with James J. Tozzi, Director of the Office of Regulatory and Information Policy, Office of Management and Budget, in Washington, D.C. (July, 1979); Mayer, *Tozzi's Task: To Turn Back Clock on Ineffective Rules*, Wash. Star, March 30, 1980, § A, at 14.

23 *See* note 28 and accompanying text *infra* for a discussion of the complexities involved in measuring these costs.

24 While federal contracting processes may include detailed procedures, a regulatory budget would not sensibly limit the costs to bidders of following those procedures, because the bids themselves, over time, would include the costs of complying with the

federal policymakers may decide to exempt certain agencies because of practical considerations involving the nature of their mandate. For example, the Internal Revenue Service and the federal law enforcement agencies may be excluded from general regulatory budget constraints because the costs they impose are fundamental to the structure of the national economy and because they do not regulate purely economic activities. This Article will assume, however, that the great majority of federal regulations intended to control the economic decisions of actors in the non-federal sector eventually could become subject to the regulatory budget.

If a regulatory budget is to be developed, Congress or the agencies themselves must develop estimates of how much regulation is costing the economy now, and a methodology for predicting the probable costs regulation will impose in the future. Understandably, proponents and opponents of the regulatory budget concept dispute the adequacy of the data base which must be, and has not yet been, generated before implementing a regulatory budget. Data collection in certain areas has been going on quietly for a number of years. For example, for more than eight years the Bureau of Economic Analysis and the Bureau of the Census, two agencies of the Department of Commerce, have been collecting data from the private sector on the costs of complying with federal environmental laws and regulations;[25] the President's Council on Environmental Quality and the Environmental Protection Agency also have conducted systematic studies on federal environmental regulatory compliance costs.[26] Few other areas of the economy, however, have been analyzed as carefully.

Therefore, the regulatory budget, if adopted, may require incremental implementation, beginning with an "informational

---

additional procedures, or because these costs would be avoided by declining to contract with the federal government. The distinction here is between government as a sovereignty and government as a purchaser of goods and services.

25 The Bureau of the Census and the Bureau of Economic Analysis have compiled cost data for environmental regulations since 1972. These data and analyses thereof have been published periodically in the U.S. DEPT. OF COMMERCE, SURVEY OF CURRENT BUSINESS. *See, e.g.*, U.S. DEP'T. OF COMMERCE, *Pollution Abatement and Control Expenditures in Constant Dollars*, SURVEY OF CURRENT BUSINESS (Feb., 1979).

26 *See, e.g.*, *Economics*, in ENVIRONMENTAL QUALITY: THE TENTH ANNUAL REPORT OF THE COUNCIL ON ENVIRONMENTAL QUALITY, ch. 12 (1979) and sources cited therein.

regulatory budget" that would require every federal regulatory agency to develop a system for estimating the reporting costs, compliance costs, and deadweight economic losses of its regulations on the various sectors of the economy.[27] Only if and when these cost-reporting systems have been developed can a working regulatory budget be implemented. The initial regulatory budget cost estimates would be very rough. Measurement of indirect costs would require estimation of many economic parameters, such as the elasticities of supply and demand in each of many markets. Therefore, in the foreseeable future, only direct costs,[28] i.e., reporting costs and compliance costs, could be included in a regulatory budget.

---

27 Alternatively, because each agency may be self-interested in overestimating or underestimating present and future compliance costs, and because certain agencies may be better equipped for data gathering than others, an "independent" agency (such as the General Accounting Office) might be designated to develop these initial estimates.

28 Implementation of even an informational regulatory budget could not proceed before government accountants and economists resolve a number of difficult questions concerning precisely which of the various costs imposed by regulations would be counted as "regulatory compliance costs" for purposes of the regulatory budget, and how those costs should be measured. The most easily measured costs are those which regulations impose directly, including reporting costs and compliance costs. Nevertheless, regulations also impose large indirect costs on the non-federal sector, many of which can be best described as deadweight economic losses.

Indirect costs of regulation result from intentional or unintentional distortion of free markets, preventing the markets from reaching optimal levels of price and output. Indirect costs may appear as reductions in productivity or economic growth, or as increases in the inflation rate. They also may appear as improper governmental conduct and restraints on freedom from governmental control. Obviously, such indirect costs of regulation are difficult to quantify.

Exactly how such direct and indirect costs eventually are borne by shareholders, taxpayers, consumers, wage earners, or others is far from clear. The distribution of such costs could be estimated only after the estimation of economic parameters such as the elasticities of supply and demand for each particular industry and product, and after development of sophisticated econometric models. Consequently, implementation of a workable regulatory budget in the near future can occur only if regulatory compliance costs are defined to include the direct costs, but *not* the indirect costs, imposed by federal regulation.

This fact does not necessarily raise a serious objection to implementation of a full regulatory budget at some time. Many agencies may impose indirect costs through regulation which are more or less proportional to the direct costs of those same regulations. Of course, this is an assumption which must be verified, and the assumption will be less true of some economic regulatory agencies, such as the Interstate Commerce Commission, which impose indirect costs on the economy out of proportion to their associated direct costs. Knowing this at the outset, rough adjustments might be made in the allowable direct costs or in the multiplier used to gross up direct costs to direct plus indirect costs, while awaiting more precise measures of indirect costs. In any case, after a trial period, during which only an *informational* regulatory budget would be in force, a *mandatory* regulatory budget, limited to direct compliance costs, could be applied to many of the federal regulatory agencies.

Several specific proposals for the implementation of a broad regulatory budget have been developed. The most widely circulated of these proposals is Senate Bill 51,[29] sponsored by Senator Lloyd Bentsen of Texas.

Senate Bill 51 would initiate the process of developing a regulatory budget for virtually every agency of the executive branch. Only the Congress, the federal courts, the government of the District of Columbia, certain military authorities, and a few other entities not normally considered as federal executive agencies would be excluded from coverage.[30] The bill would include in its coverage practically all of the rules and regulations promulgated by covered agencies. "Rule" is defined as it is in the federal Administrative Procedure Act, which is incorporated by reference:

> "rule" means the whole or a part of an agency statement of general or particular applicability and future effect designed to implement, interpret, or prescribe law or policy or describing the organization, procedure, or practice requirements of an agency.[31]

The bill would establish a "Business Advisory Council" consisting of between 25 and 50 business leaders, drawn from each

---

29 The Regulatory Budget Act of 1979, S. 51, 96th Cong., 1st Sess (1979). Representative Clarence J. Brown introduced an identical bill in the House: H.R. 76, 96th Cong., 1st Sess. (1979). The declared intent of S. 51 is to reduce the total costs of compliance with federal rules and regulations by 5 percent per year for each of the first five years of the mandatory budget. S. 51, § 1101(b).

Testifying before the Joint Economic Committee, James T. McIntyre, Jr., Director of the Office of Management and Budget (OMB), agreed that federal regulation should be performed as efficiently as possible. Although he stated that the Carter Administration did not believe that the government could measure the costs and benefits of regulations with sufficient precision to implement S. 51, he also stated that the OMB has been working with a number of agencies to develop the methodology needed to estimate regulatory compliance costs accurately and reliably. *See 1980 Economic Report of the President: Hearings before the Joint Economic Committee*, 96th Cong. 2d Sess., pt. 1 (1980). Moreover, OMB's Office of Regulatory and Information Policy has sponsored an initial feasibility study to explore the possibilities of a regulatory budget and has drafted a bill designated as the "Regulatory Cost Accounting Act," which would establish an "informational," non-mandatory version of the regulatory budget for the executive branch. *See* Demuth, *Constraining Regulatory Costs (Part II): The Regulatory Budget*, REGULATION, Mar.-April, 1980, at 29, 30. Federal regulatory agencies, however, such as the Environmental Protection Agency, have expressed strong opposition to the draft OMB bill. *See, e.g., EPA Fears Proposed to Tally Costs Will Result in "Regulatory Budget"*, II ENVIR. REP. (BNA) 38 (Current Developments May 9, 1980); *cf.* Mayer, *Tozzi's Task*, note 22 *supra*.

30 S. 51, 96th Cong., 1st Sess. § 1108(1979); *cf.* 5 U.S.C. § 551(1) (1976).

31 5 U.S.C. § 551(4) (1976).

major industrial and commercial sector in the United States, to work with the President and the OMB in establishing a structure for the regulatory budget. The President, after consulting with the Council, would determine which classes of regulations would actually be covered by a final regulatory budget, and would develop methods to estimate regulatory compliance costs.[32] "Compliance costs" are defined in the bill to include all costs imposed on the non-federal sector as a result of complying with rules or regulations promulgated by a federal agency, such as additional personnel costs, capital costs, rent, interest, and state and local taxes, due to specific rules demanding extra data collection and recordkeeping, preparation and submission of forms, purchase of necessary equipment, and change in the quality or mix of raw materials or output.[33] Thus, the bill generally attempts to avoid the thorny problems involved in measuring the various indirect costs which regulations impose by omitting reference to the deadweight economic losses caused by federal regulations.

Under Senate Bill 51, each federal agency would be required to provide annual reports to the President, the Congress, and the Comptroller General, stating (1) the regulatory compliance costs imposed by that agency on the non-federal sector during the preceding fiscal year; (2) a comparison of those compliance costs with the regulatory budget established by Congress for the agency; (3) a full explanation for any excess of compliance costs over the agency's congressionally established regulatory budget for that fiscal year; and, (4) the estimated compliance costs for the current and the succeeding fiscal year for all existing and anticipated agency regulations.[34] The Comptroller General would evaluate such agency reports and inform Congress of inadequacies or errors.[35] For every fiscal year, the President would recommend to Congress a regulatory budget for each agency at the same time he submits his fiscal budget to Congress. If the President's proposed regulatory budget for an agency were lower than the estimated total compliance costs

32 S. 51, 96th Cong., 1st Sess. § 1102 (1979).
33 *Id.* at § 1108.
34 *Id.*
35 *Id.* at § 1103.

submitted by the agency, the President, in the Budget Message, would have to recommend specific actions which could be taken during the fiscal year to reduce compliance costs to the level mandated by the regulatory budget.[36] On or before September 15 of each year, Congress would be required to complete action on a concurrent resolution to establish a regulatory budget ceiling for every agency, i.e., setting the maximum total compliance costs for all that agency's regulations during the coming fiscal year. The Committees on the Budget of the House and Senate would set these regulatory budget ceilings after considering the estimates of the agencies, the recommendations of the President, and the views of the congressional standing committees, the Joint Economic Committee, and the Joint Committee on Internal Revenue Taxation.[37]

Senate Bill 51 does not explicitly provide a mechanism for enforcing the agencies' regulatory budget ceilings, or for calling to account agencies which have imposed compliance costs on the non-federal sector in excess of the ceilings established by Congress' concurrent resolution. This is hardly surprising, since a single enforcement mechanism might not function well in every case. Here, Congress would be free to tailor remedies to account for the perceived culpability of the agency's exceeding its regulatory budget.

To coordinate the concept of regulatory budgeting with future legislation, Senate Bill 51 also requires each congressional committee to estimate the compliance costs that would be imposed by each bill or joint resolution reported upon after the effective date of the regulatory budget bill. Any House or Senate bill, resolution, or amendment which would have the effect of exceeding an agency's regulatory budget ceiling (if enacted) would be deemed to be out of order by Senate Bill 51 unless the House or Senate grants a waiver.[38] Finally, the OMB periodically would issue reports on pending legislation, estimating compliance costs and comparing them with the enforcing agencies' respective regulatory budgets.[39]

---

36 *Id.* at § 1104.
37 *Id.* at § 1105.
38 *Id.* at § 1107.
39 *Id.* at § 1106.

Senate Bill 51 shows how regulatory budget legislation can dovetail existing, statutorily defined concepts and procedures into a full plan to control federal agencies. However, the bill also illustrates the additional burdens a regulatory budget may impose on Congress' own decision-making and bookkeeping machinery, as well as on the data collection and cost estimation facilities of existing management and regulatory agencies.

II. An Initial Legal Analysis

No matter how attractive the concept of a federal regulatory budget may be at this stage, the idea cannot be implemented rationally until a number of legal questions have been answered. These questions concern the division of power between the executive and legislative branches of the federal government.

The fundamental legal issue here is, assuming that a regulatory budget is desirable, which branch of government — Congress or the President — has the constitutional authority to compel its implementation, and in what form?

A regulatory budget could be fashioned in a number of different ways, with more or less of the initiative coming from the President. To keep the analysis of separation of powers manageable, this Article considers four options for implementing a regulatory budget, presented in order of a decreasing presidential role and an increasing congressional one. Specifically, Option 1 is based on a bold, unilateral assertion of presidential power to establish a regulatory budget. Option 2 is based on an initial authorization to the President by Congress to implement a regulatory budget, and in which Congress would retain some variable degree of control. Option 3 is the same as Option 2, but it would include a "one-house congressional veto." Finally, Option 4 is based on a full fiscal-budget-type process, involving both the President and Congress, and in which Congress would retain very substantial power to fix the levels of non-federal sector costs among — and perhaps within — regulatory agencies' programs. The following subsections will describe and discuss these alternative approaches to implementing a regulatory budget.

A. *Option 1: The Executive Order and Other Unilateral Executive Action*

The President could attempt to implement a regulatory budget unilaterally, either for the executive branch as a whole or for selected federal regulatory agencies, through the use of executive orders or through informal executive action.[40] However, he or she could not effect the full range of objectives of a regulatory budget in this fashion for three reasons. First, statutes authorizing agencies to promulgate regulations generally vest discretionary authority in the agency head, not directly in the President.[41] The Supreme Court has held that when a federal officer is legally vested with discretionary authority, he or she may not be directed in the use of that discretion by a superior officer.[42] In *United States ex rel. Accardi v. Shaughnessy,* the Supreme Court ruled: "If the word 'discretion' means anything in a statutory or administrative grant of power, it means that the recipient must exercise his authority according to his own understanding and conscience."[43] The Court then held that the Attorney General could not direct the use of a subordinate's discretion, even where the Attorney General had himself granted the discretion to the subordinate and retained ultimate review of the decision for himself. Therefore, the President would have even less of a claim under unilateral action to establish a regulatory budget than did the Attorney General in *Shaughnessey:* since the President has not delegated authority to the agency head, and since he or she ordinarily has no right of ultimate review under the legislation delegating responsibility to the agency head, the President has even less power to control the discretion delegated to the agency head than if he or she had delegated the authority personally.

Second, the establishment of a federal regulatory budget by the President would, perforce, set priorities among agencies in their relative abilities to impose costs on the private sector

40 *See* text accompanying notes 8 to 20 *supra.*
41 *E.g.,* Occupational Safety and Health Act of 1970, Pub.L.No.91-596, § 6, 84 Stat. 1590 (1970) (amended 1974, 1978) (Secretary of Labor); Federal Food, Drug, and Cosmetic Act, 21 U.S.C. § 371 (1976) (then-Secretary of Health, Education, and Welfare).
42 Accardi v. Shaughnessy, 347 U.S. 260 (1954).
43 *Id.* at 266-67.

through regulations.[44] To stay within that budget, some agencies might have to forego enforcement of certain regulatory programs that were mandated by Congress. Since a regulatory budget would result in systematic enforcement of some laws, partial enforcement of others, and non-enforcement of still others, its operation would amount to executive lawmaking, a power not granted to the President by the federal Constitution.[45]

Third, the concept of a regulatory budget is analogous to a presidential impoundment of funds, since the regulatory budget might prevent the full implementation of programs mandated by Congress. "Impoundment" results from an executive determination not to expend funds appropriated by Congress. Federal courts consistently have held that for the Executive to impound funds, he or she must have the permission of Congress.[46] In *Kennedy v. Mathews*,[47] the federal district court for the District of Columbia examined an appropriation of $187.5 million to the Department of Health, Education, and Welfare. The appropriation act did not authorize the executive branch to prohibit the expenditure of the appropriated funds, and "since the executive branch [lacked] any inherent power to impound funds," the court granted plaintiffs injunctive relief, forcing all funds to be made available.[48] The court continued: *"There is no longer any doubt that in the absence of express Congressional authorization to withhold funds appropriated for implementation of a legislative program the executive branch must spend all funds."*[49] (Emphasis added.)

If the President were to attempt to implement the regulatory budget by unilateral executive action, i.e., without specific congressional authorization, the federal courts may draw the analogy to fiscal impoundment and void the regulatory budget restrictions on agency action.[50]

---

44 *But see* text following note 98 and accompanying note 99 *infra*.

45 U.S. CONST. art. I, § 1. *See* Youngstown Sheet & Tube Co. v. Sawyer, 343 U.S. 579 (1952).

46 *E.g.*, Train v. City of New York, 420 U.S. 35 (1975); Kendall v. United States, 37 U.S. (12 Pet.) 524 (1838).

47 413 F. Supp. 1240 (D.D.C. 1976).

48 *Id.* at 1245.

49 *Id.* at 1245. *See also* Train v. City of New York, 420 U.S. 35 (1975).

50 *See* text accompanying notes 91 to 105 *infra*, for a discussion of regulatory budgets and legislatively imposed restrictions on impoundments codified in the Congressional Budget and Impoundment Control Act of 1974.

Even though the President does not have authority to make a full regulatory budget legally binding throughout the executive branch, however, he or she may still be able to accomplish some of the objectives of a regulatory budget by using existing informal presidential powers. For example, the President or the OMB might be able to persuade the head of an executive agency to use the agency head's "discretion" to modify the substance of a regulation, or to delay or cease promulgation or enforcement of it. The President and the Executive Office of the President are able to exert great influence over executive branch agency heads through, *inter alia*, budgetary and political controls and the President's power to dismiss executive branch appointees with or without cause.[51]

Moreover, even though the President has no legal authority unilaterally to require agencies to implement regulatory budgets, and even though informal efforts at persuasion cannot achieve all the results expected from a regulatory budget, it might be argued that Congress has previously granted the President a discretionary authority broad enough to allow presidential implementation of the regulatory budget in the Reorganization Act of 1977 (the Act).[52] That statute gave the President broad authority to reorganize the federal executive branch until April 6, 1980, subject to possible veto by either house of Congress. However, the Act specified that "no enforcement function or statutory program shall be abolished."[53] Because a regulatory budget might effectively limit or abolish those enforcement functions which exceed a programmatic or overall agency regulatory budget ceiling, the Act would have forbidden the use of that statute to establish a regulatory budget, even if the President had acted before April 6, 1980. No other statute appears to grant the President sufficient powers today to implement a regulatory budget. Therefore, a full regulatory budget will require some affirmative grant of power from Congress.

---

51 *See, e.g.*, Humphrey's Executor v. United States, 295 U.S. 602 (1935); Myers v. United States, 272 U.S. 52 (1926); Ex Parte Hennen, 38 U.S. (Pet.) 230 (1839).

52 5 U.S.C. §§ 901-912 (Supp. II 1978). Similar grants of authority have been enacted since at least 1932. Pub. L. No. 72-212, tit. IV, 47 Stat. 382 (1932).

53 5 U.S.C. § 903(a)(2) (Supp. II 1978).

B. *Option 2: Congressionally Authorized Executive Action*

Under the second major option, Congress would enact legislation designating which federal agencies are to be covered by regulatory budgets, and establishing a process by which Congress would set the overall ceiling for the regulatory budgets of all covered agencies combined. Under this option, however, the President could determine the specific regulatory budget ceiling for each covered federal agency. Alternatively, the basic legislation could allow the President to set both the overall ceiling for the government-wide regulatory budget as well as the specific ceilings for each agency subjected to a regulatory budget by Congress. Or, the legislation could vest nearly absolute power in the President by allowing him or her to designate which federal departments and agencies would be subject to regulatory budgets, as well as allowing him or her to set the regulatory budget ceilings.

In any of these possible configurations, Congress would be delegating an extraordinary amount of authority to the President. Consequently, even though the Supreme Court has not voided as unconstitutional any congressional delegation of authority to the Executive since 1935,[54] there is still some possibility that such a sweeping delegation might be declared unconstitutional under the delegation doctrine. On the other hand, that possibility can be minimized by special care in drafting the specific language of the enabling legislation.

A statute adopting any of the possible variations of congressionally authorized executive action suggested above would effectively delegate to the President authority to determine which congressionally mandated programs would be implemented by regulation, and to what extent they would be implemented. In contrast, the "normal" congressional delegation of legislative authority to the Executive merely authorizes the President or the head of the federal department or agency to implement a more or less detailed legislative design through regulations or other mechanisms. In other words, the enabling

---

54 Panama Refining Co. v. Ryan, 293 U.S. 388 (1935); A.L.A. Schechter Poultry Corp. v. United States, 295 U.S. 495 (1935).

legislation which would authorize the President to establish a regulatory budget, without additional legislative oversight or guidance, essentially would allow the executive branch to overrule or limit select congressional decisions at the implementation level. This should be distinguished from the normal situation in which the President or agency head is authorized merely to supplement or implement legislative action. It might be argued that regulatory budgetary authority thus is different *in kind* from present grants of power encountered in ordinary delegation decisions. Assuming that Congress would adopt the Option 2 approach only if it wanted to give the President considerable flexibility in setting a regulatory budget for each affected federal agency, Congress would provide few standards by which the President's discretionary actions could be limited within constitutional bounds.[55] Since the problems with the delegation doctrine in an Option 2 arrangement would be very serious, they deserve further elaboration.

In hundreds of cases decided between the mid-19th century and the present day, the Supreme Court has upheld as constitutional significant and broadly worded congressional delegations of legislative authority to the President, and to executive or independent agencies. In all those years, the Supreme Court has voided similar congressional delegations in only two instances. The cutting edge ostensibly honed in most of these cases is that Congress must specify "meaningful standards" or an "intelligible principle" to guide the Executive in using delegated power.[56]

In 1935, the Supreme Court held unconstitutional two statutory delegations of authority to federal officials. In *Panama Refining Co. v. Ryan*,[57] the Supreme Court held unconstitutional a provision of the National Industrial Recovery Act (NIRA) which delegated to the President authority to prohibit shipment in interstate commerce of oil produced in violation of state law.

---

55 *Cf.* U.S. CONST. art. I, § 1 ("all legislative powers herein granted shall be vested in a Congress of the United States"); *Id.* art. III (powers of the Executive).

56 *See, e.g.*, American Power & Light Co. v. S.E.C., 329 U.S. 90 (1946); Opp Cotton Mills v. Adm'r, 312 U.S. 126 (1941). *See generally* 1 K. DAVIS, ADMINISTRATIVE LAW TREATISE ch. 3, § 3.9 (2d ed. 1978).

57 293 U.S. 388 (1935). (However, the act in question provided for criminal penalties under circumstances obscuring the receipt of actual notice by potential defendants. This may have influenced the Court's holding.)

The Court declared that the broad congressional delegation contained no standards to guide the President's discretion, and that the general policy statements in NIRA were not adequate to remedy this constitutional defect.

Similarly, in *A.L.A. Schechter Poultry Corp. v. United States*,[58] the Supreme Court declared unconstitutional the far-reaching NIRA provisions which had attempted to delegate to the President congressional authority to approve codes governing business activity of industries subject to federal jurisdiction. The NIRA provisions in dispute attempted to delegate to the President the authority to fix prices for trade and industry, and to formulate and approve codes governing a wide range of business activity and practice, guided only by vague standards or goals such as one "to eliminate unfair competitive practices, to promote the fullest possible utilization of the present productive capacity of industries . . . and otherwise to rehabilitate industry and to conserve natural resources."[59] According to the Supreme Court,

> In view of the scope of that broad declaration, and of the nature of the few restrictions that are imposed, the discretion of the President in approving or prescribing codes, and thus enacting laws for the government of trade and industry throughout the country, is virtually unfettered.[60]

That broad and undirected delegation was more than the Court could tolerate under the Constitution in 1935. However, since 1935, the Supreme Court has done little more than assert that the constitutional delegation standard is one of "meaningful standards." Actual holdings of the Court have failed to enforce that standard on Congress: the holdings either accept extremely vague standards, or require virtually no standards at all. For example, in 1948, in *Lichter v. United States*,[61] the Supreme Court held that the standard expressed by the term "excessive profits" was a constitutionally adequate limitation to guide an administrative agency in recovering such profits under the War Contracts Renegotiation Act.[62] In fact, delegations upheld as

58 295 U.S. 495 (1935).
59 *Id.* at 531 n.9 (Congress' declaration of policy).
60 *Id.* at 541-42.
61 334 U.S. 742 (1948).
62 The determination of sufficiency of the standard was based on the fact that

constitutional before *Panama Refinery* and *Schechter* were not much different from the delegations upheld after those cases. For example, in 1932, the Court in *New York Central Securities Corp. v. United States*[63] upheld a statute allowing consolidation of carriers when "in the public interest;" and in 1943, in *National Broadcasting Co. v. United States*,[64] the Court upheld a statute providing for the licensing of radio communications "as public convenience, interest, or necessity requires."

A number of cases demonstrate that the Court frequently has required *no* standard to limit delegated power. For example, *St. Louis & Iron Mountain Ry. v. Taylor*[65] upheld a statute which authorized the American Railway Association to set a mandatory height for drawbars on freight cars, but set no standard; *McKinley v. United States*[66] upheld a statute which authorized the Secretary of War to suppress "houses of ill fame . . . within such distance as he may deem needful of any military camp . . . ,"[67] but set no standard; the *Intermountain Rate Cases*[68] upheld a statute authorizing the Interstate Commerce Commission to grant exemptions to established rate requirements, and also set no standards. Finally, *Fahey v. Mallonee*[69] upheld a statutory delegation of authority to provide for the liquidation of savings and loan associations, but set no standards; and *Carlson v. Landon*[70] sustained a statutory delegation of authority to the Attorney General, allowing him broad discretion to grant or deny bail to aliens, and also set no standards. Therefore, despite the numerous cases in which the Supreme Court has *stated* that congressional delegations of authority must be limited by "meaningful standards," in fact the Court has upheld against delegation doctrine attack many far-reaching

---

administrative practices had already been developed to interpret the term, as considered by Congress and as reflected in the statute. *Id. See also* Yakus v. United States, 321 U.S. 414 (1944).

63 287 U.S. 12 (1932).

64 319 U.S. 190 (1943).

65 210 U.S. 281 (1908).

66 249 U.S. 397 (1919).

67 *Id.* at 398.

68 234 U.S. 476 (1914).

69 332 U.S. 245 (1947).

70 342 U.S. 524 (1952). (The Court imputed standards from other provisions of the relevant act.)

delegations of legislative authority guided by practically no principle or standard whatever.

One cannot state with certainty whether *Panama Refining* or *Schechter* has been overruled implicitly by the many Supreme Court decisions decided since 1935 which have upheld broad delegation, or whether either remains a viable precedent. Its revival, as a fundamental constitutional principle, has been debated in the scholarly literature.[71] Moreover, from time to time the Supreme Court has raised the spectre of the *Schechter* decision in *obiter dicta*, suggesting a palpable possibility that it remains available to destroy some egregious super-delegation of congressional authority, such as a regulatory budget under Option 2. For example, the opinion in *American Trucking Association v. United States*,[72] decided in 1953, cited both *Schechter* and *Panama Refining* as sound precedent. Three dissenting justices also cited these two decisions with approval in the later case of *Arizona v. California*.[73]

The delegation doctrine has received considerable recent attention in the lower federal courts as well. For example, in 1971, a three-judge panel of the federal district court of the District of Columbia reviewed the constitutionality of the Economic Stabilization Act of 1970. The court upheld the Act's extensive delegation of wage and price control authority to the President, but only after a thorough review of the Act, its legislative history, and numerous Supreme Court precedents, including *Schechter* and *Panama Refining*.[74] The court, in *Amalgamated Meat Cutters & Butcher Workmen v. Connally*,[75] relied primarily on the Supreme Court's decision in *Yakus v. United States*[76] (sustaining the constitutionality of the Emergency Price Control Act of 1942, a broad authorization for executive branch price setting). The court in *Amalgamated* cited arguments sim-

---

71 *Compare* Wright, *Beyond Discretionary Justice*, 81 YALE L.J. 575 (1975); T. LOWI, THE END OF LIBERALISM 297-99 (1969); and Merrill, *Standards — A Safeguard for the Exercise of Delegated Power*, 47 NEB. L. REV. 469 (1968).

72 344 U.S. 298, 313 (1953).

73 3783 U.S. 546, 626 (1963).

74 Amalgamated Meat Cutters & Butcher Workmen v. Connally, 337 F. Supp. 737 (D.D.C. 1971).

75 *Id.*

76 321 U.S. 414 (1944).

ilar to those used by the *Yakus* Court, and held that the 1970 statute and its legislative history provided adequate standards to ensure that the Executive would faithfully obey the expressed will of Congress.[77]

In summary, therefore, it is far from clear that the federal courts would uphold the extraordinary delegation of authority implicit in the Option 2 regulatory budget plan under a delegation doctrine challenge, since Option 2 would delegate powers notably greater and possibly more substantial in kind than those considered in every one of the delegation cases. Obviously, if Congress were to decide to adopt an Option 2-type statute despite this possibility, the statute should be carefully drafted to maximize the likelihood that it will be upheld. The drafters would be well advised to study the Supreme Court's delegation decisions thoroughly and to incorporate those features which seem determinative in withstanding delegation doctrine challenge.[78]

C. *Option 3: Executive Action with a Possible Congressional Veto*

Option 3 would be identical to Option 2, except that under Option 3 each significant exercise of presidential power would

---

77 337 F. Supp. at 746-52.

78 Congress has authorized the Supreme Court to implement rules which would supersede conflicting congressional directions. 28 U.S.C. § 2072 (1976) authorizes the Supreme Court to prescribe rules of civil procedure for the federal courts, stating that "[a]ll laws in conflict with such rules shall be of no further force or effect after such rules have taken effect." *Id.* The Supreme Court periodically has published rules pursuant to this section, implicitly upholding the section as constitutional. However, there are several factors which distinguish this statutory section from a statute authorizing the President to establish regulatory budgets. First, 28 U.S.C. § 2072 (1976), by its own terms provides that this rulemaking power may not affect "substantive rights," only procedural rights. In contrast, any regulatory budget would limit the enforcement of *substantive* provisions of law. Second, 28 U.S.C. § 2072 (1976) provides that court rules must lie before Congress 90 days before they may become effective, during which time Congress may change or "veto" them. This is in contrast to the proposed regulatory budget bills considered in the preceding sections. *But see*, the next subsection, *infra*. Third, the judiciary may be less worried about congressional delegations to *courts* than to the *executive branch*. Judges might be expected to trust judges more than politicians; and every federal judge must be confirmed by the Senate. In contrast, the President is, in the usual case, never approved by either house of Congress. *But see* U.S. CONST. art. II, § 1; U.S. CONST. amend. XII (if a majority of the Electoral College fails to choose one person, the House of Representatives decides who shall be President).

be subject to a "one-house congressional veto." For example, the President could be required to send Congress his or her proposed regulatory budget affecting each covered department or agency. If neither the Senate nor the House of Representatives passed a resolution objecting to that specific regulatory budget level within a specified period of time (e.g., 60 days), that level would become the legal ceiling for costs that could be imposed by the appropriate agency.

A number of constitutional scholars praise the one-house veto as an additional, albeit belated, check on otherwise unbridled Executive discretion.[79] The one-house veto reserves significant power in Congress, enhancing the likelihood that, if challenged, the courts would not hold the legislation to be too great a delegation of power to the President.

However, the one-house veto itself has been attacked as unconstitutional on three separate grounds:

(1) The legislative power is vested in both houses of Congress acting together, not in either house acting alone;

(2) The Constitution places the veto power in the President, not in the Congress; and

(3) All executive power belongs to the President; thus once a power is delegated it should remain in the President to avoid turning legislators into administrators.[80]

Although several courts have ruled on one-house veto provisions,[81] these cases can not be considered to resolve the issue as it would apply to a regulatory budget. Specifically, in *Atkins*

---

79 *See, e.g., Providing Reorganization Authority to the President, Hearings on H.R. 3131, H.R. 3407, and H.R. 3442 Before the Legislation and National Security Subcomm. of the House Comm. on Gov't Operations*, 95th Cong., 1st Sess. 76 (1977) (statement of Laurence H. Tribe).

80 *See* McGowan, *Congress, Court, and Control of Delegated Power*, 77 COLUM. L. REV. 1119, 1152-62 (1977); Atkins v. United States, 556 F.2d 1028, 1062 (1977). *See also* articles critical of the constitutionality of a congressional veto: Ginnane, *The Control of Federal Administration by Congressional Resolutions and Committees*, 66 HARV. L. REV. 569 (1953) *and* Watson, *Congress Steps Out: A Look at Congressional Control of the Executive*, 63 CALIF. L. REV. 983 (1975). Articles supportive of the constitutionality of a congressional veto include Cooper & Cooper, *The Legislative Veto and the Constitution*, 30 GEO. WASH. L. REV. 467 (1962); Stewart, *Constitutionality of the Legislative Veto*, 13 HARV. J. LEGIS. 593 (1976); Cooper, *The Legislative Veto: Its Promise and Its Perils* in 7 PUBLIC POLICY YEARBOOK OF ADMINISTRATION 128-74 (1957).

81 Buckley v. Valeo, 424 U.S. 1 (1976); Atkins v. United States, 556 F.2d 1028 (Ct. Cl. 1977) *cert. denied*, 434 U.S. 1009 (1978); Pressler v. Simon, 428 F. Supp. 302 (D.D.C. 1976) *vacated and remanded*, 431 U.S. 169 (1977).

v. United States,[82] the Court of Claims, in 1977, upheld the one-house veto provision as it applied to the setting of congressional salaries. However, the *Atkins* court carefully limited its decision to the case at hand; there is an especially strong argument to be made for each house's independence vis-a-vis the other house and vis-a-vis the President. The court twice noted explicitly that the case did *not* involve a regulatory matter.[83] In *Pressler v. Simon*,[84] the federal district court for the District of Columbia, in 1976, upheld the operation of the one-house veto in the context of a challenge to the Postal Revenue and Salary Act of 1967 and the Executive Salary Cost-of-Living Adjustment Act of 1975. On appeal, the Supreme Court affirmed the ruling but did so without an opinion. Justice Rehnquist filed the only written opinion and noted therein that the affirmance did not necessarily uphold the one-house veto, since it could just as easily have been based on the conclusion that petitioner Pressler did not have standing to sue. Finally, in *Buckley v. Valeo*,[85] Justice White wrote, in the context of reviewing the statute that created the Federal Election Commission and empowered it to promulgate regulations:

> I am of the view that the otherwise valid regulatory power of a properly created independent agency is not rendered constitutionally infirm, as violative of the President's veto power, by a statutory provision subjecting agency regulations to disapproval by either House of Congress.[86]

However, Justice White was speaking only for himself. No other Justice addressed the issue, and the *per curiam* opinion expressly declined to rule on it.[87] Therefore, the existing cases leave the issue of the constitutionality of the one-house veto largely unresolved. So, while the one-house veto might make Option 3 less likely than Option 2 to be struck down as an overbroad delegation of authority, the veto presents its own constitutional questions, and further uncertainty, into such a regulatory budget proposal.

---

82 556 F.2d 1028 (Ct. Cl. 1977).
83 *Id.* at 1059.
84 428 F. Supp. 302 (D.D.C. 1976) *vacated and remanded*, 431 U.S. 169 (1977).
85 424 U.S. 1 (1976).
86 *Id.* at 284.
87 *Id.* at 140 n.176.

D. *Option 4: Joint Legislative and Executive Action*

Under Option 4, Congress would enact legislation designating which federal departments and agencies are to be governed by the regulatory budget, and creating a comprehensive regulatory budget process through which to implement that budget. This process would be analogous to the present fiscal budgetary process.[88] Under this approach, the President annually would propose a regulatory budget to cover each of the designated agencies. Congress would review the proposed regulatory budget and make the adjustments it deems necessary or desirable. Congress would enact the budget, with the President's signature, or over his veto.

Option 4 maximizes congressional involvement in the regulatory budget process. Nevertheless, Option 4 still involves congressional delegation to the executive branch. Even more than in the fiscal budgetary process, in the regulatory budgetary process each affected agency head might have discretion to choose which of his or her agency's regulatory programs would be enforced. For example, even where an agency's substantive legislation provides that the agency head "shall" promulgate implementing regulations for a regulatory program, the regulatory budget may result in an overriding grant of discretion to the agency head.[89] On the other hand, where an agency's substantive legislation provides that the agency head "may" promulgate implementing regulations, the regulatory budget will not result in an additional grant of discretion in the case of *new* regulations; however, the regulatory budget may result in additional discretion to selectively enforce *existing* regulations.

In spite of this new delegation, courts almost certainly would uphold Option 4 under a delegation doctrine attack. Unlike Option 2, where discretion would be delegated to the President, here the delegation would be to agency heads. This is a difference of several orders of magnitude.[90]

---

88 S. 51, discussed at text accompanying notes 29 to 39 *supra*, seems to incorporate an Option 4 proposal for joint legislative and executive action.

89 This will be true whether the agency has already promulgated, or has yet to promulgate, regulations. The discretion of the agency head, however, will be affected only as it relates to regulatory functions. *See* text following note 98 and preceding note 99 *infra*.

90 *Cf.* text accompanying notes 42 and 43 *supra*.

E. *Comparing the Options*

From a legal perspective, Option 4 raises the least serious constitutional problems. Moreover, in addition to providing the needed congressional involvement, Option 4 has the important pragmatic advantage of being closely analogous to the current fiscal budgetary process. Since Congress and the President already are familiar with the administrative mechanisms of a fiscal budget, the regulatory budget could be adopted and implemented sooner and more certainly under Option 4 than under any of the other options.

III. THE REGULATORY BUDGET AND THE CONGRESSIONAL BUDGET AND IMPOUNDMENT CONTROL ACT OF 1974

Once a specific regulatory budget mechanism is decided upon, legislative drafters must pay close attention to the interaction of their proposed statute with existing law and legislative policy. At least one important statute would bear an obvious relationship to any regulatory budget: the Congressional Budget and Impoundment Control Act of 1974 (the 1974 Act).[91] A restrictive regulatory budget ceiling could force a federal agency to curtail or eliminate a regulatory program which Congress has authorized and funded.[92] If curtailment or elimination can be traced to presidential action in setting a low regulatory budget ceiling for the agency, that action might reasonably be construed to invoke the provisions of the 1974 Act, including the requirement that the President send a special message to Congress, specifying amounts, reasons, justifications, and other relevant information concerning funds appropriated but not spent by the agencies.

Section 1012(a) of the 1974 Act specifically states that

> Whenever the President determines that all or part of any budget authority will not be required to carry out the full objectives or scope of programs for which it is provided or that such budget authority should be rescinded for fiscal policy or other reasons (including the termination of au-

91 Pub. L. No. 93-344, 88 Stat. 297 (1974).
92 *See* text accompanying notes 46 to 50 *supra*.

> thorized projects or activities for which budget authority has been provided), or whenever all or part of budget authority provided for only one fiscal year is to be reserved from obligation for such fiscal year, the President shall transmit to both Houses of Congress a special message. . . .[93]

Similarly, section 1013(a) states that

> Whenever the President, the Director of the Office of Management and Budget, the head of any department or agency of the United States, or any officer or employee of the United States proposes to defer any budget authority provided for a specific purpose or project, the President shall transmit to the House of Representatives and the Senate a special message. . . .[94]

Since the 1974 Act specifically limits its scope to executive branch activities affecting a "budget authority," one of the threshold issues here is whether funds appropriated for regulatory programs constitute such "budget authority." Section 3 of the 1974 Act generally defines "budget authority" as "authority provided by law to enter into obligations which will result in immediate or future outlays involving Government funds."[95]

The Office of Management and Budget defines "budget authority" to include funds to be spent by agencies for regulatory programs and to be received through the appropriations process:

> *Budget authority.* Budget authority for any year represents the authority provided by law and becoming available during the year to incur obligations. [One] basic [form] of "budget authority" [is] the following:
>
> > *Appropriation.* Statutory authority that allows Federal agencies to incur obligations and to make payments out of the Treasury for specified purposes. This is the most common form of budget authority. (Note that certain types of appropriations *are not counted* as budget authority; these are appropriations: (a) to liquidate contract authority, (b) applied to the reduction of outstanding debt, (c) for refunds of receipts, and (d) for payment to the International Monetary Fund.)[96]

---

93 31 U.S.C. § 1402 (1976).
94 *Id.* at § 1403 (1976).
95 *Id.* at § 1302 (1976 & Supp. II 1978).
96 OMB Circular No. A-34, at 6 (1976)(emphasis in original).

The definition applies to "all appropriations, funds, and other authorizations, except deposit funds. . . ."[97] Thus, appropriations for federal regulatory programs would appear to constitute "budget authority." Indeed, the Conference Committee Report on the 1974 Act provides additional support for this conclusion:

> The managers intend that the definition of "budget outlays" and "budget authority" for purposes of the congressional budget process *be the same as that used for the executive budget* and that any item which is excluded by law from the executive budget may be excluded from any specification of budget outlays or budget authority in the congressional budget process.[98]

Although regulatory agency budgets generally are considered under the 1974 Act as "budget authority," unless the regulatory budget ceilings impinge on the agencies' appropriations for those regulatory programs, the reporting requirements of the 1974 Act will not be invoked. For example, if an agency has exhausted its regulatory budget for a fiscal year but still has unspent appropriated funds, the agency could avoid the reporting requirements by spending the funds for other authorized purposes. For instance, where Congress has appropriated funds to regulatory agencies in a lump sum, those agencies might avoid direct conflicts with the 1974 Act by using those funds for non-regulatory purposes, such as meeting unexpected administrative expenses on training programs for its personnel, public awareness campaigns, or research, or for funding non-regulatory social programs. Thus, where the appropriation is not limited by statute to particular regulatory purposes, the agency can shift the funds to other purposes which do not impose costs governed by the regulatory budget. The Comptroller General has held that Congress has recognized that in most instances it is desirable to maintain executive flexibility to shift around funds within a particular lump-sum appropriation account so that agencies can make necessary adjustments for "unforeseen developments, changing requirements, incorrect price estimates, wage-rate adjustments . . . and legislation enacted subsequent to appropriations."[99]

---

97 *Id.* at 3.

98 H.R. REP. No. 1101, 93d Cong., 2d Sess. 50 (1974) (emphasis added).

99 *Cf.* 55 COMP. GEN. 307, 318 (1975) (citing, in part, Fisher, *Reprogramming of Funds by the Defense Department,* 36 J. POL. 77, 78 (1974).

Breakdowns of lump-sum appropriations often are made in Committee reports, but

If a covered agency does not spend its entire appropriation, however, this action would be a "rescission" or a "deferral" as encompassed by the 1974 Act. The Congressional Budget and Impoundment Control Act's reporting requirements then would have to be met.

This statutory analysis can be succinctly summarized by a short series of questions that should be considered by policy-makers contemplating a regulatory budget:

(1) Do the regulatory program funds constitute budget authority? If the answer here is yes, as it should be in most cases, then:

(2) Does the regulatory budget seem to mandate program cuts or reductions? If yes, then:

(3) Does the agency have the authority to expend its funds for other authorized purposes? If yes, then:

(4) Will the agency actually expend the funds for other authorized purposes?

If the answer to all of these questions is yes, then a special message will not have to be sent to Congress. Negative answers will either establish that the 1974 Act is not applicable ("no" to (1)), that the regulatory budget does not trigger the 1974 Act ("no" to (2)), or that a message must be sent to Congress under the provisions of the 1974 Act ("no" to (3) or (4)).

Two illustrations suggest the usefulness of this model of statutory analysis. The Environmental Protection Agency's (EPA's) total appropriation for the abatement and control of pollution was nearly $693 million for fiscal year 1980.[100] Assume that $150 million of that amount is designated specifically for regulatory programs. Under the model analysis:

(1) Do the regulatory program funds constitute budget authority?

As suggested above, such appropriations would constitute budget authority under the 1974 Act.

(2) Does the regulatory budget seem to mandate program cuts or reduction?

Assume here that a regulatory budget of $5.5 billion has been established for these EPA programs. Assume further that the

---

according to the Comptroller General, these breakdowns are not binding on administrative officers where they are not carried into the appropriation act itself. *See* 17 COMP. GEN. 147, 150 (1937).

100 Pub. No. L. 95-392, tit. II, 92 Stat. 791 (1978).

EPA reasonably estimates that an expenditure of $100 million in program funds will impose $5.5 billion in compliance costs on non-federal entities. Thus, the EPA will have $50 million in funds earmarked for regulatory programs that cannot be expended due to the regulatory budget.

(3) Does the agency have the authority to expend the funds precluded from regulatory use, for other authorized purposes? The regulatory program of EPA is funded via a lump-sum appropriation, but its appropriation for regulatory programs comes under a line item labelled "abatement and control." Thus, the EPA would have the authority to expend or obligate the $50 million only for other authorized purposes within the scope of program authority contained in that line item. For example, the funds could be used instead to provide certain types of additional technical assistance to federal, state, local, or interstate and private entities, functions that are carried out under that "abatement and control" heading.[101] It could *not* be transferred to a municipality for the construction of a waste water treatment facility, a function that usually is carried out under the appropriations line item heading of "construction grants."[102]

(4) Will the agency actually expend the funds for other authorized pruposes?

Assume that all authorized purposes under the heading "abatement and control," except regulatory programs, have full funding available. Since the $50 million will not be expended for other authorized purposes, Congress must be notified under the 1974 Act.[103]

Consider another hypothetical, involving the Coast Guard.

(1) Do the regulatory program funds constitute "budgetary authority?"

---

101 This assumes that these sources of additional assistance either are not deemed to be regulatory under the regulatory budget legislation or do not impose any further compliance costs on the non-federal sector.

102 *See* H.R. Rep. No. 1255, 95th Cong., 2d Sess. 26, 29 (1978).

103 The Comptroller General's office has noted:

> If the Congress desires to restrict the availability of a particular appropriation to the several items and amounts thereof submitted in the budget estimates, *such control may be effected by limiting such items in the appropriation act itself.* (Emphasis added.)

Letter from Frank H. Wertzel, Assistant Comptroller General to Warren Olney III, Director of the Administrative Office of the United States Courts (June 27, 1962)(Doc. No. B-149163). (on file at the Harvard Journal on Legislation).

As suggested above, its appropriation probably would constitute budgetary authority. That appropriation is assumed to be $10 million for the fiscal year ending September 30, 1979, for "regulatory programs and functions."[104]

(2) Does the regulatory budget seem to mandate program cuts or reduction?

Assume that the agency receives a regulatory budget of $100 million and reasonably estimates that an expenditure of $8.5 million of regulatory program funds will impose the maximum $100 million in compliance costs on non-federal entities.

(3) Does the agency have the authority to expend the funds precluded from regulatory use for other authorized purposes? In this illustration the *only* "authorized purposes" are "regulatory programs and functions." If the expenditure of $8.5 million results in meeting the regulatory budget ceiling, and there is no way in which the remaining $1.5 million may be expended for "regulatory programs and functions," then the Coast Guard does not have the authority to expend the funds for any other purpose. However, unexpended regulatory funds might still be used for regulatory functions which do not impose costs on the private sector such as training programs for regulatory personnel or advertising programs to increase public awareness of regulatory requirements. When used in this way, no message to Congress is necessary.

Thus, with the appropriation itself limited to regulatory programs, an inability to expend those funds for regulatory programs will preclude their use elsewhere, and the President must so inform Congress under the provisions of the 1974 Act. There is no need to advance to issue (4) here.

These examples were, out of necessity, somewhat simplified. In practice, the designation of amounts for regulatory programs may be very difficult to ascertain, even at the committee level. The examples show that each individual case of a regulatory budget that causes the cancellation or curtailment of an agency's regulatory program will have to be closely examined to determine whether any action will be required under the 1974 Act. In some instances, holding back program funds in order to

---

104 33 U.S.C. § 157(a) (1976).

comply with the regulatory budget will result in an "impoundment" or "deferral." In order to coordinate and evaluate impact of a regulatory budget on the responses required under the 1974 Act, some "fine tuning" of the regulatory budget legislation or implementing executive action will be required. There are three possible approaches:

(1) Determine that in all cases in which program funds must be withheld due to limits imposed by a regulatory budget, such an action *will not* constitute an impoundment or deferral for purposes of the 1974 Act. This can be accomplished with appropriate language in the regulatory budget legislation.

(2) Determine that in all cases in which program funds must be withheld due to limits imposed by a regulatory budget, such an action *will* constitute an impoundment or deferral for purposes of the 1974 Act. This also can be accomplished through appropriate language in the regulatory budget legislation.

(3) Determine that in all cases in which program funds are withheld due to limits imposed by a regulatory budget, such an action *will not* constitute an impoundment or deferral for purposes of the 1974 Act; the covered agency, however, could be required to inform the appropriate committees of Congress, where it otherwise would have to send a special message to Congress. In fact, this third approach currently is being followed where lump-sum appropriations are utilized in a manner different from that intended by Congress under the 1974 Act. The Comptroller General has held that

> If an agency finds it is desirable or necessary to take advantage of that flexibility by deviating from what Congress had in mind in appropriating particular funds, the agency can be expected to so inform Congress through recognized and accepted practices.[105]

This approach can be implemented either through appropriate language in the regulatory budget legislation, or through a combination of legislation and executive order. This approach represents a compromise, balancing disclosure with flexibility. The choice among these approaches is a policy question which Congress should consider and make explicit.

---

105 55 COMP. GEN. 307, 318 (1975).

Conclusion

The concept of a regulatory budget is likely to be the subject of significant and continuing debate among federal policymakers. That debate should focus on the likelihood of success of a regulatory budget in providing a handle with which to control the level of regulation-imposed costs and in meeting constitutional and statutory dilemmas. If policymakers decide to develop a regulatory budget, the budgetary mechanism should provide for close relations between the legislative and executive branches, and the enabling legislation should explicitly resolve as many of the potential statutory conflicts as possible.

NOTE

PROTECTING PRIVATE EMPLOYEES' FREEDOM OF POLITICAL SPEECH

MARK T. CARROLL*

Two of the most highly valued freedoms in America's democratic capitalist society are the citizen's political freedom to speak his mind on issues and candidates and the businessman's economic freedom to run his business in whatever way he thinks best. In the vast majority of jurisdictions across the country, the citizen's freedom of political speech is subordinated to the businessman's managerial discretion: the general rule is that except where otherwise provided by contract, a private employer can fire employees without having to offer any reason and so, in effect, remains free to fire employees on account of political views of theirs he does not like.

While recognizing the need for employers to be able to prevent their employees from behaving in ways that disrupt business operations, Mr. Carroll argues in this Note that employees and their political freedom of speech should be protected from politically motivated firings, and advocates the adoption of either a model statute or common-law rules that would prevent employers from firing employees for non-disruptive political activity.

I. INTRODUCTION

> Those who won our independence believed that the final end of the State was to make men free to develop their faculties . . . . They valued liberty both as an end and as a means. They believed liberty to be the secret of happiness and courage to be the secret of liberty. They believed that freedom to think as you will and to speak as you think are means indispensable to the discovery and spread of political truth; that without free speech and assembly discussion would be futile; that with them, discussion affords ordinarily adequate protection against the dissemination of noxious doctrine; . . . that public discussion is a political duty; and that this should be a fundamental principle of the American government.

Mr. Justice Brandeis concurring in *Whitney v. California*.[1]

---

* B.A., Columbia University, 1977; J.D., Harvard Law School, 1980; associate lawyer with the law firm of Duane, Morris & Heckscher, Philadelphia.

1 274 U.S. 357 (1927) (footnotes omitted).

Freedom of political speech — the freedom to express one's views on how society should be organized, on what the government should do, and on who should serve in the government — is widely and properly recognized as a linchpin of a democracy. Under the First Amendment to the U.S. Constitution, the freedom to express one's political views receives stringent protection from governmental interference. That freedom, however, receives only negligible legal protection from interference by private parties. Except where employment contracts provide otherwise, a private employer generally can fire an employee for almost any reason, including the employee's expressing his political views in a manner the employer does not like. For tens of millions of persons who are genuinely dependent on private employers — dependent in the sense that they simply cannot afford to lose their current jobs — freedom of political speech can be exercised only subject to the forbearance of their employers. As economic power becomes increasingly concentrated in the hands of large corporations, citizens' dependence on private employers grows and so does the insecurity of free political speech.

The importance of free speech can hardly be overstated. To the extent that what one knows and believes hinges on what opportunities he has had to exchange ideas with others, freedom of speech is vital to free intellectual development;[2] it is a prerequisite to freedom of thought.[3] Politically, freedom of expression is essential to the conduct of a democratic government,[4] where the "citizen-critic"[5] polices the actions of government and suggests reforms. The Supreme Court once characterized the First Amendment as "the matrix, the indispensable condition, of nearly every other form of freedom."[6]

Despite their overwhelming importance, the rights encompassed within the First Amendment are given constitutional protection from only one source: governmental interference.

---

2 T. EMERSON, THE SYSTEM OF FREEDOM OF EXPRESSION 6 (1970).

3 *See* J. MILL, *On Liberty*, ch. 2, in THE PHILOSOPHY OF JOHN STUART MILL 203-48 (M. Cohen ed. 1961). Mill argued that freedom of expression is necessary to find and refine truth.

4. T. EMERSON, *supra* note 2, at 6. *See* A. MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT (1948).

5 New York Times v. Sullivan, 376 U.S. 254, 282 (1964).

6 Palko v. Connecticut, 302 U.S. 319, 327 (1937).

The First Amendment explicitly prohibits Congress from infringing the freedom of expression:

> Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press, or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.[7]

The Supreme Court has always interpreted the First Amendment's restriction on public power to apply to the entire federal government, not just to Congress. Since the 1920's, the Court has held that the First Amendment applies to state governments by way of the Fourteenth Amendment.[8] However, no private citizen or corporation is required under the Constitution to respect a citizen's right of free speech.

Occasionally, the actions of a private corporation are deemed to involve "state action," subject to First Amendment restraints. For example, the corporation, through its business operations, might be particularly intertwined with the government;[9] or facilities operated by the private corporation might take on a functional resemblance to those normally operated by a public entity.[10] But such relationships are rarely found, and the recent trend in the Supreme Court has been to limit the "state action" and "public function" analysis that would

---

7 U.S. CONST. amend. I.

8 Stromberg v. California, 283 U.S. 359, 368 (1931); Gitlow v. New York, 268 U.S. 652, 666 (1925).

9 In Holodnak v. Avco Corp., 514 F.2d 285 (2d Cir.), *cert. denied*, 423 U.S. 892 (1975), the defendant, a private defense contractor, had formed a "symbiotic relationship" with the federal government: all the land, all the buildings, and most of the equipment used in the defendant's operations were government-owned, and the defendant's main business was producing defense materiel for the armed forces. *Id.* at 289. The Court held that since the government's involvement in the defendant's operations was so great, the defendant had "contractually assumed" an obligation to respect the First Amendment rights of its employees. *Id.* at 292.

10 In Marsh v. Alabama, 326 U.S. 501 (1946), the Gulf Shipbuilding Corporation owned a "company town" that, but for private ownership, had "all the characteristics of any other American town." *Id.* at 502. Town managers appointed by the corporation forbade the distribution of certain religious literature in the town. The Court held that because the operation of the town was essentially a "public function," the rights of the owner to control speech and activity in the town were "circumscribed by the statutory and constitutional rights" of those who visit or live in the town. *Id.* at 506. The Court invalidated a state statute that criminally punished persons who violated the town managers' prohibition of religious literature; the Court held that the state's action through this statute to enforce the prohibition violated the First and Fourteenth Amendments. *Id.* at 508.

otherwise extend First Amendment duties to private entities.[11] Absent a statutory[12] or common-law rule, there is nothing to restrict private persons from using their economic leverage to inhibit the free speech of others.

This narrowness in the scope of First Amendment restrictions was understandable when the Bill of Rights was adopted, but it is no longer appropriate today. Organized government was, to the Founding Fathers, uniquely powerful as a coercive force. Private entities or individuals generally lacked sufficient economic power to control the activities of large numbers of people.[13] But the decentralized, entrepreneurial economy of eighteenth-century America has been replaced by a largely centralized, corporate economy. Private corporations today rival in size any societal structures with which the Founding Fathers were familiar.[14] Many individual corporations employ thousands — even tens of thousands — of people. Using their enormous economic leverage as major employers, such corporations can unilaterally exert strong, even coercive influence over their employees. Employees' vulnerability is heightened by employees' limited mobility[15] and by an employee's need to keep a "clean" record if he is to be able ever to obtain another job.[16]

---

11 *See* Hudgens v. NLRB, 424 U.S. 507, 520 (1976) (rejecting the "functional similarity" of a shopping mall to a municipality for First Amendment purposes); Flagg Bros. v. Brooks, 436 U.S. 149, 159-60 (1978) (pointing out that the "public function" doctrine has been limited to cases where the private entity has, in some way similar to the state, exercised an exclusive control over important human activities or opportunities). *But cf.* Peck, *Unjust Discharges From Employment: A Necessary Change in the Law*, 40 OHIO ST. L.J. 1, 13-26 (1979) [hereinafter cited as Peck], arguing that governmental regulation of business is so pervasive that employer decisions to terminate employees in some way involve state action.

12 *See, e.g.*, NLRB v. Edward G. Budd Co., 169 F.2d 571, 577 (6th Cir. 1948), *cert. denied*, 335 U.S. 908 (1949), where the Court pointed out that the First Amendment's freedom of assembly protects the organizing activities of private employees against "*governmental* infringement" (emphasis in original). But prior to the National Labor Relations Act, 29 U.S.C. §§ 151-169 (1976), there was no federal restriction on a private employer's power to interfere with these rights by discharging employees.

13 *See* D. EWING, FREEDOM INSIDE THE CORPORATION 11-12 (1977) [hereinafter cited as EWING]; Note, *California's Controls on Employer Abuse of Employee Political Rights*, 22 STAN. L. REV. 1015, 1018-19 (1970).

14 *See, e.g.*, EWING, *supra* note 13, at 12-14, comparing the total populations of the thirteen colonies to the estimated workforces of the thirteen largest corporate employers in America.

15 Blades, *Employment At Will vs. Individual Freedom: On Limiting the Abusive Exercise of Employer Power*, 67 COLUM. L. REV. 1404, 1405 (1967) [hereinafter cited as Blades].

16 Peck, *supra* note 11, at 4-6.

The existence of such broad economic power in individual employers raises the same potential for abuse that the Founding Fathers feared in the government's coercive power: the suppression of critical or unpopular speech. That corporations do indeed abridge employees' speech through subtle and not-so-subtle means has been argued and proven in many contexts.[17] Corporate coercion has been facilitated by the broad common-law rule that an employment relationship is "at will," unless otherwise specified, and as such can be terminated at any time for any reason.[18]

Of course, employers' power to control the work and nonwork activities of their employees through discharge or other economic penalties has been restricted significantly in the twentieth century. The most dramatic example of such a restriction by a legislature is the National Labor Relations Act,[19] which guarantees the rights of workers to organize unions and to bargain collectively on the terms and conditions of employment.[20] Largely as a result of union activity protected by the Act, many employees are now covered by collective-bargaining agreements that restrict employers' common-law discretion by forbidding employers from arbitrarily dismissing or disciplining employees.[21] Aside from the National Labor Relations Act and collective-bargaining agreements, there exist statutory limits on employers' discretion to penalize or discharge employees whose ideas the employer finds distasteful. Many states have criminal statutes protecting employees' voting choices and other facets of employees' political speech from employer coercion.[22]

In some states, judicial decisions reinterpreting the common law have reduced employers' traditional freedom to dismiss

---

17 *See generally* EWING, note 13 *supra*. For a particularly blatant case of employer retaliation against an employee for the employee's political views, *see* Spokane-Idaho Mining Co., 9 Lab. Arb. 749 (1947) (Cheney, Arb.) (employee discharged for membership in Communist Party and for single instance of distributing political literature at company's mine). *See also* Blades, *supra* note 15, at 1410.

18 *See* text accompanying notes 31 to 47 *infra*.

19 29 U.S.C. §§ 151-169 (1976).

20 *Id.* at § 157.

21 *See* [1976] 2 COLLECTIVE BARGAINING NEGOTIATIONS AND CONTRACTS (BNA) § 40:1 (Dec. 28, 1978). Regarding the just-cause clause in collective-bargaining agreements, *see* text accompanying notes 133 to 210 *infra*.

22 Regarding state statutes, *see* text accompanying notes 117 to 132 *infra*. For a collection of these statutes, *see* 1 LAB. L. REP. (CCH) § 1521.

employees arbitrarily.[23] Perhaps the most famous of these decisions is *Monge v. Beebe Rubber Co.*,[24] which held that an employer's bad-faith, retaliatory, or malicious termination of an at-will employee is contrary to the public interest and constitutes a breach of the employment contract.[25]

Even where applied, these limitations on employer power are grossly inadequate to protect the political speech of employees. In 1976, less than twenty-two percent of the American workforce was organized[26] and was therefore subject, under collective-bargaining agreements, to some measure of protection from arbitrarily imposed penalties. Moreover, the National Labor Relations Act merely protects for private, non-agricultural employees certain associational interests that are based on First Amendment freedoms.[27] The Act neither affords all American workers this protection nor offers protection to any employee for speech that is purely political, unrelated to labor organizing or the bargaining process.[28] The inadequacies of state statutory and common law in protecting employee political speech will be discussed more fully below.[29]

Protection and fostering of political discussion is at the very core of the First Amendment. The threat that private economic power poses to these ideals is very real. This Note concerns the clash between the rights of employers to run their businesses as they think best and the rights of privately employed citizens to express their political views freely. It will examine the limited legal protection available to political speech in the private employment relationship, consider the degree to which employer control over employee speech is necessary for the sake of efficient business operations, and discuss possible legal reforms that would accommodate the competing interests of free political speech and economic efficiency.[30]

---

23 *See* text accompanying notes 37 to 67 *infra*.

24 114 N.H. 130, 316 A.2d 549 (1974).

25 114 N.H. at 133, 316 A.2d at 551.

26 UNITED STATES DEPT. OF COMMERCE, BUREAU OF THE CENSUS, STATISTICAL ABSTRACT OF THE UNITED STATES 1977, at 387, 418, tables 625 and 678 (98th ed. 1977).

27 29 U.S.C. §§ 151-169 (1976).

28 *See* text accompanying notes 72 to 82 *infra* for a discussion of Eastex v. NLRB, 437 U.S. 556 (1978).

29 *See* text accompanying notes 65 to 67, 126 to 132 *infra*.

30 In focusing on the protection of political expression, this Note excludes from

II. Current Legal Treatment of Employees' Political Speech

Legal protection of employees' freedom of political speech has developed on three fronts: in the common law, in statutory law, and in collective-bargaining agreements. On all three fronts, the protection remains limited and unreliable.

A. *The Common Law*

At common law, "unless a definite period of service is specified in the [employment] contract, the hiring is at will, and the master may discharge and the employee may leave at any time."[31] This right of either party to terminate the employment relationship has been described as "absolute."[32] Indeed, any reason can be used to terminate the at-will employment relationship, and no reason need be given.[33] Under the traditional common law of both contract and torts, the employee at will has no cause of action against his employer for even the most arbitrary of dismissals.[34] This general rule is very broadly accepted in most jurisdictions throughout the country.[35]

---

consideration three other areas of expression that the First Amendment protects from governmental interference: (1) "obscenity," "fighting words," and other expression that is "devoid of ideas"; (2) employee expression of religious beliefs; and (3) employee speech directed toward union organizing or bargaining activity. The issue of obscenity is excluded because it involves a distinct set of arguments that cannot be transferred to the issue of political speech. Employers' curtailment of profane or offensive language can be justified in terms of the need for employees to treat employer supervisors and fellow workers with respect in order for an enterprise to function. *See* Getman, *What Price Employment?: Arbitration, the Constitution, and Personal Freedom*, in PROCEEDINGS OF THE 29TH ANNUAL MEETING OF THE NATIONAL ACADEMY OF ARBITRATORS 61, 81 (BNA 1976) (comment by Professor Dunsford) [hereinafter cited as Getman]. The issues of religious speech and of speech directed toward union organizing and bargaining are excluded because federal statutes to a large extent already protect those types of speech from employer interference. *See* National Labor Relations Act, 29 U.S.C. §§ 151-169 (1976); Civil Rights Act of 1964, 42 U.S.C. § 2000e-2 (1976).

31 Watson v. Gugino, 204 N.Y. 535, 541, 98 N.E. 18, 20 (1912).

32 204 N.Y. at 542, 98 N.E. at 20.

33 Larsen v. Motor Supply, 117 Ariz. 507, 507, 573 P.2d 907, 907 (Ct. App. 1977); Smith v. Arthur C. Baue Funeral Home, 370 S.W.2d 249, 252 (Mo. 1963).

34. Geary v. U.S. Steel, 456 Pa. 171, 181, 319 A.2d 174, 180 (1974). *See also* WILLISTON ON CONTRACTS, § 1017 (3rd ed. 1967); RESTATEMENT OF TORTS, § 762 (1939).

35 The general rule that an at-will employment contract is terminable by the employer for any reason without liability in damages has been stated in literally scores of cases. *See*, in addition to the cases already cited, Hinrichs v. Tranquilare Hospital, 352 So.

The purpose of the at-will rule is to prevent the unreasonable tying of an employee to a particular employer. The at-will rule leaves the employee free to quit his job at any time, unless he chooses to bind himself to work for a specified period. In accordance with the mutuality-of-obligation doctrine, the rule leaves the employer similarly free to terminate employment.[36]

Ironically, the at-will rule was originally thought to benefit employees by preserving their legal right to seek better wages and working conditions with new employers.[37] In 1912, in one of the earliest at-will employment cases, *Watson v. Gugino*,[38] the rule so operated to allow an employee to quit his at-will job without being liable in damages.

Nevertheless, a distinctive trait of the discharged employee in the *Watson* case points out the outmoded nature of the rationale behind the at-will rule: in *Watson*, the departing employee was in fact an entrepreneur in his own right, who ran his own business before and after leaving the employ of the plaintiff.[39] Such individualistic business enterprise is effectively beyond the reach of most employees today; indeed, it is uncommon for employees to have talents or skills that are sufficiently marketable to enable a worker to move easily from one wage-earning job to another.[40] In such an economic environment, the at-will rule operates almost exclusively to benefit employers by preventing employee recovery of damages for wrongful discharge.[41] Moreover, courts have been reluctant to infer from the usual employment contract any intention of the parties to extend the contract for a definite time or to make it terminable only upon some dereliction of the employees.[42]

2d 1130, 1131 (Ala. 1977); Wynne v. Ludman, 79 So. 2d 690, 691 (Fla. 1955); Jackson v. Minidoka Irrig. Dist., 98 Ida. 330, 333, 563 P.2d 54, 57 (1977); Roemer v. Zurich Ins. Co., 25 Ill. App. 3d 606, 610, 323 N.E.2d 582, 585-86 (Ct. App. 1975); Lorson v. Falcon Coach, 214 Kan. 670, 679, 522 P.2d 449, 457 (1974). *See* Annot., *Employment Contracts Terminable at Will*, 62 A.L.R.3d 271, 271-73 (1975).

36 *See* CORBIN ON CONTRACTS, § 684 (1960). *See also* Summers, *Individual Protection Against Unjust Dismissal: Time for a Statute*, 62 VA. L. REV. 481, 484-85 (1976) [hereinafter cited as Summers]; Blades, *supra* note 15, at 1419.

37 Blades, *supra* note 15, at 1416-18.

38 204 N.Y. 535, 98 N.E. 18 (1912).

39 204 N.Y. at 536-38, 98 N.E. at 18-19.

40 Blades, *supra* note 15, at 1416.

41 Almost every modern case invoking this rule involves a discharged employee seeking damages from his employer.

42 *See, e.g.*, Watson v. Gugino, 204 N.Y. 535, 542, 98 N.E. 18, 20 (1912); Chin v.

An example of the common-law rule's harshness toward employees is *Bell v. Faulkner*.[43] The plaintiff claimed that his former employer tried to coerce him into voting for certain candidates and pressuring his family to do the same, and that the employer fired him for refusing. The court reversed a verdict of $500 damages for the plaintiff, holding that since the employment was a contract at will, the plaintiff could be discharged at any time, for any reason.[44] In addition, the court refused to allow the plaintiff a civil cause of action under a Missouri statute that had made it a felony for an employer to use threats of discharge to coerce his employees into voting a certain way.[45] The court held that the legislature, while providing for criminal prosecution, had declined to create any civil remedies for employees.[46]

As in *Bell,* courts across the country generally follow the common-law rule that employment contracts are terminable at will unless their terms explicitly provide otherwise and that employees under such contracts can be dismissed for any reason. Courts generally have refused to make any exceptions to the rule unless specifically directed by statute to do so.[47]

However, courts in a small but growing minority of states have ameliorated the harshness of the common-law rule toward employees through liberal, policy-oriented construction of statutes. These courts have allowed employees to bring civil actions against employers for firings that involved efforts by the employers to contravene a statutorily enacted public policy. Such courts have commonly allowed employees to collect compensatory damages and occasionally have granted punitive damages as well.

The first such case was the 1959 decision by a California court of appeals in *Petermann v. International Brotherhood of Teamsters*.[48] The plaintiff had been employed as a business

---

American Telephone & Telegraph Co., 410 N.Y.S.2d 737, 739, 96 Misc. 2d 1070, 1072-73 (Sup. Ct., Spec. Term, N.Y. County 1978), *aff'd mem.*, 70 App. Div. 2d 791 (App. Div. 1979); Summers, *supra* note 36, at 488.

43 75 S.W.2d 612 (Ct. App. Mo. 1934).

44 *Id.* at 613.

45 *Id.* at 614. Concerning similar statutes, *see* text accompanying notes 115 to 124 *infra*.

46 75 S.W.2d at 614.

47 *See generally* cases cited at notes 35 & 42 *supra*.

48 174 Cal. App. 2d 184, 344 P.2d 25 (Ct. App. 1959).

agent by the defendant union. After refusing to obey his supervisor's order to testify falsely before a legislative committee, the plaintiff had been discharged. The court found that the employer's right to discharge at-will employees could be limited by considerations of public policy; the court determined the public policy of California to oppose both submission and solicitation of perjured testimony. Although the relevant statute explicitly provided only for criminal sanctions, the court construed the statute as a mandate for civil remedies as well. The court held that

> in order to more fully effectuate the state's declared policy against perjury, the civil law, too, must deny the employer his generally unlimited right to discharge an employee whose employment is for an unspecified duration, when the reason for the dismissal is the employee's refusal to commit perjury.[49]

The court found the plaintiff to have a valid cause of action for wages lost due to the wrongful discharge.

In *Monge v. Beebe Rubber Co.*,[50] decided in 1974, the Supreme Court of New Hampshire went much farther than the *Petermann* court did in departing from the common-law rule of at-will employment. The plaintiff, an at-will employee, alleged that she had been discharged for refusing to date her foreman. The court held that "a termination by the employer at will which is motivated by bad faith or malice or based on retaliation is not in the best interest of the economic system or the public good and constitutes a breach of the employment contract."[51] Thus, whereas the *Petermann* court departed from the at-will rule only on the basis that the employer had dismissed the plaintiff in an effort to frustrate public policy evinced by a specific criminal statute, the *Monge* court moved toward a rule that malicious, retaliatory, or bad-faith firings are in themselves violations of public policy. The *Monge* court inferred this policy from what it considered to be a need for balancing "the employer's interest in running his business as he sees fit . . . against the interest of his employee in maintaining his employment, and the public's interest in maintaining a proper balance between

---

49 174 Cal. App. 2d at 189, 344 P.2d at 27.
50 114 N.H. 130, 316 A.2d 549 (1974).
51 114 N.H. at 133, 316 A.2d at 551.

the two;"[52] the court did not infer this policy from any specific statute governing employer-employee relations. The *Monge* decision clearly supports the proposition that employers should not be able to dismiss employees in retaliation for political viewpoints the employer does not like. However, the precise scope of protection *Monge* offers to employees awaits clarification by the New Hampshire Supreme Court.[53]

In making public-policy exceptions to the at-will rule, courts outside of New Hampshire have followed the *Petermann* approach rather than the *Monge* approach: they have departed from the at-will rule to prevent employers from frustrating the public policies embodied in criminal statutes and in statutes designed to protect the rights of employees, but they have not moved to establish a broad common-law rule against bad-faith firings. Courts in at least two states have held that, for reasons of public policy, an employer may not discharge an employee for having served jury duty.[54] Courts in at least three states have held that employers' firings of employees in retaliation for filing of workmen's compensation claims are contrary to the policy embodied in the workmen's compensation system and so are illegal.[55] Furthermore, in the 1980 case of *Sheets v. Teddy's Frosted Foods, Inc.*,[56] the court held that an employee could not be discharged for complaining that the employer had violated a state labeling statute. In the 1978 case of *Trombetta v. Detroit, Toledo and Ironton R.R. Co.*,[57] the court held that an employee could not be fired for refusing to falsify information on a report submitted to state air-pollution-control authorities. Finally, in the 1978 case of *Harless v. First National Bank in Fairmont*,[58] the court held that an employee could not be dis-

---

52 *Id.*

53 *See* Pstragowski v. Metrolife Ins. Co., 553 F.2d 1 (1st Cir. 1977), a diversity case applying New Hampshire law. It is one of the few cases in New Hampshire to have dealt with the *Monge* rule.

54 Nees v. Hocks, 272 Or. 210, 536 P.2d 512 (1975); Reuther v. Fowler Williams, 255 Pa. Super. 28, 386 A.2d 119 (Super. Ct. 1978).

55 Kelsay v. Motorola, 74 Ill. 2d 172, 181-82, 384 N.E.2d 353, 357 (1978); Frampton v. Central Indiana Gas Co., 260 Ind. 249, 251-53, 297 N.E.2d 425, 427-28 (1973); Sventko v. Kroger, 69 Mich. App. 644, 647-49, 245 N.W.2d 151, 153-54 (Ct. App. 1976).

56 CONN L. J., Jan. 22, 1980, at 1.

57 81 Mich. App. 489, 265 N.W.2d 385 (Ct. App. 1978).

58 246 S.E.2d 270 (W.Va. 1978).

missed for complaining about the employer bank's transgressions against state and federal consumer credit laws.

In a number of cases where courts have refused to find a public policy consideration sufficient to justify a departure from the at-will rule, they have adverted to the possibility of making such a departure if offered a clearer and more specific public-policy basis for doing so.[59] Perhaps the most important of these cases is *Geary v. U.S. Steel,*[60] decided in 1974 by the Pennsylvania Supreme Court. The court phrased the possible exception primarily in terms of employer interference in employees' lives and alluded to public policy as almost an afterthought:

> It may be granted that there are areas of an employee's life in which his employer has no legitimate interest. An intrusion into one of these areas by virtue of the employer's power to discharge might plausibly give rise to a cause of action, particularly where some recognized facet of public policy is threatened.[61]

In recognizing public-policy exceptions to the at-will rule, some courts have relied on contract theory while others have relied on tort theory. In *Monge* as well as in the 1977 case of *Fortune v. National Cash Register Co.,*[62] for example, courts

---

59 Larsen v. Motor Supply, 117 Ariz 507, 508-09, 573 P.2d 907, 908-09 (Ct. App. 1977) (no clear public policy threatened by employer discharge when employee refused to take a voice stress test); Jackson v. Minidoka Irrig. Dist., 98 Ida. 330, 334, 563 P.2d 54, 58 (1977) (no public policy threatened when public employee is discharged for unauthorized use of public funds); Pierce v. Orthopharmaceutical Corp., 166 N.J. Super. 335, 342-43, 399 A.2d 1023, 1026-27 (Super. Ct. App. Div. 1979) (remand to trial court to see if a doctor's discharge for refusing to test a drug she believed to be unsafe violated public policy); Campbell v. Ford Industries, 274 Or. 243, 249-52, 546 P.2d 141, 145-47 (1976) (discharge for exercise of a statutory right of stockholder-employee to examine the employer's books does not violate public policy where the interest protected by the statute is a private, proprietary one rather than a public one).

60 456 Pa. 171, 319 A.2d 174 (1974).

61 456 Pa. at 184, 319 A.2d at 180.

62 373 Mass. 96, 364 N.E.2d 1251 (1977). In *Fortune*, the Massachusetts Supreme Court held that the parties' written contract of employment contained an implied covenant of good faith and that the defendant employer's having in bad faith discharged the employee from his position as a cash-register salesman constituted a breach of the contract. In so doing, the court said that it was "merely recognizing the general requirement in [Massachusetts] that parties to contracts and commercial transactions must act in good faith toward one another." 373 Mass. at 102, 364 N.E.2d at 1256. The only specific Massachusetts statutes the court cited for this "general requirement" were those dealing with good faith in contracts under the Uniform Commercial Code and in termination of motor-vehicle franchises. Citing *Monge* for the proposition that "in *every* contract there exists an implied covenant of good faith and fair dealing," the *Fortune* court explicitly stopped short of adopting such a broad rule but also expressed

in effect treated statutory duties and policies as implied parts of the employment contract: when employers disregarded those duties or acted to frustrate those policies, the courts found the employee plaintiffs to have a cause of action against their employers for breach of contract. Meanwhile, in *Nees v. Hocks*[63] and *Sheets v. Teddy's Frosted Foods, Inc.*,[64] for example, courts treated statutory duties and public policy not as implied terms of employment contracts but as declarations of rights vested in employees without reference to their contracts. When employees establish that there was "a demonstrably improper reason for dismissal, a reason whose impropriety is derived from some important violation of public policy,"[65] these courts allow the employees to recover damages in tort.

Of course, the possibility of collecting punitive damages is available to actions in tort but rarely to actions for breach of contract. Thus, an employee plaintiff at least theoretically is likely to win a larger settlement if he establishes that the employer's discharging him involved an intentional tort than if he only shows a breach of contract.

But as a practical matter, establishing a cause of action for violation of public policy is more difficult in courts that treat such violations as torts than in courts that treat such violations as breaches of contract. Under a contract theory, the plaintiff can recover merely by establishing that the terms of his contract — including terms implied by public policy — were violated; but a plaintiff cannot recover in tort unless he proves that his employer's violation of public policy stemmed from negligence or from specific intent to harm. Furthermore, when employers have asserted that they dismissed the employee plaintiffs for valid reasons, courts applying the tort theory have seemed more receptive to such a defense than have courts relying on the theory of implied contract. If the employer can show *any* justifiable reason for having fired the employee, he escapes liability under a tort theory, regardless of what other motivations the

---

no disapproval for such a rule; it limited its holding to the facts of the case. 373 Mass. at 100-05, 364 N.E.2d at 1255-58.

63 272 Or. 210, 536 P.2d 512 (1975).

64 CONN. L. J., Jan. 22, 1980, at 1.

65 *Id.* at 2.

employee plaintiff may have shown to have influenced the employer's actions.[66] But under a theory of implied contract, the plaintiff wins so long as he establishes that at least one of the reasons for his dismissal was antithetical to public policy.[67]

Thus, the recent judicial departures from the at-will rule are, for the most part, tied to specific public policies that are evinced in statutes. None, with the possible exception of *Monge v. Beebe Rubber Co.*, seems to support a broad prohibition against employer actions that are politically motivated or that have the effect of suppressing political speech. Employees' freedom of political speech will remain generally conditioned upon their employers' forbearance until such time as the *Monge* approach becomes ascendent in courts across the nation or until legislatures enact new statutes specifically protecting employees' political speech.

B. *Statutory Law*

1. Federal Statutes

Although no federal statute explicitly protects employee's political speech from interference by private employers, two federal statutes could plausibly be newly interpreted so as to provide expanded legal protection to employees' freedom of political speech. One of the statutes is the National Labor Relations Act. The other is 42 U.S.C. § 1985(c), originally enacted as part of the Civil Rights Act of 1871.

The National Labor Relations Act provides protection against employer retaliation for engaging in organizing and other union activities.[68] While these activities are protected by the so-called freedom-of-association clause of the First Amendment,[69] they

---

66 *See* Chin v. American Telephone & Telegraph Co., 410 N.Y.S.2d 737, 739-40, 96 Misc. 2d 1070, 1073-74 (Sup. Ct., Spec. Term, N.Y. County 1978), *aff'd mem.*, 70 App. Div. 2d 791 (App. Div. 1979).

67 In Monge v. Beebe Rubber Co., 114 N.H. 130, 316 A.2d 549 (1974), for example, the court did not require that the employer's bad-faith motivation to discharge be the exclusive reason behind the dismissal, although such a reading of the opinion is possible. 114 N.H. at 133, 316 A.2d at 551-52.

68 29 U.S.C. §§ 151-169 (1976).

69 *See* NLRB v. Edward G. Budd Mfg. Co., 169 F.2d 571, 577 (6th Cir. 1948), *cert. denied*, 335 U.S. 908 (1949).

are not truly "political speech" in the sense that term has been used throughout this Note. Organizing and bargaining activities do have societal impact, but they are mainly concerned with the betterment of individual workers through the adjustment of specific contractual relationships.

Yet the broad language of the National Labor Relations Act can be interpreted as applying to more than organizing and bargaining alone. The Act states: "Employees shall have the right to self-organization, to form, join, or assist labor organizations . . . and to engage in other concerted activities for the purpose of collective bargaining or other mutual aid or protection. . . ."[70] It is an "unfair labor practice for an employer to interfere with, restrain, or coerce employees in the exercise" of these rights.[71] To the extent that "mutual aid or protection" as used in the Act may include political speech that is not directed at unions or bargaining, the Act may be construed by courts to protect employees' political speech from employer interference.

Indeed, in 1978 in *Eastex v. NLRB*,[72] the Supreme Court took at least a preliminary step toward interpreting the "mutual aid or protection" clause of the National Labor Relations Act broadly enough for the Act to defend employees' political speech. Plaintiffs were union members whose employer had denied them permission to distribute a union newsletter in non-work areas on their own time.[73] The newsletter was divided into four sections. The first and fourth sections presented appeals to fellow employees to join and support the union. The second section urged the employees to oppose the incorporation of the Texas right-to-work statute into the new state constitution then being considered by the legislature. The third section attacked President Nixon's decision to veto an increase in the federal minimum wage. It asked workers to register to vote to "defeat our enemies and elect our friends."[74] The union had sought to distribute the newsletter in hopes of increasing membership and bargaining strength in preparation for contract negotiations with the employer. The union had filed an unfair-labor-practice

70 29 U.S.C. § 157 (1976).
71 29 U.S.C. § 158(a)(1) (1976).
72 437 U.S. 556 (1978).
73 *Id.* at 560-61.
74 *Id.* at 559-60.

charge with the National Labor Relations Board (NLRB), which had gone on to adopt an administrative law judge's findings and conclusions that under section 7 of the National Labor Relations Act all the material in the newsletter was protected as concerted activity for the "mutual aid or protection" of employees.[75] Furthermore, said the Board, there were no "special circumstances" that would have justified the employer's refusal to allow distribution of such protected material in non-work areas.[76]

The Court stated that the "mutual aid or protection" clause is not narrowly limited to specific disputes or issues between an employee and his particular employer. It covers support for other employees of other employers, as well as methods for improving terms and conditions of employment "through channels outside the immediate employer-employee relationship."[77] The Court went on to state that there are limits to the scope of the "mutual aid or protection" clause, but it declined to clarify these limits:

> It is true, of course, that some concerted activity bears a less immediate relationship to employees' interests as employees than other such activity. We may assume that at some point the relationship becomes so attenuated that an activity cannot fairly be deemed to come within the "mutual aid or protection" clause. It is neither necessary nor appropriate, however, for us to attempt to delineate precisely the boundaries of the "mutual aid or protection" clause. That task is for the Board to perform in the first instance as it considers the wide variety of cases that come before it.[78]

At this point, the Court cited as examples two NLRB cases. In the first, *Ford Motor Co.*,[79] the NLRB held that the distribution of a newsletter criticizing the United Auto Workers for its support of any traditional political party and calling for a worker's government was a "purely political tract" that was not protected under the National Labor Relations Act.[80] In *Ford*

---

75 *Id.* at 561-62.
76 *Id.*
77 *Id.* at 565.
78 *Id.* at 567-68.
79 221 N.L.R.B. 663 (1975), *enforced,* 546 F.2d 418 (3rd Cir. 1976).
80 221 N.L.R.B. at 666. This finding of the administrative law judge was adopted without change by the Board.

Motor Co. (Rouge Complex),[81] the Board's general counsel admitted before the administrative law judge that publications of the Revolutionary Communist Party, like the newsletter involved in *Ford Motor Co.*, were not protected materials under section 7 of the Act.[82]

The Court upheld the Board's determination that distribution of the second and third parts of the newletter — the parts urging opposition to the right-to-work amendment and attacking President Nixon's veto of the minimum-wage increase — constituted protected activity under the "mutual aid or protection" clause.[83] It was within the Board's discretion to find the proposed right-to-work amendment of concern to the employees as employees;[84] likewise, the agency was correct to hold that the union's concern for unorganized workers earning the minimum wage might someday win support among those workers for the union.[85] The Court concluded that, as protected activity, distribution of the newsletter could take place on the employer's non-working-area property under the facts of this case and the *Republic Aviation v. NLRB*[86] doctrine.[87]

Thus, both the Court and the NLRB are willing to read the "mutual aid or protection" clause broadly enough to encompass some political speech by employees. At the same time, as the two *Ford* cases demonstrate, the Court and the Board will

---

81 233 N.L.R.B. 698 (1977).
82 *Id.* at 705.
83 437 U.S. at 569-70.
84 *Id.* at 570.
85 *Id.* at 569-70.
86 324 U.S. 793 (1945).
87 437 U.S. at 570-76. The Court's opinion in *Eastex* cited *Republic Aviation* as having

> upheld the [NLRB's] ruling that an employer may not prohibit its employees from distributing union organizational literature in nonworking areas of its industrial property during nonworking time, absent a showing by the employer that a ban is necessary to maintain plant discipline or production. This ruling obtained even though the employees had not shown that distribution off the employer's property would be ineffective.

Id. at 570-71 (citations omitted).

In dissent, Justice Rehnquist, joined by Chief Justice Burger, contended that where the literature employees seek to distribute is not confined to narrow unionization issues, the employer has the right to prohibit distribution of the literature on his property. Their dissenting opinion says, "Congress never indented to require the opening of private property to the sort of political advocacy involved in this case." *Id.* at 580. The dissent added that no employee has "a protected right to engage in anything other than organizational activity on an employer's property." *Id.* at 583.

require that political speech bear a fairly close relationship to the immediate economic interests of "employee[s] . . . as employees."[88] Political speech aimed at broad social change, affecting employees beyond their work relationship as members of a political community, would probably fall outside section 7's protection.

The National Labor Relations Act and the *Eastex* decision give some guidance as to how closely related political speech must be to employee interests in labor conditions to be protected from employer retaliation. Unfortunately, there is less such guidance on 42 U.S.C. § 1985(c), a statute that, like the National Labor Relations Act, arguably protects employees freedom of political speech against interference by private employers.

Section 1985(c) provides in part:

> If two or more persons . . . conspire or go in disguise on the highway or on the premises of another, for the purpose of depriving, either directly or indirectly, any person or class of persons of the equal privileges and immunities under the laws; . . . the party so injured or deprived may have an action for the recovery of damages occasioned by such injury or deprivation, against any one or more of the conspirators.[89]

In 1971 in *Griffin v. Breckenridge*,[90] the Supreme Court held for the first time that Congress intended § 1985(c) to reach private conspiracies — conspiracies involving no state action or law.[91] However, the facts of the case (black citizens were assaulted by the defendants while travelling on interstate roads) involved two constitutional rights: the Thirteenth Amendment right to be free from slavery and involuntary servitude[92] and the right of interstate travel,[93] both of which by their terms are

88 *Id.* at 567.

89 42 U.S.C. § 1985(c) (1976). Until 1976, § 1985(c) was codified as 42 U.S.C. § 1985(3). Most of the cases cited below refer to this earlier numbering. Since there is no difference in language between the old and the new numberings, they are used interchangeably.

90 403 U.S. 88 (1971).

91 *Id.* at 101-02.

92 U.S. Const. amend. XIII provides: "Neither slavery nor involuntary servitude . . . shall exist within the United States. . . ."

93 The Constitution does not explicitly provide for any right of interstate travel, but the Supreme Court has held that right to be implied by the Constitution. *See, e.g.*, Shapiro v. Thompson, 394 U.S. 618, 629-31 (1969); United States v. Guest, 383 U.S. 745, 757-60 & n.17 (1966); Twining v. New Jersey, 211 U.S. 78, 97 (1908).

secured against private interference. The Court said that Congress had the constitutional authority to protect those rights through section 1985(c).[94]

But if the Court were going to extend section 1985(c)'s coverage to private conspiracies aimed at depriving others of constitutional rights, it also made clear that the statute should not become an all-purpose remedy for every tort with a conspiracy behind it. As the Court construed the statutory language,

> [t]he constitutional shoals that would lie in the path of interpreting [section 1985(c)] as a general federal tort law can be avoided by giving full effect to the Congressional purpose — by requiring, as an element of the cause of action, the kind of invidiously discriminatory motivation stressed by the sponsors of the [equal protection language in the section]. The language requiring intent to deprive of equal protection, or of equal privileges and immunities, means that there must be some racial, or perhaps otherwise class-based, invidiously discriminatory animus behind the conspirators' action. The conspiracy, in other words, must aim at a deprivation of the equal enjoyment of rights secured by the law to all.[95]

At the very least, then, a person alleging a private conspiracy to deprive him of equal protection or of Thirteenth Amendment rights has to allege an animus directed at him because of his membership in some "class." Whether the class must be racial remains unanswered by the Supreme Court, which nevertheless hinted in the quoted passage that it need not be.

But this issue of statutory construction can be resolved only after the constitutional authority question is answered: whether a constitutional right, limited by its own terms to state action,

---

94 403 U.S. at 104-06. After setting out these sources of constitutional authority for Congress' passage of § 1985(c) to cover this case, the Court stated:

> In identifying these two constitutional sources of Congressional power, we do not imply the absence of any other. More specifically, the allegations of the complaint in this case have not required consideration of the scope of the power of Congress under § 5 of the Fourteenth Amendment. By the same token, since the allegations of the complaint bring this cause of action so close to the constitutionally authorized core of the statute, there has been no occasion here to trace out its constitutionally permissible periphery.

Id. at 107 (footnotes omitted).

95 *Id.* at 102 (emphasis in original) (footnotes omitted); *but see id.* at 102 n.9: "We need not decide, given the facts of this case, whether a conspiracy motivated by invidiously discriminatory intent other than racial bias would be actionable under the portion of § 1985(3) now before us. . . ."

might nevertheless now be protected against private infringement by section 1985(c). If this were the effect of section 1985(c), then perhaps employer discharges of employees in retaliation for political speech (assuming all other elements of section 1985(c) are met)[96] might be actionable under the statute. Perhaps Congress has strengthened the right of free speech to stand against private as well as governmental infringement.

One of the very first cases to reach the federal courts of appeals after *Griffin* involved the discharge of an employee for political speech. In the 1971 case of *Richardson v. Miller*,[97] the plaintiff alleged that the defendants dismissed him in order to deprive him of equal protection of the law. The plaintiff, who was not black,[98] claimed he was discriminated against because he criticized the employer's racially discriminatory hiring practices, advocated the election of federal candidates who were most likely to remedy such racially discriminatory practices, and in general called for equal opportunity in employment.[99] The Third Circuit Court of Appeals held that these allegations were "sufficient to constitute the 'racial, or perhaps otherwise class-based invidiously discriminatory animus' required by *Griffin*."[100] The court admitted that this was a close question, since the plaintiff himself was not a member of the class allegedly discriminated against. But in light of the policy of according civil rights statutes " 'a sweep as broad as [their] language,' " the court felt that a cause of action had been stated.[101] Apparently, the court believed that this particular use of section 1985(c) had a constitutional source of power in the Thirteenth Amendment, since racially motivated discrimination was involved.

Later the same year, the Eighth Circuit decided *Action v.*

---

96 This is no small problem for a prospective § 1985(c) plaintiff-employee. If, for example, the only defendants the employee could name in such a suit were the corporate employer and its agents, the "conspiracy" requirement of § 1985(c) would be a formidable obstacle. It is generally the law that a corporation cannot conspire with its agents. *See* Bellamy v. Mason's Stores, 508 F.2d 504, 508 (4th Cir. 1974) (Boreman, J., concurring in the result); *but cf.* Great American Federal Savings and Loan Association v. Novotny, 442 U.S. 366 (1979) (assuming a corporation can conspire with its directors without deciding the issue).

97 446 F.2d 1247 (3d Cir. 1971).

98 *Id.* at 1249.

99 *Id.* at 1247.

100 *Id.* at 1249.

101 *Id.*

Gannon,[102] in which Roman Catholic church members sought an injunction under section 1985(c) against two black civil-rights groups who had seriously disrupted religious services at the plaintiffs' cathedral. The black groups intended to disrupt the services as part of their protest demanding money and action for social reform from the church. The court first held that the defendants were "stimulated to disrupt the church services by racial and economic motives," in that they were attacking a predominantly white group with demands for aid to blacks.[103] Since the constitutional rights of the plaintiffs involved in this case were based not on either the Thirteenth Amendment or the right of interstate travel, but instead on the freedom-of-religion clause in the First Amendment, the court had to find some other constitutionally based source of power for Congress to be able to apply section 1985(c) to the defendants' private conspiracy. The court found that power in sections 1 and 5 of the Fourteenth Amendment.[104] Since the Fourteenth Amendment protects the rights guaranteed by the First and since section 5 of the Fourteenth Amendment empowers Congress to reach private conspiracies that interfere with rights protected under the Amendment, the court reasoned that Congress has the power to punish private conspiracies that infringe First Amendment rights.[105]

The jump in the *Action* court's reasoning was, of course, the assumption that section 5 of the Fourteenth Amendment authorizes Congress to reach purely private conspiracies to interfere with Fourteenth Amendment rights, which, like First Amendment rights, are by their terms protected only from governmental interference. The court based its reasoning on its interpretation of *U.S. v. Guest,*[106] a 1966 case in which six justices expressed the view that section 5 did give Congress the

---

102 450 F.2d 1227 (8th Cir. 1971).

103 *Id.* at 1232.

104 *Id.* at 1233; U.S. CONST. amend. XIV, § 1 provides: "No state shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any state deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws." Section 5 states: "The Congress shall have power to enforce, by appropriate legislation, the provisions of this article."

105 *Id.* at 1235.

106 383 U.S. 745 (1966).

power to reach private conspiracies to interfere with Fourteenth Amendment rights. Since *Guest* is so uncertain in meaning on this point,[107] the *Action* court's reliance on it is clearly the weakest point in the court's chain of logic. Nevertheless, *Action* is clear in holding that Congress has the constitutional authority to use section 1985(c) to protect freedom of religion and freedom of association from private interference.

But the result and reasoning in *Action* were rejected by the Fourth Circuit in the 1974 case of *Bellamy v. Mason's Stores.*[108] In *Bellamy* an employee brought a section 1985(c) suit against his employer, alleging that he had been discharged for membership in the Ku Klux Klan. The court held that no cause of action under section 1985(c) existed for the plaintiff:

> We think the language of equal protection chosen by the 1871 Congress cannot be interpreted to mean that persons who conspire without involvement of government to deny another person the right of free association are liable under this statute. This is so because the right of association derives from the first amendment — itself framed as a prohibition against the federal government and not against private persons, and because the incorporation doctrine has never been extended by the Supreme Court to apply to private persons.[109]

The court rejected *Action*'s use of the incorporation doctrine and *Griffin*'s interpretation of section 1985(c). While the First Amendment has been incorporated into the due process clause of the Fourteenth Amendment, freedom of association is protected through restraints only on state action; the court said that enlargement of the incorporation doctrine so as to restrict private action is "an innovation that must come from the Congress or the Supreme Court."[110] Thus, two circuits have come to diametrically opposite conclusions on the issue of whether rights of free speech can be protected against infringement by private persons under section 1985(c).

Adding weight to the *Bellamy* holding is the 1975 case of *Cohen v. Illinois Institute of Technology,*[111] the opinion for

---

107 *See* text accompanying notes 232 to 237 *infra.*
108 508 F.2d 504 (4th Cir. 1974).
109 *Id.* at 506-07.
110 *Id.* at 507.
111 524 F.2d 818 (7th Cir. 1975), *cert. denied,* 425 U.S. 943 (1976).

which was written by then-Judge Stevens for the Seventh Circuit shortly before his nomination to the Supreme Court. The plaintiff was a woman suing under section 1985(c) (and also under section 1983) for alleged sex discrimination. The court held that a conspiracy to deprive a person of a constitutional right that by its terms is protected solely against state interference is not actionable under section 1985 unless the conspiracy involves state action. In other words, Congress can penalize private conspiracies against constitutional rights only when the right is constitutionally protected against private action.[112]

This holding severely restricts the reach of section 1985(c) over private conduct, since there are only two such constitutional rights: the Thirteenth Amendment right to freedom from slavery and the right of interstate travel. Furthermore, only one of these two rights — the Thirteenth Amendment right — has been specifically held to be a basis for suits under section 1985(c).

It is possible, however, that involvement of the First Amendment in a section 1985(c) suit might broaden the First Amendment's own reach beyond state action. This possibility was raised by Judge Stevens himself in the *Cohen* case. While the equal protection clause of the Fourteenth Amendment might run only against state action,[113] Judge Stevens added that "[w]e have not been faced with the questions whether [this] rationale would apply to a right protected by the First Amendment which in terms is only a protection against state action, but which is often accorded special deference."[114]

Justice Stevens has since reiterated his view on rights protected only against state action.[115] Unless the rights protected by the Fourteenth Amendment are deprived by state action, he has said, there has not even been a constitutional violation in the context of section 1985(c). Such rights can be violated only by state action.

Thus, the question of whether Congress could constitutionally protect employees' rights of free political speech against in-

---

112 *Id.* at 828-29.

113 *Id.* at 828.

114 *Id.* at 829-30 & n.33 (citations to cases omitted).

115 *See* Great American Federal Savings and Loan Association v. Novotny, 442 U.S. 366, 383 (1979) (Stevens, J., concurring).

fringement by private employers is unclear. The circuits are split on the issue. Justice Stevens seems to doubt that Congress can constitutionally restrict private action for the sake of protecting rights other than the two specified above, but hints that further consideration would be necessary when the First Amendment is involved. No court has held that employee speech unconnected with civil-rights issues is protected under section 1985(c) against a private employer's retaliation. Perhaps in the future the Supreme Court will provide more guidance in this area. For now, section 1985(c) protects employees' freedom of political speech from interference by private employers only when the speech concerns racial issues and, at that, only in the Third Circuit.

2. State Statutes

At first glance, the outlook for legally protecting employee political speech seems brighter under state statutory law than under federal law. At least thirty-seven states and Puerto Rico have statutes protecting in some way the political activities or opinions of employees from coercion or retaliation by employers.[116] Some statutes, for example Louisiana's, broadly protect an employee's political activities from employer interference:

> No . . . employer shall adopt or enforce any rule, regulation, or policy which will control, direct, or tend to control or direct the political activities or affiliations of his employees by means of threats of discharge or loss of employment in case such employees should support or become affiliated with any particular political faction or organization, or participate in political activities of any nature or character.[117]

Many states have followed a more narrow formulation. For example, several states prohibit employers (1) from placing in pay envelopes political tracts that contain threats intended to influence the political actions of employees, (2) within ninety days of an election, from posting notices that threaten a factory closing, a wage reduction, or other penalties as the conse-

---

116 The states and the texts of their statutes are collected at [1980] 1 Lab. L. Rep. (CCH) ¶ 43.045.

117 LA. REV. STAT. ANN. § 23:961 (West 1964).

quences of a particular candidate's being elected or defeated, or that make threats in general intended to influence employee "political opinions or actions."[118] Such a set of prohibitions, being so specific, is clearly aimed at practices that perhaps at one time were quite common. It is clear, however, that they do not block all possible methods by which employers could coerce employees to act a certain way in the areas of political speech and political choice. Neither do the simpler, more general statutory formulations that prohibit discharge and threats of discharge "for the purpose of influencing" the employee's vote.[119] Even where states' statutory protections of employees' political speech apply, the resulting sanctions are typically too minor to serve as a major deterrent against employers' abuses: almost all of these statutory protections are backed only by criminal sanctions consisting of small fines or imprisonment of the employer's agents.[120]

Not only are the sanctions less than formidable when imposed under these state statutes, but few suits under those statutes even reach final decisions in court. Professor Summers has pointed out one reason for the lack of decided cases: since the statutes give no civil remedies to employees, the most interested parties in such cases have no recourse in the courts, and provide no source of litigated cases.[121] As the Illinois Supreme Court put it, in an analogous context, "the imposition of a small fine, enuring to the benefit of the State does nothing to alleviate the plight of those employees who are threatened with retaliation and forego their rights. . . ."[122] Not only are criminal prosecutions by the state left as the only source of litigation under these statutes, but prosecutors generally seem to lack enthusiasm for initiating cases under the statutes: there are few reported prosecutions of employers for coercing their employees.[123]

---

118 MD. ANN. CODE art. 33, § 26-16(a)(6) (1957). *Compare* N.Y. ELEC. LAW, § 150(3) (McKinney 1978), containing almost exactly the same language. *See also* R.I. GEN. LAWS § 17-23-6 (Supp. 1979).

119 *See, e.g.*, MICH. COMP. LAWS ANN., § 168.931(d) (1967).

120 *See, e.g.*, ARIZ. REV. STAT. ANN., § 16-304(B) (1975) ($5,000 fine); MD. ANN. CODE art. 33, § 26-16(b) (1957) ($1,000 fine, one year imprisonment, or both, plus a four-year ban on holding public office).

121 Summers, *supra* note 36, at 495.

122 Kelsay v. Motorola, 74 Ill. 2d 172, 185, 384 N.E.2d 353, 359 (1978).

123 *See* Vulcan Last Co. v. State, 194 Wisc. 636, 217 N.W. 412 (1928) (upholding

There are, however, some states that do provide more than minor criminal sanctions to protect employees. California, Colorado, Delaware, Louisiana, and Puerto Rico statutorily provide civil remedies for politically coerced employees.[124] Neither California nor Louisiana — the states whose statutes broadly proscribe employer interference with employee political opinions — has had a great deal of litigation under its statutes. The lack of litigation and the accompanying lack of court decisions under the statutes seriously inhibit the statutes' usefulness to employees: without litigation and the resulting judicial clarifications of the statutes, employees do not know what statutory protections they can assert, what defenses are open to employers, or what level of coercive intent on the part of employers is necessary for an employee to recover damages.

California's broadly worded statute protecting private employees' freedom of political speech includes the following:

> *Section 1101.*No employer shall make, adopt or enforce any rule, regulation or policy . . . (b) controlling or directing, or tending to control or direct the political activities of affiliations of employees.
>
> *Section 1102.* No employer shall coerce or influence or attempt to coerce or influence his employees through or by any means of threat or discharge or loss of employment to adopt or follow or refrain from adopting or following any particular course or line of political action or political activity.
>
> *Section 1105.* Nothing in this chapter shall prevent the injured employee from recovering damages from his employer for injury suffered through a violation of this chapter.[125]

Despite this broad language, the statute has not been read or used liberally, at least not until recently. It was first addressed

---

the conviction of a corporation for attempting to influence the votes of its employees by discharge and threats of discharge); Santiago v. People of Puerto Rico, 154 F.2d 811 (1st Cir. 1946) (upholding the conviction of a plantation supervisor who refused to give work to a day laborer because of the worker's affiliation with a particular political party).

124 CAL. LAB. CODE § 1105 (West 1971); COLO. REV. STAT. § 8-2-108(2) (1973) (action limited to violation of section preventing employer interference with employees' running for office); DEL. CODE ANN. tit. 15, § 5162 (1975) (action limited to voter coercion and $500 damages); LA. REV. STAT. ANN. § 23:961 (West 1964); P. R. LAWS ANN. tit. 29, § 136 (1966) (action limited to interference with political party affiliations).

125 CAL. LAB. CODE §§ 1101, 1102, 1105 (West 1971).

by the California Supreme Court in 1946, in *Lockheed Aircraft Corp. v. Superior Court.*[126] In that case, Lockheed, a defense contractor, fired eighteen employees whose "loyalty" it was unable to verify to its own satisfaction. The employees countered with a suit under the statute, claiming they had been discharged pursuant to Lockheed rules that wrongfully regulated their political activity.[127] Lockheed contended that the statute was unconstitutionally vague, because it could not determine from the language of section 1101 whether it was required to keep persons whose loyalty it was uncertain of, or persons who advocated the violent overthrow of the government, on its payroll. In identifying the scope of political activity protected by the statute, the court relied on a broad construction of the term "political," but, by judging political activity according to its connection with the "orderly conduct" of government, limited the scope of protection offered by the statute. The court cited a dictionary definition of the term "political":

> Of or pertaining to the exercise of the rights and privileges or the influence by which the individuals of a state seek to determine or control its public policy; having to do with the organization or action of individuals, parties, or interests that seek to control the appointment or action of those who manage the affairs of state.[128]

The court went on to hold:

> In each case the interference proscribed by the statute is interference with "political activities or affiliations," and the test is not membership in or activities connected with any particular group or organization, *but whether those activities are related to or connected with the orderly conduct of government and the peaceful organization, regulation and administration of government.*[129]

Under this test, nothing in the statute prevented a defense contractor from discharging an employee who advocates the violent overthrow of the government or whose loyalty "has not been established to the satisfaction of the employer." The case was

---

126 28 Cal. 2d 481, 171 P.2d 21 (1946).
127 28 Cal. 2d at 483, 171 P.2d at 23.
128 28 Cal. 2d at 484-85, 171 P.2d at 24.
129 28 Cal. 2d at 485, 171 P.2d at 24 (emphasis added).

thus remanded for trial.[130] The California Supreme Court seemed to have construed the statutory protection of political speech so narrowly as to enable employers to fire an employee simply for being a Communist.

In the 1979 case of *Gay Law Students Association v. Pacific Telephone and Telegraph*,[131] the California Supreme Court, in a 4-to-3 decision, indicated a new willingness to read the terms "political activity" and "political line" in a manner more protective of provocative employee political speech. These terms connote the espousal of a cause, not just partisan politics, the court said. As such a cause, the struggle of homosexuals for equal rights is a "political activity." If an employer discriminated against persons who identified themselves as homosexuals or with homosexual organizations, then the employer's action tends to "control or direct the political activities or affiliations of employees," and thus violates the statute.[132] It remains to be seen if this apparent widening of the protection for employees will continue in California.

In conclusion, it is clear that despite the broad language in some of the state statutes, those statutes do not seem to offer effective protection for the political speech of employees. The lack of civil remedies translates to a lack of incentive for victimized employees to bring suit under applicable state statutes. Even states that, like California, allow a civil cause of action for suppression of political speech have failed to define the protection they give employees specifically enough for the protection to be effective. As far as employees' freedom of political speech is concerned, state statutes as well as the federal statutes discussed above have in practice carved out only marginal exceptions to the traditional and still prevalent common-law rule that an employment contract is terminable at will unless the contract itself provides otherwise.

C. *Collective-Bargaining Agreements*

Especially in view of extant statutory and common law, the most important protection organized employees have against

---

130 *Id.* The case was apparently settled out of court, since no further proceedings are reported.

131 24 Cal. 3d 458, 595 P.2d 592, 156 Cal. Rptr. 14 (1979).

132 24 Cal. 3d at 487-88, 595 P.2d at 610-11.

arbitrary or maliciously-motivated discharges is the "just cause" clause of collective-bargaining agreements. This clause typically states that an employee will be discharged or disciplined only "for just cause," "for cause," "for proper and sufficient cause," or "for good and sufficient cause"; it is enforced by an agreed upon grievance procedure that usually culminates in binding arbitration.[133] It is estimated that eighty percent of the collective-bargaining agreements in the nation have such just-cause limitations on the employer's discharge powers.[134] Since "many arbitrators would imply a just-cause limitation in any collective agreement,"[135] and since an estimated ninety-five percent of all collective-bargaining agreements contain grievance and arbitration provisions,[136] most organized employees in America can claim the protection of a just-cause clause in the event they are discharged or disciplined.

In determining whether an employer's firing or other punishment of an employee has been for "just cause,"[137] arbitrators have typically looked outside the language of the employment contract to customary notions of fairness. On a general level, arbitrators have characterized "just cause" as "related to ordinarily accepted concepts of justice between man and man,[138] as "an equity notion,"[139] or simply as a rule of reason: "whether a reasonable man would find sufficient justification in an employee's conduct to warrant discharge."[140] On a more specific level, arbitrators have looked to criteria that have traditionally been widely accepted as fair, reasonable bases for firing or otherwise disciplining employees:

> [Phrases such as "just cause"] exclude discharge for mere whim or caprice. They are, obviously, intended to include

---

133 [1976] 2 COLLECTIVE BARGAINING NEGOTIATIONS AND CONTRACTS (BNA) § 40:1 (Dec. 28, 1978).

134 *Id.*

135 E. ELKOURI & F. ELKOURI, HOW ARBITRATION WORKS 611 (3d ed. 1973) [hereinafter cited as ELKOURI].

136 U.S. BUREAU OF LABOR STATISTICS, DEP'T OF LABOR, BULL. NO. 1425-1, MAJOR COLLECTIVE BARGAINING AGREEMENTS: GRIEVANCE PROCEDURES (1965), summarized in [1965] LAB. REL. Y. B. (BNA) 34-35.

137 Throughout this Note the term "just cause" is used in lieu of "proper and sufficient cause" and all similar terms. Arbitrators have found no substantial difference between the standards embodied in "just cause" and those embodied in the various related terms.

138 Bethlehem Steel Co., 24 Lab. Arb. 852, 855 (1955) (Desmond, Arb.).

139 Spokane-Idaho Mining Co., 9 Lab. Arb. 749, 751 (1947) (Cheney, Arb.).

140 Hearst Publishing Co., 30 Lab. Arb. 642, 644 (1958) (Schedler, Arb.).

> those things for which employees have traditionally been fired. They include the traditional causes of discharge in the particular trade or industry . . . They include such duties as honesty, punctuality sobriety, or, conversely, the right to discharge for theft, repeated absence or lateness, destruction of company property, brawling and the like.[141]

At the very least, the just-cause clause in collective-bargaining agreements limits an employer's ability to discharge his workers; a neutral arbitrator, judging in terms of fairness and reasonableness, determines the validity of the employer's conduct. This limitation on the employer's discharge power has often been applied in defense of employee's political speech and activities. The just-cause clause has been construed to prevent discharges for wearing anti-right-to-work-amendment T-shirts,[142] for posting announcements concerning a May Day march on Washington, D.C.,[143] for Communist Party membership,[144] for display of an Emilio Zapata poster bearing the words, "Viva la Revolución,"[145] for political fund-raising activities by employees,[146] and for refusal to testify before the House Un-American Activities Committee.[147] Both by striking down rules[148] and by invalidating individual discharges motivated by employer dislike of an employee's political stance, the just-cause clause has functioned as a check on an employer's power to retaliate against employees whose ideas he dislikes.[149]

In arbitration cases involving employees who have been fired for political activism, a major issue commonly has been whether

141 Arbitrator McGoldrick, quoted in ELKOURI, *supra* note 135, at 612.

142 R. H. Buhrke Co., 68 Lab. Arb. 1170 (1977) (Nicholas, Arb.).

143 Wright Machinery, 64 Lab. Arb. 593 (1975) (Calhoon, Arb.).

144 Foote Bros. Gear and Machine Corp., 13 Lab. Arb. 848 (1949) (Larkin, Arb.).

145 California Processors, 56 Lab. Arb 1275 (1971) (Koven, Arb.).

146 Baltimore News, 65 Lab. Arb. 161 (1975) (Block, Arb.). This case involved an "anti-discrimination" clause. The employer's rule against political fund-raising by newspaper employees was found to be too vague to be enforceable.

147 Republic Steel Corp., 28 Lab. Arb. 810 (1957) (Platt, Arb.).

148 Arbitrators will strike down pre-set rules if the employer's action under the rule was not effected for just cause. Thus, pre-set rules offer no protection to the employer's case if they are arbitrary or otherwise repugnant to the just-cause clause. *See* ELKOURI, *supra* note 135, at 641-46.

149 There are, of course, many cases where the arbitrator decides that the just-cause clause has *not* been violated when an employer discharges an employee who is engaged in political speech. *See, e.g.*, Firestone Tire and Rubber Co., 16 Lab. Arb. 569 (1951) (Cheney, Arb.), discussed in text accompanying notes 156 to 157 *infra*. The standards used to judge whether just cause exists for the discharge of an employee in a political speech case are discussed in text accompanying notes 161 to 207 *infra*.

the employer was motivated by dislike of the employee's political beliefs or by disruptions the employee caused in the employer's business. In the 1940's cases of *Spokane-Idaho Mining Co.*[150] and *Foote Bros. Gear and Machine Corp.*,[151] the employers cited the employees' affiliation with the Communist Party as justification for discharging them. The arbitrator in each case struck down the discharges for violating the just-cause clause of the collective-bargaining agreements.[152]

In contrast to the employers in *Spokane-Idaho Mining* and *Foote Bros.*, employers have fared well in arbitration proceedings when they have shown that the discharged employee's activities had seriously interfered with the employer's business operations. In the 1972 case of *Essex International*,[153] for example, the discharged employee had travelled around the plant during working hours, investigating safety conditions and passing out literature that criticized the employer and called for a thirty-hour work week for all industrial employees.[154] The employee claimed she was discharged for her political views. The arbitrator, however, accepted the employer's argument that the employee had seriously distracted other workers on company time and thereby had unduly disrupted plant operations.[155] A similar result was reached in 1951 in *Firestone Tire and Rubber Co.*,[156] which involved an employee who had frequently stated that the Soviet Union had a better system of government than the United States and that American troops did not belong in Korea. The arbitrator found the employee to have created unrest and possibly violence among other employees, leading to "lowered morale and lowered efficiency;"[157] on this basis, the arbitrator held the firing to have been for just cause. In the 1966 case of *Baltimore Transit Co.*,[158] the discharged employee, who had been employed as a bus driver, had been serving as the Acting Grand Dragon of the Maryland branch of the Ku Klux

---

150 9 Lab. Arb. 749, 750 (1947) (Cheney, Arb.).
151 13 Lab. Arb. 848, 849 (1949) (Larkin, Arb.).
152 9 Lab. Arb. at 752; 13 Lab. Arb. at 864.
153 59 Lab. Arb. 331 (1972) (Jaffe, Arb.).
154 *Id.* at 331-32.
155 *Id.* at 333.
156 16 Lab. Arb. 569 (1951) (Cheney, Arb.).
157 *Id.* at 573.
158 47 Lab. Arb. 62 (1966) (Duff, Arb.).

Klan. Attempting to increase Klan membership, this employee had held a public press conference and a public Klan rally. These activities ultimately aroused threats of a wildcat strike, a protest by the employer's black bus drivers, plans for a boycott of the busline by black riders, and possibilities of violence against the offending employee while he was on duty.[159] The arbitrator found these threats to the company's welfare to have justified the firing of the employee.[160]

As arbitrators have sought to develop standards for testing the validity of discharges that result from employees' political activities, sharp controversy has arisen over the question of whether arbitrators should interpret the just-cause clauses in collective-bargaining agreements according to the public policies evinced by the First Amendment and other bodies of law. A key issue in this controversy has been that of how far an arbitrator should go outside the express terms of an employment contract in construing the just-cause clause of the contract.

In 1960, in *United Steelworkers of America v. Warrior and Gulf Navigation Co.*,[161] the Supreme Court held that arbitrators could interpret collective-bargaining agreements according to considerations not expressly stated in the agreements themselves. In so doing, the Court recognized that collective-bargaining agreements created "a new common law — the common law of a particular industry or of a particular plant."[162] Within this system of "industrial self government,"[163] the arbitrator functions by interpreting the contract and helping to keep peace. But

> [t]he labor arbitrator's source of law is not confined to the express provisions of the contract, as the industrial common law — the practices of the industry and the shop — is equally a part of the collective bargaining agreement although not expressed in it.[164]

Professor Cox, whose views were relied upon by the Court in the *Gulf Navigation* case, has gone even farther in endorsing

---

159 *Id.* at 62-63.
160 *Id.* at 66.
161 363 U.S. 574 (1960).
162 *Id.* at 579.
163 *Id.* at 580.
164 *Id.* at 581-82.

the use of extra-contractual factors in construing labor contracts. He argues that the power sources of arbitral standards in the interpretation of contracts include "legal doctrines, a sense of fairness, the national labor policy, past practice at the plant, and perhaps good industrial practice generally."[165]

In another context, arbitrators have been likened to an industrial "Supreme Court," who in interpreting the collective-bargaining agreement must exhibit "judicial statesmanship" similar to that required of the Supreme Court in interpreting the Constitution.[166] Arbitrators have, without doubt, imported the values and policies underlying the due-process clause of the Constitution into industrial relations when they have interpreted contracts.[167] In particular, it would seem that the contractual just-cause clause, so broad in its own terms and so important in protecting employees from arbitrary or malicious actions, is the device most likely to be used for creating "industrial common law" that draws on sources outside the collective-bargaining agreement. Professor Getman, for example, has argued,

> The language of just cause is so general that it cannot be construed in terms of the precise intent of the parties. The most that can be said is that by the use of such language, the parties have manifested an intent to refer in discharge cases to the moral standards of the community modified for the industrial setting. The Constitution and the decisions that interpret it both reflect and shape contemporary standards of morality. As such, they are valuable sources of guidance to arbitrators in determining whether specific conduct is sufficiently reprehensible to justify discharge.[168]

But other arbitrators and commentators view the issue differently. Professor Dunsford, commenting on Professor Getman's argument, rejected the idea that judicial decisions could or should provide realistic guides for arbitral decisions under the just-cause clause.[169] Arbitrator Mittenthal has argued that arbitrators should not enforce statutory rules in interpreting

---

165 Cox, *Reflections Upon Labor Arbitration*, 72 HARV. L. REV. 1482, 1500 (1959).

166 Horlacher, *Employee Job Rights versus Employer Job Control: The Arbitrator's Choice*, in PROCEEDINGS OF THE 15TH ANNUAL MEETING OF THE NATIONAL ACADEMY OF ARBITRATORS, 165, 173-74 (BNA 1962).

167 ELKOURI, *supra* note 135, at 632.

168 Getman, *supra* note 30, at 61, 66. *See also* Summers, *Individual Protection Against Unjust Dismissal: Time for a Statute*, 62 VA. L. REV. 481, 500 (1976).

169 Getman, *supra* note 30, at 79.

collective-bargaining agreements.[170] And many arbitrators refuse to handle legal issues presented to them by the parties in a contract dispute.[171]

Whether or not arbitrators feel emboldened enough to consider extra-contractual legal issues and apply them to the contract,[172] the concept of public policy is a more familiar external source of standards for the interpretation of contracts, and it is used more often. In the 1965 case of *Uddo and Taormina,*[173] for example, the arbitrator explicitly invoked the public policy against coercing citizens to vote a particular way; relying on that policy, the arbitrator invalidated an employer's voting-time decision that effectively penalized those employees who did not vote.[174] In *Foote Bros. Gear and Machine Corp.,*[175] discussed briefly above, the arbitrator considered both society's judgment that the Communist Party should be allowed a legal existence and the public policy against penalizing "hundreds of thousands" of workers for past association with leftist and front groups; on the basis of those policies, the arbitrator determined that Communist Party membership could furnish just cause for discharge.[176]

It has been argued that the determination of whether just

---

170 Mittenthal, *The Role of Law in Arbitration,* in PROCEEDINGS OF THE 21ST ANNUAL MEETING OF THE NATIONAL ACADEMY OF ARBITRATORS, 42, 54 (BNA 1968). *See also* ELKOURI, *supra* note 135, at 325-28, summarizing the debate between Meltzer, who argues that legal rules are applicable to contract interpretation but that if there is conflict, the legal rule must be ignored; and Howlett, who believes that all collective-bargaining agreements are subject to the law's commands and that arbitrators must recognize this.

171 R. H. Buhrke Co., 68 Lab. Arb. 1170, 1171 (1977) (Nicholas, Arb.) (arbitrator refusing to decide if a violation of the collective-bargaining agreement was an unfair labor practice); Los Angeles Daily News, 19 Lab. Arb. 39, 39 (1952) (Dodd, Arb.) (arbitrator refusing to recognize as an issue the legality, under California law, of the discharge of editorial writers who have failed to deny association with the Communist Party).

172 And occasionally they do. *See* Theo-Kupfer Iron Works, 67 Lab. Arb. 1203, 1213 (1976) (Gundermann, Arb.) (employee discharged for making "false and defamatory statements" about employer; arbitrator looks to libel law to determine if the statements were in fact libelous before construing the just-cause clause); Michigan Consolidated Gas Co., 58 Lab. Arb. 1058, 1059 (1972) (Shister, Arb.) (parties requested arbitrator to rule on legality, under the Ninth and Fourteenth Amendments and the Taft-Hartley Act, of a private employer's policy requiring employees to be neatly groomed).

173 45 Lab. Arb. 72 (1965) (Howard, Arb.).

174 *Id.* at 74.

175 13 Lab. Arb. 848 (1949) (Larkin, Arb.).

176 *Id.* at 859-60.

cause existed for a discharge in many cases almost necessarily blends into public-policy determinations:

> [The just-cause clause] invites a consideration of policy factors . . . [T]he line between the construction of such language in a labor agreement and a decision based on public policy is not at all clear. An arbitrator might construe the contract in light of public policy considerations. His decision simultaneously involves public policy and a construction of a particular contract.[177]

One specific public policy that some arbitrators and commentators have sought to import into contract interpretation involving the just-cause clause has been the First Amendment and its underlying values. For example, Professor Getman has argued:

> Arbitrators should recognize that certain interests, such as freedom of speech and freedom of religion, are so fundamental to individual liberty that they can be limited and made the bases for disciplinary action only when management can demonstrate an overriding economic need.[178]

The values of the First Amendment are often raised before arbitrators dealing with private-employment discharges,[179] and occassionally arbitrators explicitly refer to these values in formulating their decisions. In *International Harvester Co.*,[180] a 1972 case that involved an employee's placing on toolboxes signs expressing support for another union, the arbitrator balanced the employee's interest in freedom of expression with the employer's need for efficient production. On the basis of this balancing, the arbitrator upheld as non-discriminatory the employer's order to remove the signs and his subsequent dis-

---

177 Blumrosen, *Public Policy Considerations in Labor Arbitration Cases*, 14 RUT. L. REV. 217, 221 (1960). As an example of this argument, Professor Blumrosen points to discharge cases involving pilots or drivers where incompetence or negligence is alleged. Is there not, he asks, in all such cases, a consideration of public safety to be weighed in deciding if an employee is to be reinstated? *Id.* at 222-24. Professor Blumrosen goes on to conclude in part that arbitrators generally consider it a duty to take public policy into account in interpreting collective-bargaining agreements, but they will not ignore the facts or the contract language. *Id.* at 235-36.

178 Getman, *supra* note 30, at 64. *See also* Palmer, *Free Speech and Arbitration: Implications for the Future*, 27 LAB. L.J. 287, 293-96 (May, 1976).

179 *See, e.g.*, Great Lakes Steel Corp., 60 Lab. Arb. 860, 862 (1973) (Mittenthal, Arb.); Parker-Hannifin Corp., 63 Lab. Arb. 386, 387 (1974) (Grant, Arb.).

180 59 Lab. Arb. 219 (1972) (Block, Arb.).

ciplinary action against the employee.[181] The arbitrator noted that if he had found the management's motive to have been "bargaining self-interest," then he would have "unhesitatingly rule[d] that the employees' constitutional right of freedom of expression outweighed the interest of the employer."[182] In *Spokane-Idaho Mining Co.*,[183] discussed briefly above, the arbitrator invoked the First-Amendment value of protecting minority views against majoritarian oppression as grounds for holding that a private employee could not be discharged for distributing Communist-Party literature and being a party member:

> Our government apparently still realizes that labor or the right to labor is valuable property right, WHICH SHOULD NOT BE denied or withdrawn from laboring people because their political of economic philosophies do no accord with those of a majority of our citizens. [Emphasis in original, footnote omitted.][184]

It seems that whether or not an arbitrator wants to consider First Amendment values explicitly, he institutionally ends up doing so. As part of the "industrial common law,"[185] arbitrators have held that just cause for discharge cannot include employee activities that do not harm the employer.[186] If the political speech of employees does not fall within this category of harm to the employer, it cannot justify a discharge; no further reference to public policy, the law, or the first amendment is required.[187] Arbitrators today need use only the rules and common-sense guidelines created in the industrial common law to protect employees' political speech from interference by private employers.

Political activities of workers become just cause for discharge when they adversely affect an employer. This rule has been stated by many different arbitrators. One arbitrator has stated that worker conduct amounting to demonstrations or protest, for example, whether on or off the plant premises,

> must undergo two major tests that come at once to mind before a valid claim can be made that employees are engaged

---

181 *Id.* at 222-23.
182 *Id.* at 223.
183 9 Lab. Arb. 749 (1947) (Cheney, Arb.).
184 *Id.* at 752.
185 *See* text accompanying notes 161 to 165 *supra.*
186 *See* text accompanying notes 133 to 160 *supra.*
187 This argument is similar to Professor Blumrosen's. *See* note 177 *supra.*

> in protected activity: (a) the conduct should not be potentially disruptive of employee relations or production, and (b) it should not have an adverse impact on the Company's customers.[188]

As another arbitrator said in relation to an employee's office poster depicting Emilio Zapata and the slogan, "Viva la Revolución,"

> the ultimate test for a proper poster, in the absence of mutually agreed upon reasonable limits and rules, is whether a poster adversely affects production or the productive capacities of plant employees.[189]

And Professor Getman has argued that employer interference with such rights should be closely scrutinized by the arbitrator:

> [W]here a company seeks to limit the use of political buttons or interferes with the free exercise of religious beliefs or limits the ability of employees to make statements critical of the company, the arbitrator should require that the company demonstrate the need for such limitation.[190]

Of course, this standard of actual or potential harm to the employer's economic interests is, by itself, an important check on an employer's ability to retaliate against employees' political speech. It certainly offers the employee more protection than the common-law rule of at-will employment.[191] Moreover, for two substantial reasons, it seems that this standard is reasonable to apply to employee political speech, notwithstanding the general interest society has in fostering political speech virtually free of any limitations. First, limiting employee speech to that which is not disruptive or harmful to the employer's operations

---

188 International Harvester, 59 Lab. Arb. 219, 223 (1972) (Block, Arb.).

189 California Processors, Inc., 56 Lab. Arb. 1275, 1282 (1971) (Koven, Arb.). Arbitrator Koven thus makes the point that the union could agree, in the context of the collective-bargaining agreement, to reasonable restrictions on employee speech. Most of the political-expression cases arbitrators have dealt with, however, concern employers acting under general management-rights clauses to make either rules or ad hoc decisions that result in limitations on employee speech. When violations of the rules or ad hoc decisions result in employee discharge or discipline, the arbitrator ends up testing the reasonableness of either determination under the just-cause clause.

190 Getman, *supra* note 30, at 68. *But cf.* Safeway Stores, 57 Lab. Arb. 185, 188 (1971) (Childs, Arb.), where the company's justification for a rule against the wearing of political buttons by employees was neutrality in facing the public. This justification was accepted by the arbitrator as "reasonable."

191 *See* text accompanying notes 31 to 47 *supra*.

makes sense in the context of private employment. The purpose of a factory or office is production of goods or services through economically efficient methods,[192] rather than the promotion of free speech. A standard that promotes fulfilment of this purpose with the least possible infringement of free speech is thus necessary to the needs of the economy. Indeed, to understand the necessity of such a standard, one need only consider the converse of such a standard: that an employee may speak on political issues regardless of any disruptive effects. Plainly, such a standard is unacceptable in a system requiring industrial efficiency and discipline.

A second reason for accepting the reasonableness of a standard designating disruptiveness of an employee's political speech as just cause for discharge is that such a standard forbids employers from interfering with political speech that is orderly and non-disruptive. Just as a private individual uninjured by a speaker's message has no right to interfere with the speaker, so too should an employer refrain from suppressing employees' political speech that does not adversely affect him.

The standard, however, suffers from a failure to define its main conceptual element. What constitutes "harm" to an employer's legitimate economic interests? How much harm must there be for an employer to have just cause to fire an employee? Is a few minutes' loss of production time following an employee's provocative remark about bussing enough to justify his discharge? Or, if an employee angers a few co-workers with a pro-Communist speech during lunch hour, has the employer been harmed enough to order the employee's discharge? How much proof must the employer present as to actual or potential harm caused by the employee's political speech before the employer can satisfy the just-cause standard for discharge or discipline?

Until these questions are answered, or at least addressed somewhat within the rule, the possibility that the rule will protect employee political speech is greatly decreased. Employers with ulterior motives could make exaggerated assertions of economic injury to justify their retaliation against political views they dislike. Even an employer having no such bad-faith mo-

---

192 *See* International Harvester, 59 Lab. Arb. 219, 222 (1972) (Block, Arb.).

tivation will probably see any inconvenience to his operations caused by employee political speech sufficient reason for discharge or discipline.

Because arbitrators have failed to define the just-cause standard beyond its general contours, they have been far from uniform in assessing employer justifications of economic injury in cases of discharges for political speech or activities. Some arbitrators almost unquestioningly accept an employer's assertion of potential harm to the employer's production or sales. In *Hearst Publishing Co.,*[193] a case decided in 1958, a reporter refused to testify before the House Un-American Activities Committee as to his past affiliation with the Communist Party. In upholding his discharge from the paper on the grounds that his conduct terminated his "usefulness" to the paper and was "detrimental" to the paper's interests, the arbitrator stated:

> [H]ere, where the Publisher did not and probably could not show any loss of circulation or subscribers to indicate financial harm to itself, the Arbitrator believes that a newspaper in particular can reasonably require certainty as to technique and approach to reporting from its employees.[194]

Because the reporter had been named as a former Communist and had refused to testify about it, the paper's ability to rely on the reporter's objectivity was destroyed.[195] In *Parker-Hannifin Corp.,*[196] a 1974 case, disciplinary action based on a company rule against political buttons and leaflets was upheld as reasonable against employees who wore buttons and distributed — on their own time — leaflets calling for the impeachment of President Nixon. The arbitrator noted that "the company is under no legal obligation, at the risk of an interference with its own business, to provide an arena for the carrying on of such activity."[197]

---

193 30 Lab. Arb. 642 (1958) (Schedler, Arb.).

194 *Id.* at 645.

195 Ironically, arbitrators have required the least proof of harm in cases involving reporters or editors whose past or present affiliation with the Communist Party was seen as detrimental to the employer's interests. The theory is that newspapers, which have a great responsibility to the public to present objective news, untainted by Communist propaganda, are justified in requiring absolute certainty in their employees' willingness to present unslanted news. *See* New York Mirror, 27 Lab. Arb. 548, 551 (1956) (Turkus, Arb.); Los Angeles Daily News, 19 Lab. Arb. 39, 39 (1952) (Dodd, Arb.).

196 63 Lab. Arb. 386 (1974) (Grant, Arb.).

197 *Id.* at 388. *See also* Safeway Stores, 57 Lab. Arb. 185, 188 (1971) (Childs, Arb.),

Although many arbitrators require evidence of actual harm or a positive showing of potential harm before they uphold disciplinary actions triggered by employees' political speech, the amount of evidence some require is slight. In the 1972 case of *International Harvester,*[198] the arbitrator accepted as "not at all farfetched" the argument by the company that customers coming through its plant would think employee signs on toolboxes supporting strikes at a neighboring plant indicated labor troubles at the employer's plant. The company's judgment "of the expected reaction of its customers" had to be accepted.[199] As noted above, in *Firestone Tire and Rubber*[200] the arbitrator upheld a firing that the employer had alleged was justified by pro-Soviet statements he attributed to the discharged employee but that the employee and other witnesses insisted were never made. The arbitrator rejected the testimony of the employee and his witnesses and found that the employee had incited violence and caused unrest that undercut morale and efficiency.[201] The arbitrator made this conclusion despite evidence that the only disruption at the workplace was a disputed incident in which one angry employee threatened to strike — but did not actually strike — the employee who was discharged.[202]

Other arbitrators require a great deal of evidence — or at least more than was presented to them — before they sustain the discharge of an employee for allegedly disruptive activity. In *Republic Steel Corp.,*[203] decided in 1957, the employer's testimony showed that the grievant's refusal to testify before the House Un-American Activities Committee on his affiliation with the Communist Party caused other employees to refuse to work with him. Subsequently, the employer changed work schedules

---

holding a similar rule to be reasonable and "no more a restriction of the freedom of the individual than [the company's] requirement regarding the wearing of the apron as a uniform." No actual or potential harm to the employer's supermarket business was demonstrated. *Id.* at 188.

198 59 Lab. Arb. 219, 223 (1972) (Block, Arb.).

199 *Id.* at 223.

200 16 Lab. Arb. 569 (1951) (Cheney, Arb.); *see* text accompanying notes 156 to 157 *supra,* for additional discussion of the facts of the case.

201 16 Lab. Arb. 569, 573 (1951) (Cheney, Arb.).

202 *Id.* at 572. For a court case where the evidence of disruption required under the just-cause clause of a collective-bargaining agreement was slight, *see* United Electrical Radio & Mach. Workers v. General Electric Co., 127 F. Supp. 934 (D.D.C. 1954), *aff'd on other grounds,* 231 F.2d 259 (D.C. Cir.), *cert. denied,* 352 U.S. 872 (1956).

203 28 Lab. Arb. 810 (Platt, Arb.).

and ultimately created make-work assignments for the grievant so he could work alone. After holding that the grievant could not be discharged simply because he stood on his constitutional right not to testify, the arbitrator ruled that "the record does not establish that the aggrieved caused dissension among his fellow employees to an extent which justified his discharge."[204] According to the arbitrator, the other employees had no right to refuse to work with the aggrieved employee, and the accompanying disruption of work was not in any event the type that could justify a firing. In the 1977 case of *R. H. Buhrke Co.,*[205] the arbitrator rejected the employer's argument that the wearing of anti-right-to-work-amendment T-shirts would seriously disrupt the operations at the employer's plant; the arbitrator noted that dress was informal at the plant, that political slogans had been worn on apparel there before, and that no evidence of disruption had been presented.[206] In *Baltimore Transit Co.,*[207] the arbitrator found that there had been just cause for discharge of the employee; but he reached this conclusion only after being presented voluminous evidence that failure to discharge the employee would have triggered boycotts, protests, and public indignation that in turn would have led to serious economic harm to the bus company's business and possibly to physical harm to the employee himself.[208]

A deceptively appealing explanation of the variation in the amount of proof arbitrators require of employers seeking to justify a discharge is the distinction between political activity conducted at the workplace and that conducted away from the workplace. Since in-plant speech is more likely than outside-of-plant speech to reach other employees, it can be presumed to be more disruptive. Under this presumption, little evidence of the actual effect of the employee's speech is necessary. On the other hand, outside-of-plant speech, being less likely to reach co-workers, is also less likely to harm the employer's economic interests, so that more evidence of actual or potential

---

204 *Id.* at 814.
205 68 Lab. Arb. 1170 (1977) (Nicholas, Arb.).
206 *Id.* at 1173.
207 47 Lab. Arb. 62 (1966) (Duff, Arb.). *See* text accompanying notes 158 to 160 *supra.*
208 47 Lab. Arb. at 66.

disruption resulting from such speech is required. Compare, for example, the in-plant speech in *Parker-Hannifan*,[209] where no evidence of disruption was required, with the outside-of-plant speech in *Baltimore Transit Co.*,[210] where the employee's off-duty activities were shown to have created an unmistakable threat to the employer, justifying discharge.

Under further analysis, however, the in-plant/outside-of-plant distinction fails as a basis for arbitrators to apply varying evidentiary standards. In many cases it fails as a useful explanation because the employee's outside-of-plant speech follows him into the workplace.[211] Furthermore, even if this distinction did explain cases where in-plant speech is presumed harmful for a firing to be justified, it would still be unacceptable. No "presumption," express or implied, should be used when actual proof is readily at hand, at least in the context of arbitrators under collective-bargaining agreements. The proper focus of arbitration is on proof of facts that create the right for employers and employees to act a certain way under their contract. In most cases where an employer justifiably claims damage or potential damage to his interests, he should have little trouble gathering or presenting proof of that damage.

In sum, arbitrators frequently recognize important First-Amendment values in cases involving employee political speech, and they have developed a rule that attempts to accommodate both the employer's interest in efficient production and the employee's interest in free speech. But this rule is not universally accepted and has not defined either the amount, type, or potential of economic injury the employer must show before he may discharge an employee for political speech. Consequently, the rule is inadequate for making the necessary accommodation between employer and employee interests. In effect, it tells neither the employer nor the employee how much speech is allowed or who must bear the burden for it beyond a certain point.

---

209 63 Lab. Arb. 386 (1974) (Grant, Arb.).
210 47 Lab. Arb. 62 (1966) (Duff, Arb.).
211 *See* the newspaper cases cited at note 195 *supra*.

III. Proposals for Protecting Private Employees' Freedom of Political Speech

As was explained in the preceding section, employees' freedom of political speech under the traditional common-law rule of at-will employment received no legal protection from interference by private employers: the rule allowed an employer to fire or otherwise penalize an at-will employee for disagreeing with the employee's political views or for any other reason. Although most courts still adhere to the at-will rule in its traditional form, private employees' freedom of political speech has gained at least marginal legal protection from three sources: (1) common-law departures by a few courts from the at-will rule; (2) federal and state statutes carving exceptions into the rule; and (3) inclusion of a just-cause clause in employment contracts, particularly in collective-bargaining agreements.

Conceptually, at least, the just-cause clause as applied in arbitration proceedings deserves recognition as the most promising source of protection for employees' freedom of political speech. The just-cause clause is flexible and allows for consideration of the pro-free-speech policy underlying the First Amendment. It also lends itself to a balancing of the competing interests of employees in free political speech and of employers in smooth, efficient operation of their businesses. Such a balancing has been conspicuously absent from most of the statutes and common-law decisions that have sought to protect employees from the harshness of the at-will rule.

In terms of their potential for achieving comprehensiveness in protecting employees' freedom of political speech, however, just-cause clauses of employment contracts are inferior to common-law decisions and to statutory reforms. Only a minority of private employees are protected by contracts containing a just-cause clause, and this severe limitation in coverage is not likely ever to be substantially reduced, let alone eliminated. Unlike employment contracts, common-law decisions and statutory enactments are potential means of protecting the political speech of broad categories of workers: depending on their specific terms, a court's common-law interpretation of employees' legal rights or a legislature's statutory enactments concerning

those rights may cover all employees in a particular industry, in a given type of job, in an entire state, or even in the entire nation.

Below are two proposals for reforming the legal treatment of private employees' freedom of political speech. One proposal calls for Congress or the various state legislatures to enact a statute embodying a balancing test that resembles, but is much more detailed and refined than, the balancing tests arbitrators have applied under just-cause clauses of employment contracts. The other proposal would have state courts promulgate new common-law rules preventing employers from discharging employees for non-disruptive political speech.

A. *A Proposed Statute*

The following statute is proposed as a way to accommodate both free-speech interests and economic-efficiency interests in the employee/employer relationship:

> *Section 1. Prohibition of Employers from Politically Coercing Employees.* No employer shall
>
> (a) make any express or implied threats to discharge, suspend, or otherwise discipline any employee, or
>
> (b) discharge, suspend, or otherwise discipline any employee,
>
> when such action is intended or tends to influence, control, direct, or retaliate against the political views, expressions, affiliations, or activities of any employee.
>
> *Section 2. Remedies Available to Injured Employees.* Any employer who violates the provisions of section 1 shall be liable to the injured employee in a civil suit for actual damages incurred, plus a sum of punitive damages not to exceed $10,000. If the employee has been discharged, he may seek reinstatement with back pay in lieu of punitive damages. Such reinstatement shall be rendered without any lapse in or loss of any rights accrued or acquired by the employee prior to his discharge.
>
> *Section 3. Defenses Available to Employers.* In any suit brought under section 2, the employer shall have the defense that the employee's political views, expressions, affiliations, or activities
>
> (a) substantially injured the employer's ability to produce his goods or services, or

(b) caused the loss of a substantial amount of business from customers or suppliers, or

(c) created a reasonable likelihood of the immediate occurrence of the events specified in sub-paragraphs (a) and (b) of this section.

The employer shall prove any or all of these defenses by a preponderance of the evidence. Proof of any one of these defenses by a preponderance of the evidence shall relieve the employer of all liability under the Act as to the particular suit in question.

Section 4. Limitations on Construction of this Act. Nothing in this Act shall prevent any employer from making any reasonable and neutral rule, enforced in a nondiscriminatory manner, regulating employee communications during the employee's working hours or in work areas; at the same time, nothing in this Act shall be construed to revoke, limit, or redefine any rights granted under the National Labor Relations Act as administrated by the National Labor Relations Board or interpreted by the courts.[212]

Section 5. Definitions.

(a) "Activities" includes, but is not limited to, voting in any public election.

(b) "employee communications" means any method of communication, including the distribution of printed matter.

(c) "employer" . . . . [This term can be defined by individual legislatures. Many legislatures may wish to define "employer" so as to exempt businesses they deem too small to be regulated by this statute.][213]

(d) "political" means of or having to do with issues, ideas, arguments, ideologies, or positions that deal with broad social policy choices, the organization, conduct, and powers of government, and similar matters of concern to the general public.

(e) "working hours" means those times when em-

---

212 This proviso will only be necessary if the statute is passed by Congress. It is assumed that the supremacy clause, U.S. CONST. art. VI, cl. 2, would prevent state passage of the proposed statute from restricting rights defined by the National Labor Relations Act.

213 *See, e.g.*, LA. REV. STAT. ANN. § 23:961 (West 1964), exempting employers with fewer than 20 employers from the prohibition against employer policies preventing employees from participating in politics or from becoming candidates for political office. (Section 23:962 covers all employees, providing a penalty against any employer who discharges an employee for his political opinions or attempts to control the vote of an employee through contractual agreement.)

ployees are scheduled to be at their jobs and engaged exclusively in job-related tasks.

(f) "work areas" means those areas of the employer's property, or those areas assigned to the employee, which are dedicated exclusively as far as the employee is concerned to the performance of the employee's job.

Section 6. Prohibition Against Waivers. No employee or labor organization may waive any of the rights granted under this Act.

Section 1 of the statute defines the employer's duty to the employee. The "express or implied threats" language, copied from statutes like New York's,[214] is intended to prevent subtle forms of coercion by the employer, such as suggesting that if a particular candidate is not elected, the shop might close.[215] The terms "discharge" and "suspend" should be clear in most cases. "Discipline" implies punishment of some sort, and should be fairly discernible in most cases (e.g., a "warning letter" in an employee's file). The employer's action must tend to or be intended to coerce the employee. This language, copied from statutes like California's,[216] is significant. It places on the employer two restrictions: he cannot consciously take any action that has a coercive effect. The seeming harshness of this rule is mitigated by the broad defenses allowed the employer in section 3 of the statute: any employer may take actions that have politically coercive consequences if he has a justifiable reason.

The rest of section 1 is worded in a general manner so as to preclude employers from interfering with any type of political speech or activity. "Political" is defined further in section 5. The usage of the term "any employee" is similar to that in the National Labor Relations Act.[217] It is meant to cover employees who have not been hired as agents of an employer, but who are nonetheless subject to an employer's control.

Section 2 attempts to remedy what is described above in this Note[218] as a major defect of most of the statutes protecting

---

214 N.Y. ELEC. LAW § 17-150(3) (McKinney 1978).

215 A practice more specifically aimed at in statutes like New York's, *id. Compare* MD. ANN. CODE, art. 33, § 26-16(a)(6) (1957), containing almost exactly the same language. *See also* R. I. GEN. LAWS § 17-23-6 (Supp. 1979).

216 CAL. LABOR CODE § 1101(b) (West 1971).

217 *See* National Labor Relations Act, 29 U.S.C. § 152(3) (1976).

218 *See* text accompanying notes 120 to 123 *supra*.

political speech: the lack of a civil remedy for the employee. Here, the employee can recover actual damages and, if he has been dismissed, can choose in addition to receive either punitive damages or reinstatement. The imposition of both punitive damages and a reinstatement order on an employer for an unjustified dismissal might constitute an excessively harsh sanction. While punitive damages may prevent or deter similar conduct by the employer in the future, the discharged employee may find that he would be better served by reinstatement. For many employees, especially those who are near retirement or lack job mobility for other reasons, reinstatement with no loss in accrued rights would be more valuable than the amount recoverable as punitive damages.

Section 3 sets out the defenses to an employee's action under the statute. It attempts to define more precisely the standards arbitrators have used under the just-cause clause for political-speech cases. Subsection (c), by designating speculative injury (as opposed to actual, provable injury) as a valid reason for discharge, might give the employer more latitude than is necessary to protect his economic interest. On the other hand, it would be unjust to require the employer to wait until his injuries had been made real, and possibly devastating, before allowing him to dismiss an employee who was the cause of those damages.[219] Furthermore, the employer is required to prove his defenses by a preponderance of the evidence, a heavier burden of proof than arbitrators have required in the past.[220] The employer must show "substantial" injury to his "ability to produce goods and services" under subsection (a). This language is meant to apply generally to the employer's work plant and to disruptions that an employee's speech or other political activity may cause in the production of goods or services. Injury is a relative term, and it is expected that in individual cases the amount of disruption that will be considered "substantial" will vary.

Section 4 is included in recognition of the fact that employer-promulgated rules on employee speech during working hours or in work areas, as defined in section 5, are often necessary

---

219 *See* Baltimore Transit Co., 47 Lab. Arb. 62 (1966) (Duff, Arb.)
220 *See* text accompanying notes 198 to 208 *supra*.

for the effective functioning of the enterprise. Such rules are allowed if they apply on their face to all types of communication and are not discriminatorily enforced against certain employees or certain types of political speech. A rule limiting political speech in non-work areas or during non-work time would be impermissible under the statute unless the employer could justify the rule's application under section 3 to some specific instance of speech. In order to prevent any implication that organizing rights are threatened by the allowance of work-time/work-area rules regulating employee communications, an explicit reference is made to rights granted under the National Labor Relations Act, explicitly and through its interpretation in such cases as *Republic Aviation v. NLRB*.[221]

Finally, section 6 prevents employers from forcing employees to waive the rights granted under the statute.[222]

Congress could, with the appropriate jurisdictional additions, pass the proposed statute.[223] The interstate commerce clause[224] of the Constitution appears to vest Congress with the authority to pass a statute protecting the political speech of employees against interference by private employers. Section 6 of the Fourteenth Amendment[225] may also provide Congress with the constitutional power to enact such a statute.

In the 1937 case of *NLRB v. Jones and Laughlin Steel Corp.*,[226] the Supreme Court determined that Congress possessed broad powers under the interstate commerce clause and upheld Congress' passage of the National Labor Relations Act. In particular, the Court stated that the clause was not limited specifically to "commerce," but that the regulatory power granted by the clause allowed Congress to reach "burdens" on commerce as well:

> The congressional authority to protect interstate commerce from burdens and obstructions is not limited to trans-

221 324 U.S. 793 (1945).

222 *See, e.g.*, 15 U.S.C. § 77n (1976).

223 The cause of action given to employees would be similar to that given to unions in the Labor Management Relations Act, as codified at 29 U.S.C. § 185 (1976).

224 U.S. CONST. art. I, § 8, cl. 3, provides that the Congress shall have power "[t]o regulate Commerce with foreign Nations, and among the several States. . . ."

225 Section 5 of the Fourteenth Amendment provides that "Congress shall have power to enforce, by appropriate legislation, the provisions of this article." U.S. CONST. amend. XIV, § 5; *see* note 104 *supra*.

226 301 U.S. 1 (1937).

> actions which can be deemed to be an essential part of a "flow" of interstate or foreign commerce. Burdens and obstructions may be due to an injurious action springing from other sources. The fundamental principle is that the power to regulate commerce is the power to enact "all appropriate legislation" for "its protection and advancement. . . ."[227]

As it was believed by Congress that regulation of the organizing activities of private employees would promote industrial peace[228] and hence promote interstate commerce, so too Congress could find that a statute protecting all employees against dismissal on grounds that have nothing to do with their abilities as workers is necessary for the maintenance of industrial peace and the furtherance of interstate commerce.

However, the *Jones and Laughlin* decision indicated that the interstate commerce clause gives Congress authority to regulate employers' discharges of employees only as necessary to alleviate harms that the objectionable employment practices have caused to interstate commerce.[229] Thus, congressional passage of the statute outlined above would withstand constitutional scrutiny by the courts under the interstate commerce clause purely to the extent that politically coercive or politically motivated actions by employers against employees could be shown harmful to interstate commerce.

Section 5 of the Fourteenth Amendment is much more problematic as a source of authority for Congress to pass the proposed statute. By its terms, Section 1 of the Fourteenth Amendment applies only to state action depriving persons of equal protection or due process. It is unclear whether section 5 of the Fourteenth Amendment can be read to grant Congress the power to reach wholly *private* deprivations of equal protection or due process.[230] If it can, perhaps freedom of speech could

---

227 *Id.* at 36-37 (footnotes and citations omitted).

228 *Id.* at 41-43.

229 *Id.* at 37.

230 This issue is different from the one the Court decided in Griffin v. Breckenridge, 403 U.S. 88 (1971); *see* text accompanying notes 90 to 95 *supra*. In *Griffin*, the plaintiffs were blacks claiming private interference with their right of interstate travel and other injuries based on racial animus. As such, they in effect brought their § 1985(c) suit under two constitutional rights explicitly protected against private interference: the Thirteenth Amendment right to be free from slavery, and the right of interstate travel. The Fourteenth Amendment's possible grant of power to Congress to reach private deprivations of Fourteenth Amendment rights did not have to be reached by the Court. Indeed, the Court explicitly reserved that issue. *Id.* at 107.

be included as one of the rights protected against private interference. In *U.S. v. Guest,*[231] private defendants were charged under 18 U.S.C. § 241[232] with a conspiracy to deprive blacks of rights of access to public accommodations and the right of interstate travel; the defendants were accused of seeking to carry out their conspiracy by assault, by murder, and by causing the false arrest of blacks.[233] Speaking for the Court, Justice Stewart addressed the question of whether Congress had the constitutional authority to reach private actions depriving persons of rights protected by the Fourteenth Amendment from state interference. He found sufficient evidence of state involvement in the conspiracy to make it subject to Congress's power under the Fourteenth Amendment. But he also specifically rejected any suggestion that the Court in this case had reached or was required to reach the issue of whether section 5 of the Fourteenth Amendment enabled Congress to reach purely private conspiracies to deprive persons of Fourteenth-Amendment rights.[234] In a concurring opinion, however, Justices Clark, Black, and Fortas declared that "Section 5 empowers the Congress to enact laws punishing all conspiracies — with or without state action — that interfere with 14th amendment rights."[235] Justices Brennan, Douglas, and Chief Justice Warren agreed, arguing in another concurring opinion that section 5 empowers Congress to prohibit *all* conspiracies that interfere with Fourteenth-Amendment rights, such as the right of equal access to state facilities.[236] Thus, a "majority" of the Court held, in one opinion or another, that section 5 gave Congress the power to reach wholly private conspiracies to deprive persons of Fourteenth Amendment rights.

This "holding" of the six concurring justices in *Guest* is troublesome, mainly because it carried majority support but has not been adopted by the Court in a majority opinion. If *Guest* does stand for the proposition that section 5 enpowers Congress to reach private actions that deprive a person of his constitu-

231 383 U.S. 745 (1966).
232 Section 241 is the criminal analogue of 42 U.S.C. § 1985(c) (1976); *see* text accompanying notes 88 to 115 *supra.*
233 383 U.S. at 747-48.
234 *Id.* at 755.
235 *Id.* at 762.
236 *Id.* at 782.

tional rights under the Fourteenth Amendment, perhaps it also supports the proposition that section 5 empowers Congress to protect freedom of political speech from private interference just as the First Amendment guarantees it against governnental interference. In 1971, in *Action v. Gannon,* the Court of Appeals for the Eighth Circuit adopted such an interpretation.[237] The Eighth Circuit, however, is the only court of appeals to have done so. With such scant authority for construing section 5 to empower Congress to regulate purely private action, the interstate commerce clause seems to be a much more reliable constitutional basis for congressional passage of a statute protecting employees' freedom of political speech from interference by private employers.

B. *Proposed Common-Law Rules*

An alternative to a statute protecting the political speech of employees is for the state courts to formulate new common-law limits on the employer's discretion to discharge at-will employees for non-disruptive political speech. Acting consistently with the respect traditionally accorded to public-policy considerations in judicial interpretation of tort law and contract law, courts could adopt common-law rules that would serve much the same purpose as the proposed statute.

In recognition of both the general importance of free political speech and the desirability of protecting employee's political speech without seriously undercutting smooth operations of employers' businesses, state courts across the country should recognize a common-law cause of action for wrongful discharges that are shown to have been motivated primarily by an employer's disagreement with an employee's political views. Judgments for employees in such an action should be granted only if the following two conditions are met: (1) that the employer cannot show by a prepondrance of the evidence that the employee's speech substantially injured the employer in its business operations or in its relations with customers or suppliers;

---

237 450 F.2d 1227, 1233-37 (8th Cir. 1971). *But cf.* Bellamy v. Mason's Stores, 508 F.2d 504, 507 (4th Cir. 1974), rejecting this approach as too innovative for a panel below the Supreme Court.

and (2) that the employer cannot establish that the action was taken against the employee only under a neutral rule, applicable to all employees without discrimination, forbidding political speech or activity during working hours or in working areas.

This wrongful-discharge action is essentially an action in tort, but the employer's motivation need be directed only *primarily,* rather than entirely, toward punishing the employee. Because the employer is given defenses similar to those discussed in the preceding section on a statutory remedy, he need not worry about being prevented from firing politically active employees who seriously disrupt the employer's business operations. An employee cannot win a wrongful-discharge suit merely by citing the political nature of his speech or activities.

In awarding damages, courts allowing the proposed action for wrongful discharge should, at the least, let an unjustifiably released employee recover back pay, computed from the time of the discharge to the time the suit is brought, minus any wages received in the interim from other employment. Reinstatement should be available essentially as an equitable remedy, but of course only at the discretion of the judge. While reinstatement of an employee in the common-law context might be rare, certainly the availability of this remedy for more than forty years under the National Labor Relations Act[238] provides a useful analogue of support and experience for common-law courts. Finally, as a deterrent measure, punitive damages ought to be available. But the amount can be limited by judicial rule, perhaps to bear some relationship to actual damages if the court fears punitive liability will be too great a burden on employers.

In justifying adoption into the common law of this proposed cause of action for wrongful discharge, courts could stress that common law has always originated and evolved with strong influence from public-policy considerations. Courts could also stress that a few jurisdictions have already used public policy to justify major exceptions to the at-will rule. As discussed previously, the New Hampshire Supreme Court in *Monge*[239] held that, notwithstanding the explicit terms of the employment contract, a discharge motivated by bad faith, malice, or retal-

---

238 29 U.S.C. § 160(c) (1976).

239 114 N.H. 130, 316 A.2d 549 (1974); *see* notes 51 to 53, 67 and accompanying text *supra.*

iatory intent is a breach of the employment contract.[240] In the area of definite-term employment contracts (as distinguished from at-will contracts), courts have long consulted extra-contractual public policy in determining when an employer or an employee is justified in terminating the contract.[241] In so doing these courts have in effect applied a doctrine of "just cause" for terminations of employment — a doctrine that, even if not explicitly labeled "just cause" by courts, nonetheless closely resembles that established by arbitrators of labor disputes.[242] This doctrine forbids employers from discharging employees for reasons that are not relevant to the employment relationship; it prohibits employers from being arbitrary or capricious in firings.

In applying tort doctrine, some courts have held that an employer who intentionally inflicts injury on an employee for a reason not recognized by the law as a justification has committed a tort against the employee.[243] For example, employees discharged for serving jury duty have successfully brought actions in tort for damages or reinstatement.[244] Clearly, employees incur significant loss when discharged from their jobs; in this sense, one can assume generally that an employer knowingly and intentionally imposes an "injury" on any employee he fires. Equally clearly, the ability to criticize policies and officials, to suggest alternative solutions to social problems, and to advocate reform is as important to our democratic political institutions as the availability of citizens for jury duty is for our judicial system.[245] Indeed, when an employer discharges a politically active employee purely due to disagreement with the employee's political views, the effect is to penalize the employee for exercising one of the most important freedoms in our democratic society: freedom of political speech.

There are strong grounds for considering protection of free speech and encouragement of vigorous debate or political issues

---

240 114 N.H. at 133, 316 A.2d at 551.

241 *See., e.g.,* Seco Chemicals v. Stewart, 349 N.E.2d 733 (Ct. App. Ind. 1976).

242 *Id.* at 739.

243 Nees v. Hocks, 272 Or. 210, 536 P.2d 512 (1975). *See also* Sheets v. Teddy's Frosted Foods, Inc., CONN. L.J., Jan. 22, 1980, at 1, 3-4.

244 *Id.*

245 *Cf.* Nees v. Hocks, 272 Or. 210, 536 P.2d 512 (1975); Reuther v. Fowler Williams, 255 Pa. Super. 28, 386 A.2d 119 (Super. Ct. Pa. 1978).

to be major policies of the federal government and of the various state governments. Not only does the First Amendment protect freedom of speech and freedom "to petition the Government for a redress of grievances," but most if not all of the state constitutions have similar provisions protecting state citizens' rights of free speech.[246] There is no more fundamental way for the nation or the various states to establish official policies than to enshrine them in a constitution.

In drawing on public policy to create new common-law rules, a number of state courts have created civil remedies for actions that under existing statutes were subject only to criminal sanctions. Several state courts have done so as a means of protecting private employees' rights of free speech.

While some courts have heretofore required that any public-policy exception be based on some *statutorily* evinced societal interest,[247] such a basis for a free-speech exception to the at-will rule already exists in states with employee free-speech statutes. The fact that the statute sets out only criminal penalties should not deter the court from declaring additional civil remedies. Although the legislature has determined sanctions for the violations of the statute, where these sanctions are inadequate to protect the fundamental, legislatively-enacted public-policy choices from being frustrated, the courts should not be afraid to step in to provide more effective protection for the public policy. This is exactly the reasoning adopted by the Illinois Supreme Court in *Kelsay v. Motorola, Inc.*,[248] and other courts have not been afraid to create civil remedies for discharged employees when to do so would promote the public policies of statutes otherwise enforced by criminal sanctions.[249]

Apart from the question of whether existing legal doctrine is adequate to justify adoption of the proposed common-law rules on employees' freedom of political speech, some courts have expressed fears that creation of a common-law cause of action

---

246 *See* B. SACHS, FUNDAMENTAL LIBERTIES AND RIGHTS: A 50-STATE INDEX 39-41. This index was prepared as a companion to CONSTITUTIONS OF THE UNITED STATES, NATIONAL AND STATE (Columbia University Legislative Drafting Research Fund 1980).

247 *See, e.g.*, Becket v. Welton Becket and Associates, 39 Cal. App. 3d 875, 114 Cal. Rptr. 31 (Ct. App. 1974); Campbell v. Ford Industries, 274 Or. 243, 546 P.2d 141 (1976).

248 74 Ill. 2d 172, 184-85, 384 N.E.2d 353, 357-58 (1978).

249 Sheets v. Teddy's Frosted Foods, Inc., CONN. L.J., Jan. 22, 1980, at 1, 3-4; Harless v. First National Bank in Fairmont, 246 S.E.2d 270 (W.Va. 1978).

for wrongful discharge from employment would unleash a flood of litigation, swamping the courts and drastically reducing employers' abilities to manage their operations effectively.[250] These fears, however, have not been borne out in states whose supreme courts have adopted general public-policy exceptions to the at-will rule — exceptions not limited merely to political speech.[251] In these states there has been no flood of litigation, and there does not seem to have been any extraordinary burden placed on the employers' management capabilities. As was indicated above, New Hampshire, the state with the broadest exception to the at-will rule, has experienced little such litigation arising under the exception since the exception was established.[252] Likewise, California and Puerto Rico, both of which allow a broad cause of action to injured employees under political-speech statutes,[253] have both been far from overwhelmed by litigation on the issue.[254]

V. Conclusion

The time is long past when the law would allow an employer to treat his employees with as much arbitrariness or malice as he desired. Since the late nineteenth century, society has made a steady progression of decisions to give the worker greater rights and protection against his employer. When a right so fundamental as free speech remains effectively subordinated to an employer's power over a job, it is time for society to address that aspect of the employment relationship as well. The rights of workers to think and speak on political issues can be accommodated to the needs of employers for an efficient work force. The making of this accommodation would be in the best tradition of American economic and legal history.

---

250 Geary v. U.S. Steel, 456 Pa. 171, 180-83, 319 A.2d 174, 179-80 (1974).

251 For leading examples of such holdings by state supreme courts, *see* Monge v. Beebe Rubber Co., 114 N.H. 130, 316 A.2d 549 (1974); Fortune v. National Cash Register Co., 373 Mass. 96, 364 N.E.2d 1251 (1977); Tameny v. Atlantic Richfield Co., 27 C.3d 167, 610 P.2d 1330, which relied on the seminal case of Petermann v. International Brotherhood of Teamsters, 174 Cal. App. 2d 184, 344 P.2d 25 (Ct. App. 1959). *See also* the discussions pertaining to *Monge, Fortune,* and *Petermann* in text accompanying notes 47 to 53 *supra.*

252 *See* text accompanying notes 51 to 53 *supra.*

253 *See* text accompanying notes 124 to 132 *supra.*

254 The lack of decided cases, however, does not indicate the lack of need for protection. *See* Note, *California's Controls on Employer Abuse of Employee Political Rights,* 22 Stan. L. Rev. 1015, 1021 (1970).

NOTE

THE SOFT DRINK INTERBRAND COMPETITION ACT OF 1980: ANTITRUST LOSES ITS FIZZ

LEONARD R. STEIN*

The Soft Drink Interbrand Competition Act of 1980 was enacted in response to two 1978 Federal Trade Commission decisions invalidating certain vertical restraints common in the soft drink industry. The Act protects such restraints from attack under the antitrust laws when "substantial and effective" competition exists in the relevant market. In this Note, Mr. Stein argues that the practices covered by the Act reduce competition without effecting the benefits claimed by their defenders. He maintains further that the Act was not meant to confer antitrust immunity on these practices and that Congress intended the "substantial and effective" test to be vigorously applied by the courts. A truly vigorous application, he argues, will eventually combine with the economic self-interest of industry members to eliminate the very restraints the industry sought to protect in calling for congressional action.

Introduction

The Soft Drink Interbrand Competition Act of 1980[1] [hereinafter Soft Drink Act] insulates certain soft drink industry prac-

---

* B.A., M.A., Yale University, 1978; Member, Class of 1981, Harvard Law School.

1 The Soft Drink Interbrand Competition Act of 1980, Pub. L. No. 96-308, 94 Stat. 939 (1980) [hereinafter cited as the Soft Drink Act]. The law, as enacted, provides as follows:

> *Section 1.* This Act may be cited as the "Soft Drink Interbrand Competition Act."
>
> *Section 2.* Nothing contained in any antitrust law shall render unlawful the inclusion and enforcement in any trademark licensing contract or agreement, pursuant to which the licensee engages in the manufacture (including manufacture by a sublicensee, agent, or subcontractor), distribution, and sale of a trademarked soft drink product, of provisions granting the licensee the sole and exclusive right to manufacture, distribute, and sell such product in a defined geographic area or limiting the licensee, directly or indirectly, to the manufacture, distribution, and sale of such product only for ultimate resale to consumers within a defined geographic area: *Provided,* That such product is in substantial and effective competition with other products of the same general class in the relevant market or markets.
>
> *Section 3.* Nothing in this Act shall be construed to legalize the enforcement of provisions described in section 2 of this Act in trademark licensing contracts or agreements described in that section by means of price fixing agreements, horizontal restraints of trade, or group boycotts, if such agreements, restraints, or boycotts would otherwise be unlawful.

tices from attack under the United States antitrust laws. The new law is a direct response to the Federal Trade Commission's 1978 rulings in the *Coca-Cola*[2] and *Pepsico*[3] cases. In those proceedings, the FTC invalidated all territorial limitations and customer restrictions in connection with the sale or distribution of soft drinks as unfair methods of competition under section 5 of the Federal Trade Commission Act.[4] Such restrictions have been standard within the industry for over 75 years, and industry members argue that their abolition would prompt higher prices, the destruction of small businesses, and the disappearance of the returnable bottle.[5] On the other hand, industry critics and some economists contend that removal of these restrictions would augment competition, increase efficiency, and maximize consumer welfare.[6]

The new law attempts to resolve these conflicting assessments by instructing the courts to evaluate the competitive environment in individual markets. Only where competition is "substantial and effective" will territorial and customer restrictions be presumed lawful under the antitrust laws.[7] This approach rejects the FTC's wholesale invalidation of these restrictions

---

Section 4. In the case of any proceeding instituted by the United States described in subsection (i) of section 5 of the Clayton Act (relating to suspension of the statute of limitations on the institution of proceedings by the United States) (15 U.S.C. § 16(i)) which is pending on the date of the enactment of this Act, that subsection shall not apply with respect to any right of action referred to in that subsection based in whole or in part on any matter complained of in that proceeding consisting of the existence or enforcement of any provision described in section 2 of this Act in any trademark licensing contract or agreement described in that section.

Section 5. As used in this Act, the term "antitrust law" means the Sherman Act (15 U.S.C. §§ 1 et seq.), the Clayton Act (15 U.S.C. §§ 12 et seq.), and the Federal Trade Commission Act (15 U.S.C. §§ 41 et seq.).

2 In re Coca-Cola Co., 91 F.T.C. 517 (1978), No. 78-1364 (D.C. Cir., heard Oct. 28, 1978).

3 In re Pepsico, Inc., 91 F.T.C. 680 (1978), No. 78-1544 (D.C. Cir., heard Oct. 28, 1978).

4 Both Coca-Cola and Pepsico appealed the Commission's decision to the United States Court of Appeals for the District of Columbia Circuit. The court heard argument on the case in October, 1978, but has yet to render its decision.

5 *See* text accompanying notes 88 to 124 *infra*. More generally, some commentators have argued that the field of vertical restrictions should be "swept clear of legal inhibition" because such restrictions promote economic efficiency. Bork, *Vertical Restraints: Schwinn Overruled*, 1977 SUP. CT. REV. 171.

6 *See* Comanor, *Vertical Territorial and Customer Restrictions: White Motor and Its Aftermath*, 81 HARV. L. REV. 1419 (1968).

7 Soft Drink Act, *supra* note 1, at § 2.

without embracing the rule of per se legality that had been sought by the industry.[8]

The purpose of this Note is to demonstrate that despite the protection the new law purports to give vertical restraints in the soft drink industry, it will eventually lead to their disappearance. This conclusion is based on four premises. First, Congress intended to validate vertical restrictions only where vigorous and atomistic competition is present. Second, the standard contained in the Soft Drink Act is no more generous than the rule-of-reason test announced by the Supreme Court in *Continental T.V., Inc. v. GTE Sylvania Inc.*;[9] indeed, the substantial-and-effective-competition standard is arguably more restrictive than the *Sylvania* test, since Congress has indicated that the relevant focus should be on extant competitive conditions within a market or markets and not on the relative economic efficiency of particular modes of distribution.[10] Third, the courts will give the substantial-and-effective-competition test serious weight in their deliberations. Fourth, the absence of competitive conditions in some markets will lead to the invalidation of territorial restrictions in those markets. Once this occurs, producers' self-interest will lead them to seek elimination of restrictions in other markets.

The following discussion seeks to justify these four propositions. Part I examines the law of vertical non-price restraints as established by Supreme Court precedent. Part II provides a short history of vertical restraints in the soft drink industry from their inception around 1900 until the Federal Trade Commission's decisions in the *Coca-Cola* and *Pepsico* cases. Part III analyzes the asserted benefits and costs of immunizing ver-

---

8 The FTC's order applied only to restrictions involving nonreturnable-bottle sales. The rationale for this limitation is examined at text accompanying note 121 *infra*.

9 433 U.S. 36 (1977).

10 It is possible to argue that Congress intended the courts to apply a two-step test. First, the court should examine whether substantial and effective competition exists in the relevant geographic and product markets. If competitive conditions prevail, the inquiry stops and the restrictions are validated. If not, the court should proceed to a *Sylvania* balancing test and weigh all the relevant factors. However, the latter inquiry cannot ignore the court's initial finding of imperfect competition. The very existence of a non-competitive/oligopolistic market should tilt the court's rule-of-reason determination toward a holding that the restraints are unreasonably anti-competitive. Thus, the court's rule-of-reason inquiry in a soft drink vertical restraint case must necessarily differ from the standard balancing test.

tical restraints in the soft drink industry from antitrust-law scrutiny. Finally, Part IV explicitly considers the reasonableness of the four premises proposed above.

I. COMPETITION AND VERTICAL NON-PRICE RESTRAINTS

Antitrust law distinguishes two types of market restraint. Vertical restraints are agreements between manufacturers and distributors concerning the manner in which the manufacturers' products will be distributed or sold. Horizontal restraints involve arrangements among competitors to fix prices, allocate markets, or otherwise restrain trade.[11] A further distinction is usually drawn between restraints which involve prices and those which do not. The courts have generally held horizontal restraints [12] as well as vertical price restraints[13] to be per se violations of the antitrust laws. In *Northern Pacific Railway v. United States,* Justice Black explained the rationale behind the per se concept:

> [T]here are certain agreements or practices which because of their pernicious effect on competition and lack of any redeeming virtue are conclusively presumed to be unrea-

---

11 A third type of agreement is known as "dual distribution." The court in Krehl v. Baskin Robbins Ice Cream Co., 1979-2 Trade Cas. ¶ 62,806, at 78,703-04 (C.D. Cal. 1979), defined this phenomenon as follows:

> In a dual distribution system, a single entity operates at two distinct levels of the market structure: the entity both provides a product to the market level one step nearer to the ultimate consumer and to the market level two steps nearer to the ultimate consumer or the consumer himself. For example, a manufacturer may both provide its product to a distributor for subsequent sale to consumers and provide the product directly to the consumer.

The current status of such restraints under the antitrust laws is unclear, although the presence of other restraints may cause them to be characterized as horizontal restraints. *See Krehl,* note 11 *supra,* and American Motor Inns, Inc. v. Holiday Inns, Inc., 521 F.2d 1230 (3d Cir. 1975).

12 United States v. Trenton Potteries Co., 273 U.S. 392 (1927); and United States v. Socony-Vacuum Oil Co., 310 U.S. 150 (1940) (price-fixing a per se violation of Sherman Act). Addyston Pipe and Steel Co. v. United States, 175 U.S. 211 (1899); Timken Roller Bearing Co. v. United States, 341 U.S. 593 (1951); United States v. Sealy, Inc., 388 U.S. 350 (1967); and United States v. Topco Associates, Inc., 405 U.S. 596 (1972) (horizontal market division, whether or not accompanied by price-fixing, is per se illegal). Klor's, Inc. v. Broadway-Hale Stores, Inc., 359 U.S. 207 (1959); and United States v. General Motors Corp., 384 U.S. 127 (1966) (group boycotts per se illegal).

13 Dr. Miles Medical Co. v. John D. Park and Sons, 220 U.S. 373 (1911) (resale price maintenance per se unlawful); Simpson v. Union Oil Co., 377 U.S. 13 (1964) (vertical price-fixing per se unlawful).

> sonable and therefore illegal without elaborate inquiry as to the precise harm they have caused or the business excuse for their use. This principle of per se unreasonableness not only makes the type of restraints which are proscribed by the Sherman Act more certain to the benefit of everyone concerned, but it also avoids the necessity for an incredibly complicated and prolonged economic investigation into the entire history of the industry involved, as well as related industries, in an effort to determine at large whether a particular restraint has been unreasonable — an inquiry so often wholly fruitless when undertaken.[14]

With one major exception,[15] the Supreme Court has evaluated vertical non-price restrictions under the more hospitable rule of reason. This standard, first announced in the *Standard Oil* case[16] and later elaborated upon in *Chicago Board of Trade v. United States,*[17] recognized that:

> Every agreement concerning trade, every regulation of trade, restrains. To bind, to restrain is of their very essence. The true test of legality is whether the restraint imposed is such as merely regulates and perhaps thereby promotes competition, or whether it is such as may suppress or even destroy competition.[18]

Prior to the passage of the Soft Drink Act, the-rule-of reason test governed the legality of the restraints present in the soft drink industry.[19]

The Soft Drink Act addresses agreements between soft drink bottlers, who distribute the finished product, and syrup manufacturers, granting an individual bottler the exclusive right to distribute and sell the manufacturer's product in a defined geo-

---

14 356 U.S. 1, 5 (1958).

15 United States v. Arnold, Schwinn & Co., 388 U.S. 365 (1967). *See* text accompanying notes 40 to 51 *infra.*

16 Standard Oil Co. v. United States, 221 U.S. 1 (1911).

17 246 U.S. 231 (1918).

18 *Id.* at 238.

19 Although the *Schwinn* rule declared most territorial restraints per se illegal, *see* text accompanying notes 40 to 51 *infra,* the lower courts read it narrowly and almost uniformly applied the rule of reason. *See* American Bar Association, *Vertical Restrictions Limiting Intrabrand Competition,* 14-20 (Monograph No. 2, 1977) [hereinafter cited as ABA Monograph No. 2]. The Supreme Court's decision in *Sylvania,* 433 U.S. 36 (1977), overruled the decision in *Schwinn* and restored rule-of-reason treatment for vertical non-price restraints. *See* text accompanying notes 52 to 71 *infra.* The FTC's 1978 rulings in the *Coca-Cola* and *Pepsico* cases were based on the test set forth in *Sylvania.*

graphic area. The agreement also requires the bottler to limit his sales to customers within his designated territory.[20] The former arrangement is commonly referred to as a territorial limitation and the latter as a customer restriction. Territorial limitations and customer restrictions differ from other vertical non-price restraints such as location clauses,[21] areas of primary responsibility,[22] and profit pass-over agreements[23] in that they foreclose *intra*brand competition[24] entirely. As a result, producers can survive rule-of-reason scrutiny only by showing that such restrictions provide significant benefits to *inter*brand competition.[25] The tension between the total elimination of intra-

---

20 Soft Drink Act, *supra* note 1, at § 2. The FTC's complaint in the *Coca-Cola* case contains a typical license agreement between Coca-Cola, the syrup manufacturer, and its bottlers:

> . . . COMPANY agrees to furnish to BOTTLER, and only to furnish for the territory herein referred to, sufficient syrup for bottling purposes to meet the requirements of BOTTLER in the territory herein described.
>
> . . . .
>
> . . . COMPANY does hereby select BOTTLER as its sole and exclusive customer and licensee for the purpose of bottling the Bottlers' bottle syrup, COCA-COLA, in the territory described.
>
> [BOTTLER agrees] not to use trade-marks COCA-COLA or COKE, nor bottle nor vend said product except in the territory herein referred to. This limitation, however, is not to prevent BOTTLER from acquiring similar rights for other territory.
>
> . . . .
>
> [BOTTLER agrees] not to use said distinctive [COCA-COLA] bottle for any other purpose than the bottling of COCA-COLA, and not in any territory except as herein referred to.

91 F.T.C. at 521.

21 A location clause is an agreement between a manufacturer and a dealer which specifies the physical location from which the dealer may sell or distribute the manufacturer's goods. Such provisions do not eliminate intrabrand competition because they do not specify the persons to whom the dealer may sell. This is the type of restriction which was at issue in the *Sylvania* case.

22 Areas of primary responsibility usually involve an agreement by the distributor to exert his best efforts to meet certain sales goals within a defined area. Although the dealer is not prohibited from making sales outside his area, failure to satisfy the stated goals may result in termination of the franchise. As a result, dealers generally concentrate their efforts within their own territories, and diminished intrabrand competition results. (Presumably, manufacturers do not assign overlapping territories.) 16G VON KALINOWSKI, BUSINESS ORGANIZATIONS § 62.01[1] (1978).

23 Profit pass-over agreements require a dealer who makes sales outside his own territory to compensate the dealer in whose territory the customer was located. Such clauses are intended to prevent dealers from acting as "free riders," *see* text accompanying note 165 *infra,* and act as a powerful deterrent to extraterritorial dealer initiatives.

24 Intrabrand competition denotes competition between different sellers of the same product.

25 Interbrand competition denotes competition between different sellers of similar products produced by independent manufacturers.

brand competition on the one hand, and the asserted gains to interbrand competition on the other, is reflected in the contrasting conclusions reached by Congress and the FTC in considering the nature of competition in the soft drink industry.

A. White Motor's *Uncertain Teaching*

Prior to 1963, the Supreme Court confronted vertical non-price restrictions in only one case, *United States v. Bausch and Lomb Optical Co.*[26] In that case, the Supreme Court held that "a distributor of a trademarked article may not lawfully limit by agreement, express or implied, the price at which or the persons to whom its purchaser may resell."[27] *Bausch and Lomb* involved a resale price maintenance scheme as well as non-price restraints, and the Court did not indicate whether the latter standing alone would constitute a per se violation of section 1 of the Sherman Act. Nevertheless, the Department of Justice interpreted the *Bausch and Lomb* decision as holding that both territorial limitations and customer restrictions were per se unlawful, and subsequently obtained several consent judgments on this basis.[28]

United States v. White Motor Co.[29] provided the Supreme Court with its first opportunity to address the legality of vertical territorial limitations. White Motor manufactured trucks and truck parts. The company's franchise contracts with its dealers (1) restricted the geographic areas within which distributors and dealers were permitted to sell trucks and parts, (2) restricted the persons to whom distributors and dealers could sell trucks for resale, and (3) precluded distributors and dealers from selling to any federal, state, or municipal government, as well as to certain other large customers, without the manufacturer's express permission. Furthermore, the contracts fixed the resale

---

26 321 U.S. 707 (1944). Prior to World War II, vertical territorial limitations and customer restrictions were uniformly upheld by the lower courts. ABA Monograph No. 2, *supra* note 19, at 6, 7 & n.14.

27 321 U.S. at 721 (1944). Soft-lite was the exclusive dealer for Bausch and Lomb pink tinted eyeglass lenses. Soft-lite, in turn, sold lenses to wholesalers who were permitted to resell only to dealers expressly approved by Soft-lite and at prices which conformed with published price lists determined by Soft-lite.

28 ABA Monograph No. 2, *supra* note 19, at 7 n.17.

29 372 U.S. 253 (1963).

price for trucks, parts, and accessories sold by distributors and dealers to certain customers.[30] In the trial court, White Motor argued that such restrictions were necessary to enable it to compete effectively, that franchisees were entitled to be protected against invasions of their territory, and that such arrangements insured that dealers would exert maximum sales efforts within their own territories.[31] The district court held that the territorial limitations, customer restrictions, and price-fixing provisions constituted per se violations of sections 1 and 3 of the Sherman Act.[32] Accordingly, it granted summary judgment for the Government. White Motor chose not to appeal the adverse judgment on the price-fixing count, and the only question presented upon appeal was the legality of the territorial and customer restraints.

The Supreme Court reversed and remanded the case for trial without stating any opinion on the legality of the restraints. The Court justified its failure to apply either the rule-of-reason or a per se standard as follows:

> We do not know enough of the economic and business stuff out of which these arrangements emerge .... They may be too dangerous to sanction or they may be allowable protections against aggressive competitors or the only practicable means a small company has for breaking into or staying in business . . . and within the "rule of reason." We need to know more than we do about the actual impact of these arrangements on competition to decide whether they have such a "pernicious effect on competition and lack . . . any redeeming virtue" . . . and therefore should be classified as per se violations of the Sherman Act.[33]

Justice Clark, joined by the Chief Justice and Justice Black, filed a sharply worded dissent which chastised the majority for not declaring vertical territorial and customer restraints per se unlawful under the Sherman Act. Justice Clark argued that a customer seeking to purchase a White truck "must deal with only one seller who . . . has the sole power . . . to set prices, determine terms and even to refuse to sell to a particular cus-

30 *Id.* at 254-56.
31 *Id.* at 256-57.
32 *Id.* at 256.
33 *Id.* at 263.

tomer."[34] Although the consumer could buy another brand of truck, "the existence of interbrand competition has never been a justification for an explicit agreement to eliminate competition."[35] In the view of Justice Clark, such agreements should have been declared per se unlawful. On remand, White Motor entered into a consent decree.[36]

Although the majority expressly refused to adopt either a per se or rule-of-reason approach to territorial and customer restraints,[37] two courts subsequently interpreted the Court's action as a *sub rosa* endorsement of the rule of reason.[38] This set the stage for the Supreme Court's next encounter with vertical non-price restraints four years later in *United States v. Arnold, Schwinn and Co.*[39]

B. Schwinn *and the Per Se Rule.*

As in *White Motor,* the relevant agreements in *Schwinn* prohibited wholesalers from dealing with customers outside their territories and from selling to non-franchised dealers. Schwinn, previously the United States' largest bicycle manufacturer, had suffered a decline in its market share from 22.5 percent in 1951 to 12.8 percent in 1961. During the same period, the market share of its closest rival had grown from 11.6 percent to 22.8

34 *Id.* at 278.
35 *Id.*
36 United States v. White Motor Co., 1964 Trade Cas. ¶ 71,195 (N.D. Ohio 1964).
37 *But see* Mr. Justice Brennan's concurring opinion. *White Motor,* 372 U.S. at 264-66.
38 Snap-On Tools Corp. v. FTC, 321 F.2d 825, 828 (7th Cir. 1963) ("[*White Motor*] determined that vertical allocations of dealer territory are not *per se* violations of the Sherman Act. Such a holding necessitates an inquiry into the reasonableness of Snap-On's dealer arrangements. . . ."). Sandura v. FTC, 339 F.2d 847, 849-50 (6th Cir. 1964) "[We must] refuse to find Sandura's arrangements illegal without examining their particular effect on competition *and* the facts offered to justify the resulting restraint.") *See* 16G Von Kalinowski, Business Organizations § 62.02, at 62-32, 62-34 (1978) (emphasis in original). Preston, *Restrictive Distribution Arrangements: Economic Analysis and Public Policy Standards,* 30 L. & Contemp. Prob. 506 (1965), argues in favor of a balancing test which weighs the pro- and anti-competitive effects of vertical territorial and customer restrictions. He maintains that, in general, courts should permit vertical non-price restraints on distribution where they (1) substantially increase market coverage without direct reduction in the availability of products from other suppliers, and (2) do not result in supra-normal supplier or distributor profits. *Id.* at 519. By Preston's standard, the restraints present in *White Motor* were reasonable. *Id.* at 525.
39 388 U.S. 365 (1967).

percent.[40] As a result, Schwinn decided to revamp its distribution system. It cut the number of its retail dealers by 63 percent and set up a regional wholesale distribution system comprised of 22 separate territories. The wholesalers dealt with Schwinn in three ways. First, they resold bicycles purchased from Schwinn to authorized dealers in their own territories. Second, they acted as agents for Schwinn by arranging direct sales between Schwinn and dealers; under this arrangement, known as the "Schwinn Plan," they received a commission for their efforts. Third, they served as consignment agents for a small volume of bicycles which Schwinn insured and retained title to; these stocks permitted dealers to obtain immediate inventories of bicycles directly from Schwinn. (As with the Schwinn Plan, payment was made directly to Schwinn, with the wholesaler receiving a commission.[41]) The district court upheld the franchising system, the customer restraints, and the territorial limitations imposed on wholesalers when they were acting as Schwinn agents and consignees. The court struck down the territorial limitations relating to bicycles sold by Schwinn to its wholesalers as a per se violation of Sherman Act, section 1.[42]

The Supreme Court reversed and held that where a manufacturer has parted with title, dominion, or risk with respect to a product, both territorial limitations and customer restrictions on subsequent resales are per se violations of Sherman Act, section 1:

> Under the Sherman Act, it is unreasonable without more for a manufacturer to seek to restrict and confine areas or persons with whom an article may be traded after the manufacturer has parted with dominion over it. . . . Such restraints are so obviously destructive of competition that their mere existence is enough.[43]

Where the manufacturer retained ownership over the product, the Court held that a rule of per se illegality might unfairly

---

40 *Id.* at 368-70.
41 *Id.* at 370.
42 *Id.* at 367-68. The United States did not appeal the district court's rejection of the price-fixing charge, nor did Schwinn appeal the finding that restraints on resale by distributors who purchase products from Schwinn were per se illegal under the Sherman Act.
43 *Id.* at 378-79.

disadvantage small enterprises or accelerate a trend toward vertical integration. In such cases, vertical non-price restraints would be deemed unlawful only when they were "unreasonably" restrictive of competition.[44] In this case, the Schwinn agency and consignment agreements satisfied the rule of reason.[45]

The *Schwinn* decision generated significant criticism.[46] Some commentators complained that the Court had inadequately explained what it meant when it said that non-price restraints were unreasonable "without more."[47] Others echoed Justice Stewart's dissenting opinion, in which he criticized the majority's reliance upon "the ancient rule against restraints on alienation"[48] as well as its unexplainable distinction between agency-and-consignment transactions on one hand and sales on the other.[49] On a more practical level, *Schwinn* raised fears that the abolition of the practices which it condemned, which were endemic in a wide range of American industries, would prompt many businesses to abandon independent systems of distribution.[50] Widespread invalidation did not materialize, however, largely because subsequent reviewing courts chose to read the *Schwinn* decision narrowly, often limiting the case to its facts.[51] Furthermore, independent systems of distribution provide many efficiencies for the manufacturer (e.g., risk-spreading, increased sources of investment capital, and division of labor) which prob-

---

44 *Id.* at 380.

45 *Id.* at 381. Justice Stewart filed a separate opinion, concurring in part and dissenting in part. His opinion expressed the fear that abolition of vertical restraints would foster a trend toward vertical integration and the simultaneous disappearance of the small, independent merchant. This notion played a central role in the debate over the adoption of the Soft Drink Act.

46 *See, e.g.*, Handler, *Twenty-Five Years of Antitrust*, 73 COLUM. L. REV. 415 (1973); Note, *Territorial Restrictions and Per Se Rules — A Reevaluation of the Schwinn and Sealy Doctrines*, 70 MICH. L. REV. 616 (1972); Posner, *Antitrust Policy and the Supreme Court: An Analysis of the Restricted Distribution, Horizontal Merger and Potential Competition Decisions*, 75 COLUM. L. REV. 282 (1975); Pollock, *Alternative Distribution Methods After Schwinn*, 63 NW. L. REV. 595 (1968); and sources cited in ABA Monograph No. 2, *supra* note 19, at 9 n.24.

47 388 U.S. at 379. *See* sources cited in ABA Monograph No. 2, *supra* note 19, at 9 n.24.

48 388 U.S. at 380. *See* Handler, *supra* note 46, at 458-59; Posner, *supra* note 46, at 295-96.

49 388 U.S. at 379-81. *See* Pollock, *supra* note 46, at 606-08; Posner, *supra* note 46, at 296-97.

50 Pollock, *supra* note 46, at 609-10.

51 ABA Monograph No. 2, *supra* note 19, at 14-20; VON KALINOWSKI, *supra* note 38, at § 62.04.

ably outweigh the marginal loss in control from the elimination of vertical restraints. Finally, at least in the short run, few firms are likely to have the massive resources necessary to supplant an independent chain of distribution with a comparable internalized system.

C. Sylvania *and the Return to the Rule of Reason*

Schwinn remained good law until the Supreme Court decided *Continental T.V., Inc. v. GTE Sylvania Inc.*[52] in 1977. The *Sylvania* decision announced that vertical non-price restrictions would thereafter be judged under the rule of reason. The Court expressly overruled *Schwinn* and held that vertical non-price restraints would be deemed per se unlawful only if they had a "pernicious effect on competition" and lacked "any redeeming virtue."[53]

Sylvania, a manufacturer of television sets, sought to improve its market position by attracting more aggressive and competent retailers. As part of this scheme, Sylvania limited the number of retail franchises granted in a given area and required each retailer to sell its products only from the location or locations in which it was franchised. The plaintiff, one of Sylvania's franchised dealers, brought suit claiming that the franchise violated section 1 of the Sherman Act.[54] The jury returned a verdict for the plaintiff after the district judge, relying on the *Schwinn* case, instructed it that location restrictions constituted a per se violation of the Sherman Act.[55] The Ninth Circuit, sitting en banc, reversed and distinguished the case from *Schwinn*.[56] The court reasoned that Sylvania's location clauses had less potential for anti-competitive effects than the airtight territorial and customer restraints in *Schwinn* and should therefore be judged under the rule of reason.[57]

The Supreme Court affirmed the judgment of the Ninth Cir-

---

52 433 U.S. 36 (1977).

53 *Id.* at 50 (quoting Northern Pacific Railway v. United States, 356 U.S. 1, 5 (1958)). *See* text accompanying note 14 *supra*.

54 433 U.S. at 38–42.

55 As reported in 537 F.2d 980, 985–88 (1976).

56 *Id.*

57 433 U.S. at 41–42.

cuit, but differed in its comparison of the *Sylvania* and *Schwinn* restraints. The majority felt that Sylvania's location clauses fell directly within the ambit of the per se rule in *Schwinn*. Thus, the Court was limited to two courses of action: either it could grant *Schwinn* its full stare decisis effect[58] and strike down the *Sylvania* restraints as per se unlawful, or it could entirely reconsider the *Schwinn* rule. The Court, reasoning that "*Schwinn* itself was an abrupt and largely unexplained departure from *White Motor*,"[59] chose the latter course.

In evaluating the per se rule announced in *Schwinn*, the Court focused on the effects which vertical non-price restraints have on intrabrand and interbrand competition.[60] The distinction in *Schwinn* between sale and non-sale transactions failed to address adequately the potential for intrabrand competitive harm and interbrand competitive gain created by various vertical non-price restraints. While it was true that vertical restraints reduced intrabrand competition by limiting the number of dealers who could compete for the patronage a given group of customers, such restrictions also permitted manufacturers to achieve efficiencies in the distribution of their products. These benefits included (1) enhancing the ability of new entrants to attract investment in products unfamiliar to consumers, and (2) encouraging dealers of established manufacturers to provide service facilities and engage in promotional activities without the fear that competing dealers would benefit from these expenditures at no cost (the "free rider" effect). Although the Court left open the possibility that particular applications of vertical non-price restraints might justify per se prohibition,[61] it held that, in general, such restraints should be judged under the rule of reason.[62]

---

58 This is the position which Mr. Justice White advocated in his dissent. *Id.* at 59–71.

59 *Id.* at 47.

60 In a footnote, the Court observed that "interbrand competition . . . is the primary concern of antitrust law." *Id.* at 52 n.19.

61 *Id.* at 58.

62 *Id.* at 59. The Court also cited approvingly the rule-of-reason test laid down in the *Chicago Board of Trade* case:

> To determine [whether the rule of reason is satisfied] the court must ordinarily consider the facts peculiar to the business to which the restraint is applied; its condition before and after the restraint was imposed; the nature of the restraint and its effect, actual or probable. The history of the restraint, the

Although the *Sylvania* decision signifies a more flexible approach to vertical non-price restrictions, several problems remain.[63] The rule-of-reason approach undeniably has the virtue of allowing courts to consider a multiplicity of economic factors in reaching their decisions. The discretionary nature of this inquiry, however, undermines the notion that the trial court is applying established legal principles and not simply making ad hoc determinations of fact, degree and value. While the rule of reason may entail closer and more reasoned analysis of the facts before the court, it does not provide a clear guide to future conduct for other similarly situated individuals. For example, the majority correctly indicated that a rule-of-reason analysis of non-price restrictions must focus on the relative effects which these restraints have on intrabrand and interbrand competition. The Court further provided a partial catalogue of particular anti-competitive and pro-competitive effects which might accrue from such arrangements.[64] However, the Court never explicitly detailed which pro-competitive and anti-competitive factors were determinative in *Sylvania*. This omission gave the district courts little guidance as to which effects should be assigned considerable weight and which should be ignored. As a result, several commentators have suggested the need for a more clearly defined rule of reason.[65]

Another troubling aspect of *Sylvania* is that it couches its argument almost exclusively in terms of economic justifica-

---

> evil believed to exist, the reason for adopting the particular remedy, the purpose or end sought to be attained, are all relevant facts. This is not because a good intention will save an otherwise objectionable regulation or the reverse; but because knowledge of intent may help the court to interpret facts and to predict consequences."

Id. at 49 (quoting 246 U.S. 231, 238 (1918)).

63 Like *Schwinn*, *Sylvania* has prompted extensive commentary. *See, e.g.*, Posner, *The Rule of Reason and the Economic Approach: Reflections on the Sylvania Decision*, 45 U. CHI. L. REV. 1 (1977); Bork, *Vertical Restraints: Schwinn Overruled*, 1977 SUP. CT. REV. 171; Pitofsky, *The Sylvania Case: Antitrust Analysis of Non-Price Vertical Restrictions*, 78 COLUM. L. REV. 1 (1978); and Note, *Vertical Restraints and the Demise of Schwinn: Sylvania and the Rule of Reason*, 42 ALBANY L. REV. 137 (1977).

64 *See* text accompanying notes 57 to 59 *supra*.

65 Strassner, *Vertical Territorial Restraints After Sylvania: A Policy Analysis and Proposed New Rule*, 1977 DUKE L.J. 775; Zelek, Stern & Dunfee, *A Rule of Reason Decision Model After Sylvania*, 68 CALIF. L. REV. 13 (1980). Pitofsky, note 60 *supra*, advocates retention of a per se approach to certain vertical non-price restrictions which are particularly damaging to intrabrand competition.

tion.[66] While this approach has been applauded by commentators who generally favor the importation of microeconomic analysis into antitrust law,[67] a welfare-maximization analysis often ignores the social concerns underlying the antitrust laws. As the discussion below will demonstrate, one of the most compelling arguments in favor of vertical non-price restraints is that they permit the preservation of small independent businesses.[68] This purpose would not be well-served by rules based solely on economic considerations.[69]

Finally, *Sylvania* does not attach sufficient importance to the efficiencies resulting from intrabrand competition. Interbrand competition may be the critical concern where vertical integration is an imminent possibility.[70] Fully integrated systems of distribution, however, are neither achievable nor desirable for many businesses. In some industries, an independent chain of distribution can promote efficiency by allowing risk-spreading and facilitating capital formation. Furthermore, most manufacturers lack the capital resources necessary to create an effective, internalized regional or national distribution network. Where manufacturers will distribute their products through independent entities whether or not they can exercise control through vertical restrictions, intrabrand competition is as important a concern as interbrand competition, since an increase or decrease in the number of intrabrand competitors will have the same effect on competition as an increase or decrease in the number of competing interbrand entities.[71]

Nevertheless, the *Sylvania* decision remains an admirable

---

66 The Court does, however, recognize that safety considerations might also play a role. 433 U.S. at 55 n.23.

67 *Cf.* Bork, note 61 *supra*.

68 Judge Learned Hand stated this view in United States v. Aluminum Co. of America, 148 F.2d 416, 429 (2d Cir. 1945):

> [I]t has been constantly assumed that one of [the antitrust laws'] purposes was to perpetuate and preserve for its own sake and in spite of possible cost, an organization of industry in small units which can effectively compete with each other.

69 Exclusive reliance on efficiency criteria could also produce undesirable results in cases involving vertical integration and vertical mergers, where the generation of scale economies would run directly counter to the preservation of small business.

70 The presence of intrabrand competition is illusory where the firm can fully internalize its system of distribution at any time.

71 This is particularly true where only minor product differentiation exists between brands.

attempt to reconcile the competing legal and policy questions raised by vertical non-price restraints. The complexity of the problem is clearly demonstrated by the Federal Trade Commission's application of the *Sylvania* standard to the soft drink industry.

II. VERTICAL RESTRAINTS IN THE SOFT DRINK INDUSTRY

This section examines the nature and history of competition and vertical restraints in the soft drink[72] industry from its inception to the Federal Trade Commission's 1978 decisions in the *Coca-Cola*[73] and *Pepsico*[74] cases. The focus will be upon the development of the Coca-Cola company. This approach is appropriate for several reasons. First, Coca-Cola is the largest producer of soft drinks in the United States.[75] Second, Coca-Cola's experience is typical of other soft drink manufacturers' and fully illustrates the methods of distribution which are standard in the industry. Third, the Federal Trade Commission's decision in the *Coca-Cola* case was the primary stimulus to the enactment of the Soft Drink Interbrand Competition Act of 1980.

A. *The Development of Territorial Restrictions in the Soft Drink Industry*

Coca-Cola was invented and first marketed in 1886.[76] Other early entrants into the industry included Dr. Pepper, Pepsi, Hires, and Clicquot Club.[77] Most of these manufacturers originally sold fountain or "post-mix" syrup directly to soda fountains which mixed the syrup with carbonated water to form the

---

72 The term "soft drink" as used in this Note refers solely to carbonated, flavored soft drinks and does not include either unflavored and carbonated beverages (such as mineral water) or flavored and uncarbonated beverages (such as grape juice).

73 91 F.T.C. 517 (1978).

74 91 F.T.C. 680 (1978).

75 *Soft Drink Interbrand Competition Act of 1980: Hearings on S. 598 Before the Subcomm. on Antitrust, Monopoly and Business of the Senate Comm. on the Judiciary*, 96th Cong., 1st Sess. 111 (1979) (statement of William S. Comanor) [hereinafter cited as *Hearings on S. 598*].

76 91 F.T.C. at 527.

77 *Id.* at 531–32.

final product. Around the turn of the century, industry members sought to gain broader distribution of their products by licensing local businessmen to bottle and distribute the product at the wholesale level. Exclusive territorial provisions were thought necessary to encourage businessmen to invest the necessary capital and to exert sufficient effort to gain widespread consumer acceptance of soft drinks.[78] As a result, territorial restrictions became a standard element of bottlers' franchise agreements. By 1920, over 880 bottling companies were licensed to sell Coca-Cola nationwide.[79] That same year a federal district court upheld Coca-Cola's territorial licenses as reasonable ancillary restraints.[80]

The standard bottler's trademark agreement has changed little in almost 80 years. The syrup company manufactures the syrup and sells it to the bottlers. The bottlers produce and bottle the final product according to the manufacturer's specifications. The bottler chooses his plant and equipment, determines his own marketing strategy (including the choice of container sizes) and sets his own wholesale prices.[81] In addition, a bottler may bottle and distribute the products of competing manufacturers. This practice, known as "piggybacking," is discussed further below.[82]

Typically, the product is delivered to retailers through what is known as the "store-door" system. Route trucks will service accounts on a regular basis, making deliveries, retrieving empty bottles for re-use, rotating stock, and maintaining product displays.[83]

Over the years, the nature of soft drink consumption has changed significantly. In 1900, 70 percent of soft drinks were consumed on the vendor's premises and 30 percent in homes. The emergence of supermarkets, "six-packs," and coin-operated vending machines sharply altered consumption patterns. By 1975, food stores accounted for between 85 and 90 percent of sales for home consumption, or 50 percent of total soft drink sales. Vending machines alone accounted for 18 percent of total

---

78 *Id.* at 532–33.

79 Brief for Respondent at 11, Coca-Cola Co. v. FTC, No. 78–1364 (D.C. Cir., heard Oct. 28, 1978).

80 Coca-Cola Bottling Co. v. Coca-Cola Co., 269 F. Supp. 796 (D. Del. 1920).

81 91 F.T.C. at 533.

82 *See* text accompanying notes 107 to 110 *supra*.

83 Brief for Petitioner at 8–10, Coca-Cola Co. v. FTC, No. 78–1364 (D.C. Cir., heard Oct. 28, 1978).

industry sales. Per capita consumption of soft drinks increased from 1.6 eight ounce containers in 1849 to 100.1 containers in 1940 and 429.6 containers in 1973.[84]

Despite the fact that hundreds of private label soft drink brands have appeared since the early 1960's, total concentration in the industry continues to increase.[85] The percentage of sales made by the four largest firms in the industry rose from 47 percent in 1935 to 62 percent in 1963 and 72.4 percent in 1977.[86] Furthermore, between 1963 and 1978, the number of domestic bottlers shrank from 3,569 to 1,724.[87] Given the tremendous changes in consumption patterns, productive technology, and industry structure, it is curious that the nature of the manufacturer-bottler relationship has survived virtually unaltered for almost 80 years.

B. *The Federal Trade Commission Proceeding*

On July 15, 1971, the Federal Trade Commission filed complaints against seven soft drink syrup companies[88] alleging, *inter alia,* that territorial provisions which restricted soft drink sales outside a designated geographic area were unlawful because they eliminated intrabrand competition among bottlers in the distribution and sale of soft drinks. In 1975, the administrative law judge reached an initial decision in the *Coca-Cola* and *Pepsico* cases.[89] He concluded that territorial restrictions promoted interbrand competition and were therefore not unreasonable restraints of trade. Accordingly, the complaints against both Coca-Cola and Pepsico were dismissed. Complaint counsel for the FTC appealed the decision to the full Commission. In 1978, the Commission reversed, by a 2 to 1 vote, and, relying in part on the Supreme Court's decision in *Sylvania,* held that certain of the territorial restrictions were unlawful.[90]

---

84 91 F.T.C. at 536–38.
85 *Hearings on S.598, supra* note 75, at 111.
86 *Id.* at 112.
87 *Id.* at 47 (statement of Lee S. Preston).
88 The seven were Coca-Cola, PepsiCo, Crush International, Dr. Pepper Company, The Seven-Up Company, Royal Crown Cola Company, and National Industries, Inc.
89 91 F.T.C. 517 (1978); 91 F.T.C. 680 (1978).
90 91 F.T.C. 607 (1978).

1. The Initial Decision

Administrative Law Judge Dufresne's initial opinion in the *Coca-Cola* case contained 195 detailed findings of fact concerning the nature of the soft drink industry and the effects of territorial restraints on the distribution and sale of soft drinks. The opinion first examined the history and structure of the soft drink industry. It depicted the industry as a dynamic, growing business comprised of manufacturers and bottlers who competed vigorously for consumer patronage. Although the territorial restrictions foreclosed intrabrand competition between licensed Coca-Cola dealers in price, package size, and service, the availability of a wide variety of brands in local markets made for "intense" interbrand competition in these areas.[91] In addition, Coca-Cola competed with other national brands in product availability,[92] merchandising,[93] new product entry,[94] and fountain drinks.[95] The administrative law judge found that elimination of territorial restrictions might lead to short-term price reductions for chain store customers as competition for the larger accounts increased and bottlers switched to warehouse delivery.[96] This competition, however, could also lead to increased prices for soft drinks purchased at smaller outlets, and might result in the disappearance of some brands from certain markets. In Judge Dufresne's view, increased costs and cream-skimming by larger bottlers would inevitably force hundreds of bottlers out of business, producing a "substantial" impact on the economic health of many communities.[97]

After presenting his findings, Judge Dufresne turned his at-

---

91 *Id.* at 548–70.

92 *Id.* at 564. The term "product availability" refers to the total number of outlets from which a consumer may purchase a product.

93 *Id.* at 565–66.

94 *Id.* at 568–69. Thus, soft drink manufacturers must constantly contend with the introduction of new flavors or products by competitors. The advent of low-calorie soft drinks such as "Tab" and "Fresca" provides one example.

95 *Id.* at 563. Soft drinks are still sold at soda fountains and through cup vending machines. Individual bottlers face substantial intrabrand and interbrand competition from these sources.

96 Under the warehouse delivery system bottlers would make one-stop bulk deliveries directly from the bottling plant to the chain store's warehouse. The chain store would then distribute the product to its individual outlets. Bottlers have repeatedly argued that a shift to warehouse delivery would cause a decline in the availability of soft drinks by making service to small accounts unprofitable.

97 91 F.T.C. at 572–79.

tention to the legal issues in the case. Citing *Chicago Board of Trade,*[98] he held that the rule of reason governed the legality of Coca-Cola's territorial restraints. To meet this standard, Coca-Cola had to show (1) that the purpose of the restraints was to further legitimate business goals rather than to hinder competition, (2) that the restraints were no broader than reasonably necessary, and (3) that the restraints promoted competition by regulating it. The fact that the territorial restrictions entirely foreclosed intrabrand competition between licensed Coca-Cola bottlers was insufficient by itself to sustain a finding of unreasonableness.[99]

Judge Dufresne found that Coca-Cola had adopted the territorial restraints for a legitimate purpose — that of broadening the distribution of Coca-Cola by encouraging local businessmen to invest capital and effort into the bottling of Coke. In his view, the restraints continued to serve a useful purpose by making additional investment attractive, fostering high levels of product quality and availability, and encouraging the use of returnable bottles. Moreover, the restraints were no broader than necessary to provide the bottler with a territory he could serve and develop. Finally, the restraints reasonably regulated trade by allowing bottlers to focus on their own territories and by precluding other Coca-Cola dealers from obtaining without cost the benefit of a local dealer's promotional activities.[100]

After reiterating his belief that abolition of the territorial provisions would lead to the failure of small bottlers and an increase in consumer prices, Judge Dufresne concluded that the restrictions did not unreasonably restrain trade. In the end, their adverse effect on intrabrand competition was outweighed by their beneficial effects on interbrand competition in national and local soft drink markets.[101]

2. The Federal Trade Commission Opinion

The first issue which the Commission considered in its opinion was the impact of Coca-Cola's territorial restrictions on intra-

98 246 U.S. 231 (1918).
99 91 F.T.C. at 585–86.
100 *Id.*
101 *Id.* at 588–89.

band competition. Three effects were deemed particularly harmful. First was the simple absence of intrabrand price competition within a single territory; indeed, testimony introduced at trial indicated that bottlers even exerted pressure on dealers in neighboring territories not to undercut their prices.[102] Without direct competition among Coca-Cola bottlers, consumers were denied the opportunity to purchase Coke at competitive prices. Second, the Commission found that the restrictions locked the industry into the "store-door" delivery system[103] and prevented the development of a more efficient central-warehouse-and-plant-pickup distribution system[104] While the latter method of delivery might inconvenience some customers,[105] it would promote substantial competition in the absence of territorial restraints. The maintenance of the store-door system was one aspect of the "quiet life" that the bottlers were enjoying in a noncompetitive environment.[106] Finally, territorial restrictions prevented geographic market expansion. The Commission found that although bottlers could expand their markets by mergers, acquisitions, or agency agreements to supply neighboring bottlers, the territorial provisions constituted an artificial constraint on natural market evolution.[107]

The Commission next examined the bottlers' asserted justifications for territorial restrictions. Noting that the company had failed to present evidence that it and its bottlers could not operate profitably in the absence of territorial restrictions, the Commission rejected the argument that restraints were necessary to promote investment in the industry.[108] Furthermore, Coca-Cola was neither a new entrant nor a failing firm. Had this been the case, the restraints might have been justified as serving the interests of competition.[109] The Commission concluded that "one of the most popular consumer product lines in American industry"[110] would remain an attractive investment even in the absence of restraints.

102 *Id.* at 625.
103 *See* notes 83 and 96 and accompanying text *supra*.
104 91 F.T.C. at 624–25 n.26; *see* note 96 *supra*.
105 91 F.T.C. at 624.
106 *See* note 104 *supra*.
107 91 F.T.C. at 619.
108 *Id.* at 626–27.
109 *Id.* at 627.
110 *Id.*

The bottlers also argued that the route delivery system and exclusive territories permitted availability of Coca-Cola "in every conceivable location."[111] Even though service to such small accounts as vending machines was unprofitable, these outlets developed a consumer preference for Coca-Cola. In addition, bottlers typically charged the same price to all their customers regardless of the cost of serving them. The Commission concluded that the uniform pricing was no more than an internal subsidy of the uneconomic accounts, funded by charges in excess of cost to large-volume customers. As for the market penetration argument, the elimination of restraints would certainly prompt some restructuring in this area, but the benefits from maintaining the status quo did not outweigh the consumer's interest in increased competition.[112]

The bottlers maintained further that territorial restrictions prevented the extraction of "free rider" benefits from local bottlers' advertising and promotional expenditures. The Commission distinguished the Supreme Court's consideration of this argument in the *Sylvania* case by noting that in *Sylvania* the restraints involved persons who sold products directly to the public. The Coca-Cola bottlers, on the other hand, were wholesalers who had no control over the prices ultimately charged to consumers, and who were therefore unlikely to engage in substantial price advertising. Indeed, on the wholesale level, increased intrabrand competition would probably result in the provision of more, rather than less, information to potential customers since bottlers would now be forced to compete for retailers' patronage.[113]

The Commission also rejected the industry's argument that territorial restrictions were necessary to prevent small bottlers from going out of business. In fact, many small bottlers were more efficient than their larger counterparts, and territorial restrictions prevented them from reaping the benefits which accrue to the most efficient producers in an open and competitive market. Furthermore, it was not clear that large, non-local bottlers had a competitive edge over small, local bottlers in ob-

---

111 *Id.*
112 *Id.* at 629.
113 *Id.* at 629–31.

taining larger accounts. The local bottler's proximity to clients as well as a prior course of dealing would counteract any advantages which large bottlers might possess. Even if small bottlers lost central warehouse accounts, there was still ample economic opportunity in competing for customers for whom store-door delivery remained efficient.[114]

Finally, bottlers contended that territorial restrictions promoted product quality by simplifying the syrup company's quality-control program and allowing bottlers to monitor the freshness of their product on supermarket shelves. The Commission found that the syrup manufacturer's ability to guarantee high product quality would not be affected by abolition of territorial restraints. Methods less destructive of competition remained available to protect the manufacturer's interest.[115]

The Commission next considered the role of interbrand competition in the industry. Although in general many brands were available in local markets, this factor alone was an inadequate measure of the intensity of competition since individual bottlers generally distribute soft drink brands trademarked by two or more syrup companies. This practice, known as "piggybacking," is used extensively in the soft drink industry and can decrease dramatically the number of actual competitors in a market.[116] Piggybacking allows the bottler to minimize competition among his own brands by controlling price and by formulating a unified packaging and marketing strategy, and even to prevent the introduction of potentially competing brands into his market:

> It is a policy of the Coca-Cola Company not to license its allied products to bottlers other than Coca-Cola bottlers. . . . Consequently, bottlers have at times elected to piggyback certain brands or flavors of another syrup company, knowing that Coca-Cola allied products would not be intro-

---

114 *Id.* at 653–62.

115 For example, the Commission suggested that Coca-Cola could require its bottlers to place an identifying mark on each bottle, bottle top, or can. This practice would permit spot checks for quality. *Id.* at 633–34.

116. In 1970, 438 of the 726 domestic Coca-Cola bottlers also distributed at least one other soft drink brand. Thus, such brands as Dr. Pepper and 7-Up are distributed by Coca-Cola in certain territories. Similarly, Nestea canned iced tea is marketed under a territorial licensing system composed of about 60 Pepsi-Cola bottlers, about 50 7-Up bottlers, and 30 Coca-Cola bottlers. *Id.* at 636 n.38.

> duced as competitive brands in their territory. Thus Mr. PiBB, respondent's Pepper-type drink, was not introduced in New York City because the Coca-Cola bottler there elected instead to distribute Dr. Pepper. . . . Had Dr. Pepper entered New York via another bottler, such as the Pepsi bottler, the New York Coca-Cola bottler could have responded by introducing Mr. PiBB. Pursuant to The Coca-Cola Company's policy, however, the New York Coca-Cola bottler not only acquired control of the Dr. Pepper brand, it knew that no other bottler would have access to the competing Pepper-type drink, Mr. PiBB. Similarly, Coca-Cola bottlers who handle their own flavor lines understand that competing Fanta flavors will not be introduced in their territories by any other bottler.[117]

Piggybacking increases the concentration of soft drink brands controlled by the largest bottlers in a territory, while the geographic restrictions protect them from the competition of extra-territorial interbrand bottlers[118] The New York Coca-Cola bottler succinctly described the benefits of piggybacking: "We would rather compete with ourselves than have somebody else compete with us."[119]

The Commission found that the territorial restrictions reduced interbrand competition because Coca-Cola bottlers in one market were unable to compete with Pepsico, Dr. Pepper, and other brands in other markets. Indeed, the industrywide nature of such restrictions meant that not only Coca-Cola bottlers but all bottlers were insulated completely from the competition of potential interbrand bottlers not having access to their territories. In view of this fact, and in light of the deleterious effect of the territorial restrictions on intrabrand competition, the Commission concluded that the restrictions substantially lessened competition among soft drink suppliers across the nation in violation of section 5 of the Federal Trade Commission Act.[120]

The relief stage of the proceeding was complicated by the industry's contention that abolition of territorial restraints, coupled with the adoption of the central warehouse distribution

---

117 *Id.* at 637 n.39. The Commission's opinion discusses other instances where piggybacking has led to high concentration.

118 *Id.* at 639.

119 *Id.* at 637.

120 *Id.* at 644.

system, would sound the death knell for the returnable bottle. The Commission found this argument persuasive and issued an order which struck down territorial restrictions insofar as they applied to non-refillable containers, but upheld them insofar as they applied to returnable bottles. The Commission noted that refillable containers — which were both more efficient and less harmful to the environment than refillables — were economically feasible only where manufacturers could be reasonably certain of recapturing their bottles after each use. Extensive intrabrand competition would make this task substantially more difficult.[121]

3. The Impact of the Commission's Decision

The Coca-Cola Company and Pepsico appealed the Commission's orders to the D.C. Circuit Court of Appeals, where the cases are still pending.[122] At the same time, the industry intensified its effort to obtain remedial legislation which would override the Commission's orders and clearly establish the legality of vertical non-price restrictions in the soft drink industry. In its favor, the industry could point to several lower court decisions which had upheld territorial restraints in the soft drink industry;[123] Judge Dufresne's initial opinion; the pervasiveness of such practices throughout the economy; the *Sylvania* decision; and the legal, social, and economic arguments that Coca-Cola and Pepsico had asserted, with mixed results, before the full Commission. Arrayed against the industry was a strong congressional tradition of allowing exceptions to the antitrust laws only in the presence of strong and compelling circumstances. Furthermore, congressional intervention before final adjudication of the Coca-Cola and Pepsico suits would be not only inappropriate, but also wasteful if the FTC's decision were

---

121 *Id.* at 645–52. Commissioner Clanton filed a dissenting statement which recommended that the Commission remand the case for a fuller consideration of the competitive state of the soft drink industry. *Id.* at 589–606.

122 *See* note 4 *supra.*

123 See Coca Cola Bottling Co. v. The Coca-Cola Co., 269 F. 796 (D. Del. 1920); Tomac Inc. v. The Coca-Cola Co., 418 F. Supp. 359 (C.D. Cal. 1976); First Beverages, Inc. of Las Vegas v. Royal Crown Cola Co., 1980-1 Trade Cas. ¶ 63,162 (9th Cir. 1980).

to be overturned in either the court of appeals or the Supreme Court. Remedial legislation might also encourage other parties who were aggrieved by adverse administrative decisions to seek direct relief in Congress instead of the courts. This result would place considerable pressure on an already scarce commodity — congressional time — and would undermine a constitutional structure which intends the courts and not the Congress to be the presumptive aribiter of legal disputes.[124]

III. VERTICAL RESTRAINTS AND THE SOFT DRINK INTERBRAND COMPETITION ACT OF 1980

The industry received a warm welcome in Congress. On March 8, 1979, Senators Bayh of Indiana and Cochran of Mississippi, along with 77 co-sponsors, introduced S. 598, The Soft Drink Interbrand Competition Act of 1980, in the United States Senate. Shortly thereafter, a companion bill, H. R. 3567, was introduced in the House of Representatives, and eventually 310 congressmen joined as co-sponsors.[125] A little over a year later, the Senate and the House passed an amended version of the House bill by margins of 89 to 3 and 377 to 34, respectively. After rejecting advice from his top consumer affairs adviser, the Justice Department, and the Federal Trade Commission, President Carter signed the Act into law on July 10, 1980.[126]

Despite an express purpose to "clarify the circumstances under which territorial provisions in licenses to manufacture, distribute, and sell trademarked soft drink products are lawful under the antitrust laws,"[127] the Soft Drink Acts leaves the courts considerable discretion in assessing the legality of vertical restraints in the soft drink industry. The new law does not declare vertical territorial limitations and customer restraints

---

124 This is not to imply that individuals should not seek to change laws with which they disagree. The courts, however, are the most appropriate *initial* source of relief in cases in which the controversy in question is amenable to judicial resolution. A system based upon separation of powers cannot function effectively where the Congress is forced to be the court of last resort in every case.

125 S. REP. No. 645, 96th Cong., 2d Sess. 9–10 (1980) [hereinafter cited as SENATE REPORT].

126 [1980] ANTITRUST & TRADE REG. REP. (BNA) A–14.

127 Pub. L. No. 96–308, 94 STAT. 939 (1980).

to be lawful. Rather, it provides that these restrictions will be upheld only when a given brand is in *substantial and effective competition in the relevant market or markets.*[128] Furthermore, Congress intended this standard to have real bite. The House Report clearly states that the Soft Drink Act does not create antitrust immunities.[129] It cannot be argued that this statement simply refers to other anti-competitive practices condemned by the antitrust laws, since these are specifically excepted from the scope of the new law.[130] Rather, the clear implication is that Congress did not wish to foreclose the condemnation of vertical territorial restraints in those markets which lacked the vigorous and open competition protected by the antitrust laws. The object of the Soft Drink Act is not to legitimate, but rather to palliate.

A. *Congressional Intent and the Legality of Vertical Restraints*

Analysis of the amended House bill[131] and original Senate bill[132] as well as the Committee reports which accompanied them

---

128 *Id.* at § 2.

129 *Id.* at § 3.

130 *Id.*

131 The amended House bill became the Soft Drink Interbrand Competition Act of 1980. *See* note 1 *supra.*

132 The original Senate version was as follows:

Section 1. This Act may be cited as the "Soft Drink Interbrand Competition Act."

Section 2. Nothing contained in any antitrust law shall render unlawful the inclusion and enforcement in any trademark licensing contract or agreement, pursuant to which the licensee engages in the manufacture (including manufacture by a sublicensee, agent, or subcontractor), distribution, and sale of a trademarked soft drink product, of provisions granting the licensee the sole and exclusive right to manufacture, distribute, and sell such product in a defined geographic area or limiting the licensee, directly or indirectly, to the manufacture, distribution, and sale of such product only for ultimate resale to consumers within a defined geographic area: *Provided,* That such product is in substantial and effective competition with other products of the same general class.

Section 3. The existence or enforcement of territorial provisions in a trademark licensing agreement for the manufacture, distribution, and sale of a trademarked soft drink product prior to any final determination that such provisions are unlawful shall not be the basis for recovery under section 4 of the Act entitled "An Act to supplement existing laws against unlawful restraints and monopolies, and for other purposes," approved October 15, 1914.

Section 4. As used in this Act, the term "antitrust law" means the Act entitled "An Act to protect trade and commerce against unlawful restraints

indicates that Congressional solicitude toward the industry weakened significantly between the original and final versions of the bill. For example, both houses were concerned that the long pendency of the government's action (initiated in 1971) might expose bottlers to significant treble damage liability if the decision striking down territorial restraints were upheld. In that case, bottlers would be liable for damages extending back to 1967, since the institution of a government suit tolled the statute of limitations for treble damages. The original Senate bill completely barred the recovery of treble damages under section 4 of the Clayton Act for violations which occurred prior to a final determination that the challenged practices were unlawful.[133] The amended House version substituted a provision which simply suspended the tolling statute with regard to the FTC proceeding.[134]

The amended House bill added a new provision which specifically excluded price-fixing agreements, horizontal restraints of trade, and group boycotts from the scope of the new law.[135] This change made it clear that Congress intended to deal with a very narrow range of practices and did not wish the courts to infer any weakening of "our beleaguered antitrust laws."[136] Similarly, the "substantial and effective competition" clause of section 2 was strengthened to include the phrase "in the relevant market or markets."

The Senate and House Committee reports present a much sharper contrast in legislative point of view.[137] The report of the Senate Committee on the Judiciary characterizes the soft drink industry as worthy of rescue:

> and monopolies" (the Sherman Act), approved July 2, 1890, the Federal Trade Commission Act, approved September 26, 1914, and the Act entitled "An Act to supplement existing laws against unlawful restraints and monopolies, and for other purposes" (the Clayton Act), approved October 15, 1914, and all amendments to such Acts and any other Acts in pari materia.

[hereinafter referred to as the original Senate bill].

133 *Id.* at § 3.

134 Soft Drink Act, *supra* note 1, at § 4.

135 *Id.* at § 3.

136 H. R. REP. No. 118, 96th Cong., 2d Sess. 7 (1980) [hereinafter cited as HOUSE REPORT].

137 Since the Senate adopted the text of the House bill, the House Committee report is dispositive on the issue of congressional intent. It is instructive, however, to examine the Senate report for an example of an approach which Congress specifically declined to take.

> The committee has concluded that the present territorial franchise system in the soft drink bottling industry can foster effective competition. The committee recognizes that the destruction of the system is likely to depress the value of the franchised bottling plants and cause tremendous economic harm to hundreds of small bottlers who have depended on this system for many years.[138]

The House report, on the other hand, takes a less hospitable view:

> This legislation restates the rule of reason approach followed by the Supreme Court in *Continental T.V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36 (1977). The clarification eliminates uncertainty in the law that has plagued the industry, particularly smaller bottlers, during the last decade. It does not grant antitrust immunities. Indeed, the legislation will apply only in situations in which there is "substantial and effective competition" among soft drink bottlers and among their syrup manufacturers in the relevant product and geographic markets.[139]

Although the Senate report also conditions protection of territorial restraints upon the presence of substantial and effective competition, the more carefully articulated House version sets a somewhat higher standard of conduct. Both reports instruct the courts to weigh the number of brands, types, and flavors of available soft drinks; the degree of service; the number of available retail price options; and the ease of entry.[140] The Senate report also indicates that the courts should consider the number and strength of sellers of competing products in the market; the persistence of long-run *monopoly* profits; the *persistence* of inefficiency and waste; the failure of output levels to respond to consumer demands; and lack of opportunity to introduce more efficient methods and processes.[141] The House version emphasizes the persistence of long-run *anti-competitive* profits; the *existence* of inefficiency and waste; the number and strength of *directly* competing products in a relevant market; and the availability of various forms of containers or packag-

---

138 SENATE REPORT, *supra* note 125, at 2.

139 HOUSE REPORT, *supra* note 136, at 2.

140 *Id.* at 5; SENATE REPORT, *supra* note 125, at 10.

141 SENATE REPORT, *supra* note 125, at 10-11. The Senate report implies that "substantial and effective" competition refers only to interbrand competition. *Id.* at 7.

ing.[142] Although the differences in these two tests are a matter of degree, it is clear that in adopting the text of the House bill, Congress intended the courts to give close scrutiny to competitive conditions in individual markets. Therefore, one should not argue that the Soft Drink Act creates a strong presumption in favor of the legality of territorial restraints. The history of the Act indicates instead that such restrictions should be condoned only where market behavior conforns to the competitive norms protected by the antitrust laws.

The House report asserts that the amended bill merely codifies the *Sylvania* decision.[143] Reference to the text of the bill, however, indicates that Congress has modified the *Sylvania* test in a subtle, yet important, way. The rule-of-reason test endorsed in *Sylvania* focuses upon the nature of the restraint and *its* net effect on competition. Thus, the reviewing court is concerned with the marginal effect of removing the restraint and *not* with the pre-existing level and intensity of competition in the relevant market.[144] Under the Soft Drink Act, the latter factor is all-important. Where the court finds that competition is substantial and effective, the territorial restrictions must be upheld. Where market conditions deviate from the competitive norm, the legality of territorial limitations may be questioned. In this case, the court will presumably proceed to analyze the restraint under the rule of reason. However, this inquiry cannot help but be influenced by the prior determination that the market in question is imperfectly competitive. Consequently, some restraints which might have withstood pre-Soft Drink Act and post-*Sylvania* rule-of-reason analysis may be struck down as unreasonable. For support, the court need look no further than Justice Black's admonition in *Northern Pacific Railway v. United States* that the Sherman Act rests:

> on the premise that the unrestrained interaction of competitive forces will yield the best allocation of our economic resources, the lowest prices, the highest quality and the greatest material progress, while at the same time providing

---

142 HOUSE REPORT, *supra* note 136, at 5-6.

143 *Id.* at 4.

144 Even the vague rule of reason test announced in the *Chicago Board of Trade* case, 246 U.S. 231 (1918), did not direct courts to examine the existing competitive conditions in the relevant industry or market.

> an environment conducive to the preservation of our democratic institutions. But even were that premise open to question, the policy unequivocally laid down by the Act is competition.[145]

The dissenting and additional remarks to the respective Committee reports provide several contrasting views of Congress' true intent. The dissenting statement of Congressmen Edwards, Conyers, and Drinan characterized the amended House bill as "truly legislation of the industry, by the industry, and for the industry."[146] The statement went on to question both the timing of and necessity for the legislation. Legislative change prior to the completion of judicial proceedings would not only disrupt the judicial process, but encourage others to try their cases before Congress. Furthermore, the congressmen disputed the majority's contention that the Act would preserve the small-business character of the industry:

> The facts show that, in any event, a restructured industry is inevitable. The large multiplant firm, not the small bottler, will be the primary beneficiary of this legislation. The industry has gone from 6000 bottlers in 1950 to less than 2000 bottlers today. Coca Cola-owned franchises sell 15 percent of the Coca Cola brands sold nationally; for PepsiCo, a conglomerate, the figure is 20 percent. IC Industries, formerly the Illinois Central Railroad, boasts in a recent Wall Street Journal advertisement that their franchises for one brand alone serve over 18 million people. Coca Cola of New York, a Fortune 500 company, recently spent $85 million on an acquisition program. Other giants, such as General Cinema, General Tire and Rubber, Wotemco, Liggett (formerly Liggett and Myers), Warner Communications, also own bottlers. This is not a list of corporations in need of special treatment from Congress.[147]

However, the congressmen welcomed the addition of the amendments which considerably narrowed the bill's scope. They attached particular significance to the exclusion of arrangements which were per se unlawful under the antitrust laws, since the Supreme Court in *Sylvania* specifically left open the possibility that certain vertical territorial restrictions might still

---

145 356 U.S. 1, 14 (1956).
146 HOUSE REPORT, *supra* note 136, at 14.
147 *Id.* at 15.

fall within the scope of the per se prohibition.[148] Despite such provisions, however, the congressmen's ultimate conclusion was that the American public would be the big loser if the bill became law.

Congressman Hall and twelve colleagues[149] filed a statement of additional views which endorsed the interpretations contained in the Senate report. The congressmen asserted that the current state of the industry satisfied the substantial-and-effective-competition test and that the weight of legal and economic authority favored upholding the restrictions. Furthermore, they asserted that courts should not consider the absence of intrabrand competition as a factor in judging the reasonableness of territorial limitations.[150] The tone of their remarks indicates concern that the bill does not insulate the industry's questionable practices from antitrust scrutiny. Their suggestion that courts should not consider the detrimental effects of territorial restraints on intrabrand competition stands out as an obvious attempt to stack the judicial deck in the industry's favor.[151]

Two other congressmen filed supplemental views. Congressman McClory asserted that the bill simply assured that the rule of reason would govern consideration of territorial restraints, with the legislation changing only the formulation and not the substance of the rule-of-reason standard.[152] The congressman also felt the addition of "in the relevant market or markets" to section 2 of the Act would allow effective judicial review of territorial restrictions.[153] Congressman Railsback applauded the bill for placing the burden of proof on the party seeking to establish that a vertical non-price restriction constituted an antitrust violation. The FTC proceeding had placed the burden of showing business justification for the restraints on the in-

---

148 *Id.* at 16. Although the text of the Soft Drink Act does not expressly exclude per se unlawful activities from the scope of the Act, the majority report of the House indicates that this exclusion was intended by the Congress. *Id.* at 6.

149 Congressmen Brooks, Volkmer, Synar, Glickman, Evans, Fish, Butler, Hyde, Kindness, Sawyer, Lungren, and Sensenbrenner.

150 House Report, *supra* note 136, at 18.

151 *Id.*

152 *Id.* at 21.

153 Congressman McClosky also stated that this phrase "must be narrowly construed as a matter of logic." *Id.* at 22.

dustry.[154] In Congressman Railsback's view, this was a misapplication of *Sylvania*.[155]

The Senate report on the initial version of the bill prompted an acerbic response from Senators Metzenbaum and Kennedy. The senators first argued that the courts should remain the final arbiters of the reasonableness of the restraints. If the net effect of the restraints was pro-competitive, then the courts would ultimately hold them lawful. Second, fair enforcement of the antitrust laws necessitated the application of the same rule-of-reason test to all industries. Absent a showing of unique and compelling circumstances, the industry did not merit special treatment. Finally, the senators contended that the trend toward concentration both at the syrup manufacturing and bottler levels of the industry, as well as the absence of price competition in many local markets, meant that perpetuation of the soft drink industry's territorial restrictions would continue to work to the serious detriment of consumers.[156]

Since the Senate eventually adopted the text of the amended House bill, the House report is dispositive on the issue of congressional intent. The House report indicates that Congress intended to codify the rule-of-reason test announced in *Sylvania* without immunizing industry practices from antitrust scrutiny. The law does not formally reverse the Federal Trade Commission's decisions in the *Coca-Cola* and *Pepsico* cases. It does, however, increase the possibility that the D.C. Circuit Court of Appeals will remand the cases to the Commission to consider whether substantial and effective competition is present in the industry and to determine the scope of the relevant product and geographic markets. While nothing in the new law prevents the Commission from determining that the nation as a whole is the relevant market, such a finding would be imprudent given the strength of the congressional reaction to the wholesale invalidation of territorial restraints in the *Coca-Cola* and *Pepsico* cases. Instead, the Commission should limit its industrywide investigation to one or a few fairly localized markets. This type

---

154 *Id.* at 24; *Coca-Cola*, 91 F.T.C. at 670.
155 House Report, *supra* note 136, at 24.
156 Senate Report, *supra* note 125, at 18.

of inquiry would permit the Commission to construct a much clearer picture of both the intensity of competition and the effects of territorial restraints on interbrand and intrabrand competition. The conclusions which resulted from such an inquiry, whether positive or negative, would be less amenable to attack in the courts and considerably more persuasive to those whose economic interests are affected.

Since the Soft Drink Act leaves several important factual and legal issues unresolved, it is useful to re-examine the asserted justifications for and arguments against legitimizing vertical non-price restraints. Most of the arguments discussed below were raised in various forms either during the FTC proceeding or before Congress. The force of each was affected by the new law.

B. *Asserted Justifications for Legitimizing Vertical Non-Price Restraints*

Although many of the arguments which follow can be subsumed under the heading of "microeconomic efficiencies," the discussion is not intended to suggest that such arguments merit exclusive or even primary weight in legislative and judicial determinations concerning vertical non-price restraints. The framers of the antitrust laws were concerned with social as well as economic factors, and the goal of a decentralized and fair economic system remains important today.

1. Vertical Restraints Protect Bottlers' Expectations

This argument maintains that it is unfair to defeat expectations upon which bottlers have based investment decisions for over 75 years. It recognizes that the bottler's intrabrand monopoly reduces the financial risk of doing business. The value of this reduced risk is reflected in the market value of the bottler's franchise rights. Thus, abolition of territorial restrictions would assertedly inflict an unfair capital loss on current franchise holders.

The easy response to this argument is that the market price which the bottler paid for his intrabrand monopoly reflected the

discounted probability that the territorial restraints would be held to violate the antitrust laws. In this case, the bottler simply lost his gamble. One cannot, however, avoid the fact that many individuals have placed real and substantial reliance upon their expected ability to recover the value of their franchise investment. While this reliance must be weighed against the loss in consumer welfare generated by inefficient constraints on competition, a basic notion of fairness favors finding some way to ease the blow. The most convenient answer to this dilemma — to leave the status quo intact — fails because it allows bottlers to continue to benefit at the expense of efficiency and consumer welfare. A more acceptable approach would be a gradual restructuring of the current system. This solution would recognize the interests of both the public and the franchise holder: it would reduce the efficiency losses resulting from imperfect competition while permitting franchises to return some short-run residual value.[157]

2. Vertical Restraints Preserve Small Business

Proponents of the "protect small business" view inevitably point to Judge Learned Hand's observation in the *Alcoa* case that one purpose of the antitrust laws is to "perpetuate a . . . decentralized economy, even if this entails a measure of economic inefficiency."[158] One industry witness, in testimony before the Senate Committee on the Judiciary, predicted that the following developments would occur if territorial restraints were eliminated:

---

157 Assume that the Supreme Court, in 1981, declares territorial restraints in the soft drink industry unlawful under the rule of reason (after first finding that soft drinks are not in substantial and effective competition on a nationwide basis). If Congress concurred in this assessment, yet desired to ease the blow on industry, it could pass hardship legislation which would effectively enjoin enforcement of the Court's judgment for, say, a period of five years. This would allow the bottlers to recover some of their investment and to adjust to the new competitive rules of the game. Invalidation of restraints would pose some of the same problems as deregulation in the transportation industry. In determining the course of deregulation, Congress had to grapple with the destruction of 40-year-old "property rights." There are, of course, tremendous differences between antitrust regulation of private contractual arrangements, and direct government intervention and control in cases of perceived market failure.

158 *See* note 68 *supra*.

(1) extensive vertical integration by the major-brand franchisors,

(2) backward integration by major chain stores,

(3) geographic market expansion by the largest and strongest established bottlers,

(4) disappearance of minor brands from the market, and

(5) disappearance or substantial contraction of a large number of smaller, but currently viable and profitable, bottling firms throughout the country[159]

This prediction, however, rests on the questionable assumption that small entities cannot compete effectively with large enterprises in the absence of territorial restrictions. Furthermore, the argument that abolition of restraints would result in increased concentration has force only insofar as centralization would *not* occur if restraints were retained.

In the *Coca-Cola* and *Pepsico* cases, the Federal Trade Commission concluded that small bottlers are often just as efficient as — and sometimes more efficient than — their larger counterparts.[160] The fact that local bottling plants are likely to be closer to chain-store and other large-account warehouses gives them an advantage in transportation costs. In addition, local bottlers are likely to have built up substantial good will with their principal customers through the years. These facts belie the notion that small bottlers would not survive the elimination of territorial restraints. Indeed, one might argue that such restraints actually harm small bottlers by preventing them from expanding into neighboring territories and achieving sufficient volume to capture scale economies in production.

The "cream-skimming" argument — that large bottlers will invade the markets of their smaller competitors and take the most profitable accounts — raises a deeper question. The very existence of "cream" to skim indicates that bottlers have been able to charge certain customers prices in excess of reasonable costs. Bottlers respond that extra revenue from these large accounts is used to subsidize such uneconomic outlets as vending machines and "Mom-and-Pop" groceries. Should the source of

---

159 *Hearings on S. 598, supra* note 75, at 53 (statement of Lee S. Preston).
160 *See* text accompanying notes 111 to 114 *supra*.

the subsidy disappear, the argument goes, bottlers would be forced to discontinue serving these outlets. However, elimination of the cross-subsidy need not lead to discontinuance of service. If soft drinks were priced to reflect their true costs, the price of supermarket soft drinks would decline and the price of convenience soft drinks would increase. As a result, consumption of supermarket soft drinks would increase and convenience soft drink consumption would decline. This poses yet another problem since most vending machine operators and convenience outlet owners are themselves small businesses. The elimination of territorial limitations on bottlers, if it led to efficient pricing, might jeopardize an entirely different group of small businessmen. But even these effects may not occur since vending machine owners might welcome a "walls down" market with the intent of expanding into the bottling business themselves.

One could easily argue that the preceding analysis of second- and third-order effects of eliminating territorial limitations pushes the small-business argument beyond its useful limit. However, this analysis does indicate the complexity of the effects involved. Thus, while a deconcentrated economy may be a laudable goal, it is impossible to say that elimination of territorial restraints would injure small businesses.

Moreover, it is quite likely that concentration of the bottling industry will occur even if territorial restraints are retained. Thhere is little disagreement that such concentration is already increasing at a rapid rate. The same industry witness who decried the consolidation that would result from invalidation of territorial restraints informed the Committee earlier in his testimony that:

> Almost 2,000 small soft drink bottling plants have disappeared over the period 1963-1978, while the number of larger plants has increased substantially. Census data reveal that a predominant share of both employment and value added by manufacturing has correspondingly shifted from small to larger plants over the same period.[161]

Between 1963 and 1978 the number of bottling companies decreased from 3,569 to 1,724.[162] At the syrup-manufacturing level,

161 *Hearings on S. 598*, *supra* note 75, at 46 (statement of Lee S. Preston).
162 *Id.* at 47.

the four largest companies comprised 62 percent of the market in 1963 and 72.4 percent in 1977.[163] The twenty largest companies commanded 70 percent in 1963 and 95.1 in 1977.[164] These figures hardly support the proposition that territorial restrictions have had a significant effect in preventing the disappearance of small businesses.

Whether abolition of territorial restrictions would lead to increased vertical integration by syrup companies or backward integration by supermarket chains is quite speculative, although the occurrence of either or both would inarguably increase the size and economic power of firms in the industry. The primary argument against the realization of either eventuality is that the present system does not prevent either syrup manufacturers or supermarkets from purchasing existing bottlers. Simple business considerations constitute the most effective damper on such activity. Supermarkets deal in thousands of individual commodities. It seems unlikely that they would be any more interested in bottling Coke for the soda aisle than in raising cows to stock their meat counters. Syrup manufacturers also gain substantial benefits through the use of independent distributors. Bottlers share the financial risk of the business, constitute a ready source of capital, and possess a knowledge of local conditions which the parent company could not duplicate. This state of affairs would not change dramatically in the absence of territorial restrictions.

The one development which vertical territorial restrictions do prevent is the geographic-market expansion of soft drink bottlers. However, removal of the restrictions is unlikely to lead to expansion by only the largest and best-established bottlers. Many small bottlers have remained small simply because territorial limitations have prevented them from growing. Similarly, bottlers may be financially weak because vertical restraints have prevented them from expanding in search of more fertile markets. Finally, a "walls down" market might encourage entirely new enterprises to attempt to enter the market. The industry

---

163 These figures may overstate the degree of market concentration because most major supermarket chains market private-label soft drinks. However, most of these drinks are produced under contract with independent bottlers and not by entities controlled by the supermarkets themselves. *Id.* at 112 (statement of William S. Comanor).

164 *Id.*

has consistently avoided the conclusion that unfettered geographic market expansion is essentially pro-competitive.

The bottom line is that in all likelihood territorial restraints do not foster the preservation of small business units and a decentralized economy. In the absence of convincing evidence to the contrary, the presumption should be in favor of a free marketplace.

3. Vertical Restraints Promote Interbrand Competition

The most commonly asserted argument in favor of vertical restraints is that they foster interbrand competition. In the *Sylvania* case, the Supreme Court suggested that vertical non-price restraints might (1) aid new entrants in raising capital and winning consumer acceptance of an unfamiliar product, and (2) encourage established dealers to provide valuable services without fear of intrabrand competitors' gaining a "free ride."[165] In the soft drink industry, these benefits are likely to be minimal. The new entrant gains no advantage by offering territorial restrictions when pre-existing industry members employ similar restrictions: the new entrant's restraints merely offer the same reduction in risk that potential dealers could obtain by investing their capital and energies in any of the established enterprises. Thus the new entrant is in no better shape than where all market participants are precluded from employing restraints.

Insofar as the "free-rider" argument relates to the provision of service or repairs, it has merit only in those industries where services are important. Such amenities play an insignificant role in the soft drink industry. A more pertinent inquiry concerns the potential for "free-rider" effects in the provision of advertising. Where a particular bottler has invested considerable sums of money to develop a favorable consumer attitude toward his product, it would clearly seem unfair for a neighboring competitor to reap the benefits of these expenditures by selling to customers in the first bottler's territory. To preclude this possibility, the industry argues, bottlers will cease to advertise their products, and consumers will be deprived of useful information.

---

165 *See* text accompanying notes 111 to 114 *supra*.

However, the social value of bottler-sponsored advertising is questionable, since the two types of advertising which are prevalent in the industry — image advertising and price advertising — are likely to be provided in the absence of bottler participation. Manufacturers have not only a strong incentive to promote strong image identification for their products, but also an efficient means of spreading this cost among all those who benefit from such expenditures. Retailers are the primary source for price advertising since they, and not bottlers, are the ones who determine the ultimate price of the product.

The industry has also maintained that territorial limitations promote interbrand competition by improving market coverage. The essence of this argument is that dealers will do a better job of developing new business if they are only permitted to focus on a narrow geographic area. The industry argues that if bottlers are allowed to sell wherever they wish, only the large and more profitable accounts will be served. But the likely result of increased competition for such "cream" accounts is that prices will fall and these accounts will lose their relative attractiveness over smaller, less profitable accounts. Thus, bottlers will find it worthwhile to pursue smaller accounts with increased vigor. In addition, removal of the current cross-subsidy of small accounts will inevitably lead to an increase in the price of soft drinks purchased through convenience outlets. In economic terms, this is not an undesirable result, since the current underpricing of these soft drinks has caused a misallocation of resources.

Bottlers argue further that widespread availability of their products promotes consumer awareness and encourages additional purchases through other outlets. The elimination of territorial restraints would not necessarily force bottlers to reduce sales through these outlets. It would simply force the industry to view such services as a promotional activity and not as a profit-producing activity. Moreover, it is far from clear that vending-machine and convenience-outlet sales are inherently unprofitable. The introduction of a true competitive spur might lead to the development of new production and marketing techniques which could make such activities financially attractive in their own right.

Finally, bottlers argue that territorial restraints aid interbrand

competition by reducing the costs of supervising product quality. Syrup companies currently monitor product quality through surprise plant visits and spot-sampling of retailer inventories. The industry asserts that the cost of this program would increase dramatically in the absence of territorial restraints, since syrup companies would be unable to identify the source of a given batch of soft drinks. In the *Coca-Cola* case, the FTC indicated that this problem could be eliminated by simply requiring each bottler to place an identifying mark on the bottle or bottle cap.[166] The presence of this cheap, less restrictive alternative suggests that the actual effect of territorial restraints in the product-quality area is at best slight.

The total benefit of territorial restraints on interbrand competition in the soft drink industry is minimal. This is not to say that such restraints will never have substantial pro-competitive effects. Indeed, the arguments made by the bottlers may have considerable force in industries where, as in *Sylvania,* manufacturers impose vertical non-price restraints upon franchisees who deal directly with the consuming public. The argument in favor of vertical restraints is also stronger in industries where the consequences of defective manufacture, inadequate product availability, or the "free-rider" phenomenon are particularly severe.

4. Vertical Restraints Protect the Returnable Bottle

The use of the returnable bottle is typically justified on ecological and economic grounds. The ecological impact of the returnable bottle is dependent upon the average number of times it can be re-used. At two trips, the refillable bottle has approximately the same effect on the environment as the non-refillable, non-returnable bottle. At four trips, its impact equals that of the conventional steel can. At five trips its impact is roughly that of an aluminum can which is recycled at an 80 percent rate. A returnable bottle with ten or more trips is more energy efficient than any other type of drink package.[167]

---

166 *See* text accompanying note 115 *supra*.
167 91 F.T.C. at 666-67.

The returnable bottle's economic advantage is its low cost. Bottlers can charge retailers less for returnables because they are able to spread the package cost over a larger number of trips whereas the total cost of a non-returnable must be recovered in one shot. As a result, soft drinks in returnable bottles are generally priced significantly lower than soft drinks in non-returnables.[168]

Territorial restrictions are relevant for two reasons. First, they enhance bottlers' ability to recapture used bottles, thus increasing the average number of trips and reducing the amount which bottlers have to invest in bottle inventories. Second, returnable bottles are primarily marketed by small bottlers. Territorial restrictions encourage the continued use of the returnable to the extent which they preserve the existence of the small bottler. The Federal Trade Commission found both the economic and ecological arguments persuasive. Accordingly, the Commission upheld territorial restrictions insofar as they related to soft drink distribution in returnable packages.

The Commission's decision is vulnerable in two respects. First, retention of territorial restraints does not guarantee that small bottlers — or any bottlers — will continue to use returnable bottles in the future. Second, environmental goals are best served directly, not through the manipulation of the antitrust laws. The Federal Trade Commission only undermines its institutional legitimacy when it bases its judgments on factors outside its designated area of expertise. Few would disagree that encouragement of ecologically sound production methods is a laudable social policy. However, the proper way to achieve this goal is through direct legislation which would require the use of returnable packages, and not through subversion of the marketplace.

C. *The Perceived Evils of Vertical Restraints*

Opponents of territorial limitations typically assert four arguments: (1) airtight territorial limitations eliminate intrabrand competition and reduce interbrand competition; (2) territorial limitations, when combined with extensive brand "piggyback-

168 *Id.* at 646-47.

ing," increase market concentration and impose barriers to entry; (3) territorial limitations increase product differentiation; and (4) vertical territorial limitations are really horizontal market allocation schemes in disguise.

1. Airtight Restraints Eliminate Intrabrand Competition and Reduce Interbrand Competition

By their very nature, territorial limitations preclude intrabrand competition. While the FTC accorded this factor considerable weight in the *Coca-Cola* and *Pepsico* cases, both the Supreme Court and Congress have indicated that the primary focus of antitrust concern should be on the vitality of interbrand and not intrabrand competition.[169]

However, the FTC found that in the soft drink industry territorial restraints also reduced interbrand competition. The detriment to the competitive process arises from the fact that brand A soft drink bottlers in territory X are unable to compete with the bottlers of brands B, C, D, and E in territory Y. In other words, removal of territorial limitations would increase the number of competing sellers in a given market and would improve the quality of competition in that market over and above any beneficial effects accruing from increased intrabrand competition between bottlers of brand A. Neither the Supreme Court in *Sylvania* nor the Congress explicitly considered this argument. Such considerations are important in the soft drink industry since in many markets the entry of even one additional seller could improve substantially the competitive character of those markets.

2. Territorial Limitations, When Combined with "Piggybacking," Increase Market Concentration and Raise Barriers to Entry

Industry spokesmen typically argue that piggybacking[170] increases interbrand competition by easing entry of new brands into local markets. By licensing an existing bottler, the new

169 *See* text accompanying notes 52 to 71 *supra*.
170 *See* text accompanying notes 116 to 119 *supra*.

entrant gains the advantage of an existing system of distribution, an established clientele, and experience in producing and distributing similar products. However, the number of sellers, not the number of brands, determines the degree of competition. Individual bottlers are not apt to compete against themselves; instead, they are likely to design their marketing strategy to maximize their full sales at the expense of other bottlers' lines. Furthermore, some bottlers have used piggybacking to foreclose entry of new brands into local markets. The presence of many brands in a market creates the illusion of competition.[171] The mere illusion of competition does not guarantee that competition is in fact occurring or that consumers are paying competitive prices.

Extensive piggybacking may also raise significant entry barriers against new bottlers. In general, retailers prefer to deal with a small number of wholesalers. In markets where retailers can satisfy their demands for a full line of flavors by dealing with two or three multi-brand bottlers, they are less likely to purchase from a new entrant who can only offer one brand. This suggests that piggybacking may foreclose entry by new bottlers on less than a full line basis (i.e., one cola, one ginger ale, one orange, one diet, etc.). The presence of territorial limitations exacerbates this problem by precluding competition from bottlers whose total product might vary from pre-existing bottlers. For example, suppose a retailer seeks to stock his shelves with flavors A, B, C, D, E, and F. Assume further that three bottlers are present in the market. Bottler X sells XA and XB, bottler Y sells YC and YD, and bottler Z sells ZE and ZF. Our retailer satisfies his needs by purchasing XA and XB from A, YC and YD from Y and ZE and ZF from Z. Now suppose a new entrant, N, seeks to market an equivalent version of D called ND. Unless he can offer the retailer a price advantage over YD, the retailer is likely to continue purchasing his D from Y, if only for the convenience of having to deal with three rather than four bottlers. In a market governed by territorial restrictions, this is the end of the story unless our new entrant

171 This illusion was convincing enough to lead the administrative law judge in the *Coca-Cola* case to find the existence of intense interbrand competition in the soft drink industry on the basis of this fact. 91 F.T.C. at 549-53.

can acquire the capital to produce NC as well. Abolition of territorial restraints would provide the opportunity for other existing bottlers to expand into the market. In our example, N's chances of success would greatly improve if bottler MX, who markets XA, XB, and YC in a neighboring community, enters the market. In such a case, our retailer could shift his purchases from X and Y to M and N without increasing the number of merchants he has to deal with. Similarly, the entry of another bottler P, who offers YC, YD, and ZF would create the opportunity for successful entry by a new E producer. The ultimate result would be a market which better conforms to classical notions of dynamic competition.

3. Territorial Restrictions Increase Product Differentiation

Product differentiation[172] occurs where a manufacturer seeks to create the impression that his product is unique so that consumers will be less likely to switch to other brands if the price is increased. For example, assume that consumers generally drink either M cola or O cola; the colas taste similar; and they each cost 35 cents per can. Presumably, an increase in M cola to 50 cents per can would prompt consumers to drink more O cola. Now suppose that the M company changes its name to Millionaire's Cola and undertakes an advertising campaign designed to convey the notion that successful people drink Millionaire's. If M's campaign were successful, it could raise its price without causing consumers to flock to O cola. Product differentiation not based on differences in product quality is inefficient and results in a misallocation of economic resources.

The battle for product differentiation is quite intense in the soft drink industry. Among soft drink manufacturers, product differentiation is seen not as a means of raising prices to increase revenues, but as a hedge against price reductions by competitors. To the extent that vertical territorial restrictions allow bottlers to obtain higher mark-ups, they intensify product differentiation by providing capital to fund the requisite promotional activities.[173] The high degree of product differentiation in

---

172 For a sophisticated discussion of this subject, *see* Comanor, note 6 *supra*.
173 *Id.*

the soft drink industry implies that the prices of most brands exceed the levels which would obtain in a perfectly competitive market. The abolition of vertical restraints could mitigate the effects of product differentiation by stimulating greater competition within the industry.

4. Vertical Territorial Limitations Are Actually Horizontal Market Allocation Schemes in Disguise

One way to argue that vertical territorial restrictions are unlawful is to analogize them to a practice which is condemned per se under the antitrust laws: horizontal market division.[174] Although it is clear that territorial restrictions in the soft drink industry were originally imposed by manufacturers, the Supreme Court has recognized that arrangements which are initially innocuous may ripen into unlawful restraints on trade.[175] This is precisely the suggestion made by Senators Kennedy and Metzenbaum in their dissenting statement to the original Senate bill:

> The noticeable absence of support from syrup manufacturers like the Coca-Cola Company and PepsiCo., Inc., in marked contrast to the strong support for S. 598 aggressively and continually voiced by representatives of the bottling companies, strongly suggests that the soft drink industry's territorial restrictions no longer promote interbrand competition, but instead give cloaks of legitimacy to clearly illegal horizontal market divisions.[176]

Antitrust law generally affords vertical restraints more generous treatment than horizontal market divisions. This policy reflects the notion that, in the case of vertical restrictions, the manufacturer's economic interest dictates that dealer mark-ups be minimized and retail prices kept down. Horizontal divisions among dealers, on the other hand, typically have only one goal

---

174 United States v. Topco Associates, Inc., 405 U.S. 596 (1972); United States v. Sealy, Inc., 388 U.S. 350 (1967).

175 United States v. General Motors Corp., 384 U.S. 127, 136 (1966); United States v. Jerrold Electronics, 187 F. Supp. 545 (E.D. Pa. 1960), *aff'd per curiam*, 365 U.S. 567 (1961).

176 SENATE REPORT, *supra* note 125, at 21. The FTC rejected a similar argument in the *Coca-Cola* case.

— to maximize dealer mark-ups. This distinction breaks down where the manufacturer also has an interest in increasing dealer mark-ups. Such an interest may arise in markets characterized by a high degree of product differentiation.[177] Where such markets are concerned, it would be imprudent to assume that vertical restraints are less pernicious than horizontal restraints. Some would argue that this reasoning could be applied to the soft drink industry.

IV. The Long Run Impact of the Soft Drink Act

At the outset, this Note argued that despite the motivations that lay behind its enactment, the Soft Drink Act would lead to the disappearance of vertical territorial limitations from the industry. This conclusion is based on four assumptions: (1) Congress validated territorial limitations only where competition is vigorous and atomistic; (2) the substantial and effective competition test is less generous than the *Sylvania* test in that courts must now focus upon the absolute level of competition in a market in addition to the net effect a given restraint has upon that competition; (3) courts will give the substantial and effective competition test serious weight in their deliberations; and (4) the absence of competitive conditions in some markets will lead to judicial invalidation of some restrictions and the eventual disappearance of territorial limitations. The validity of these premises will now be examined.

That Congress did not intend to immunize vertical non-price restraints in the soft drink industry from antitrust scrutiny is apparent both from the text of the law and from the House report.[178] Indeed, it was probably the narrow scope of protection afforded by the final version that prompted Congressman Hall's additional views. The congressman and his colleagues had apparently hoped to pass a bill which would insure that industry practices remain inviolate regardless of competitive conditions in the marketplace. However, it seems clear that in the end Congress' acceptance of territorial restraints was conditioned

177 Comanor, *supra* note 6, at 1427.
178 *See* text accompanying notes 135 to 142 *supra*.

upon what it perceived as a fairly high level of interbrand competition — a level sufficient to protect consumers from serious welfare losses resulting from excessive prices.[179] It is equally clear that Congress did not desire courts to uphold vertical non-price restraints where interbrand competition no longer imposed competitive discipline on the marketplace. Once vigorous and atomistic competition has disappeared, vertical non-price restraints are to be viewed with great suspicion.

The second assumption may be restated as follows: the Soft Drink Act is no more, and may be significantly less, protective of territorial restraints than the rule-of-reason test endorsed in *Sylvania*. That inquiry required courts to weigh a restraint's detrimental intrabrand effects against its pro-competitive interbrand benefits and determine whether the net effect was to favor efficiency and competition. Congress intended to codify this basic approach in the Soft Drink Act.[180] But it made a significant change in the test by restating it in terms of extant levels of competition in the relevant market. Under the new law, the rule-of-reason analysis will be applied only where the court has already found that the relevant market is not competitive. The absence of competition will make it harder for the industry to prove that territorial limitations in fact benefit competition. If the *Sylvania* test does not, in its present form, preclude courts from striking down territorial limitations in oligopolistic markets, the substantial-and-effective-competition test compels it.

While the Soft Drink Act clearly establishes substantial and effective competition as the test for determining whether territorial limitations are vulnerable to antitrust scrutiny, Congress failed to provide a bright-line standard for determining whether this test is satisfied. To a great extent, the continued legitimacy of the restraints depends on how strictly the courts apply the many criteria enumerated by Congress.[181] The easy route for a court to take is to interpret the relevant indicia broadly in order to find substantial competition and thereby preserve the status quo. But this approach is clearly not what Congress intended. The courts should recognize that the Soft Drink Act

---

179 *See* text accompanying notes 131 to 156 *supra*.
180 House Report, *supra* note 136, at 2.
181 *See* text accompanying notes 140 to 145 *supra*.

is but one piece of the larger fabric of antitrust law and should apply its mandate with vigor.

The irony of the Soft Drink Act is that it seals the fate of the small bottlers whom it was designed to protect. Had Congress upheld the Federal Trade Commission's order, small efficient bottlers would have been given the chance to expand into new markets, exploit scale economies, and, in all likelihood, survive. The new law precludes such healthy competition by locking small bottlers into their territories and, in many cases, into inefficient production methods. The current trend toward concentration in the industry will undoubtedly continue as capital-rich investors buy out small bottlers and acquire larger territories. It is unlikely that any of the present small bottlers will be able to outbid larger bottlers and conglomerates for additional franchise rights. Bottler support for the new law was not motivated by concern for the long-run economic viability of small businesses, or even by the fear that small bottlers could not successfully compete with their larger counterparts. Rather, their only goal was to assure that they could capture the value of their intrabrand monopolies. Clearly, large bottlers will be willing to pay a premium to acquire neighboring territories which will allow them to capture scale economies in production in all their markets. Furthermore, they will recognize that territorial consolidation greatly reduces intrabrand competition which occurs near territorial dividing lines.

For example, assume the Langdell River divides the city of Ames into two neighborhoods, East Ames and West Ames. Suppose further that bottler X holds the franchise for YokoCola in East Ames and bottler Y the franchise for West Ames. Bottler X's YokoCola prices provide a competitive check on bottler Y's prices in the following way. If Y raises the West Ames YokoCola price too far above the East Ames price, consumers will simply cross the bridge into East Ames and buy their YokoCola there. If Y buys out X, however, he can raise YokoCola prices up to the level where consumers will start switching to other brands of cola. If Y has successfully created a unique image for his product in the minds of consumers (i.e., product differentiation), the range of pricing discretion could be substantial. This example illustrates how both the acquiring and acquired bottler benefit from increased industry consoli-

dation. The acquiring bottler gains scale economies and an aggregate intrabrand monopoly which affords greater market power than the sum of the two separate franchises. The acquired bottler not only recoups his initial investment in purchasing his intrabrand monopoly but also receives a premium reflecting part of the synergistic gains which accrue to the acquiring bottler.[182] Note that in the absence of territorial limitations, X's franchise is worth considerably less to Y since nothing prevents a new YokoCola bottler from entering either Ames market at any time. In that case, X will lose the ability to earn a premium on the sale of his territorial rights.

Eventually, the continuing consolidation of the industry will cause competition in certain markets to dip clearly below the substantial and effective level. If restraints are then challenged, the reviewing court will apply a *Sylvania* analysis. By that point, several justifications for vertical territorial limitations will have disappeared or lost their persuasive force. For example, the disappearance of small firms from the industry removes one of the most powerful arguments in favor of territorial restraints. Similarly, technological advances may obviate the ecological and economic advantages of the returnable bottle. In this context, the court is likely to find that vertical territorial restraints have outlived their social and economic usefulness and hold them unlawful. The invalidation of restrictions as to one bottler gives him the freedom to invade other markets. The astute bottler will try to enter new markets in ways which do not violate the existing dealer's franchise agreement with the syrup manufacturer. Alternatively, the unrestricted bottler might seek to enter highly concentrated markets, in which legal challenges based on the territorial provisions are likely to fail. Either way, bottlers still subject to the restraints will now desire to expand into the markets of those who are already free to sell to whomever they please. The eventual result will be the elimination of territorial restrictions throughout the industry. The unfortunate aspect of this scenario is that competition will indeed be restored, but that the nature of the industry is likely to have changed dramatically. The characteristic industry unit will no

---

182 Of course, these gains are over and above gains Y makes through his increased share of the greater Ames market.

longer be the family-operated or local bottling company, but rather the multi-billion dollar conglomerate.

V. Conclusion

Although it is clear that the Soft Drink Act does not create antitrust immunity for the soft drink industry, it does create a separate standard of conduct under the antitrust laws. This result is unfortunate in three respects. First, it erodes the notion that the antitrust laws provide a uniform and non-discriminatory standard of conduct for all industries. Second, it encourages other industries to seek similar protection for practices "unique" to and "essential to the survival" of those industries. If these pleas are heeded, the antitrust laws will become nothing more than a hopeless patchwork of incomplete rules and unprincipled exceptions. Finally, the Act will probably not ameliorate the concerns which underlay its passage. The Act does not guarantee the continued use of the returnable bottle or the survival of small bottlers. Further, it does not address the ever-increasing concentration at the syrup manufacturing level. Finally, it does not even assure the continued validity of vertical territorial restraints in the industry. The considerable loss in consumer welfare, in competitiveness of the marketplace, and in uniformity of the antitrust laws can only lead to the conclusion that the Soft Drink Act was a regrettable response to the problem of vertical restraints in the soft drink industry.

NOTE

RESTRICTIVE MEDICAL MALPRACTICE COMPENSATION SCHEMES: A CONSTITUTIONAL "QUID PRO QUO" ANALYSIS TO SAFEGUARD INDIVIDUAL LIBERTIES

HOWARD ALAN LEARNER*

During the mid-1970's, increasing medical malpractice insurance rates raised concerns about higher costs to patients and the availability of physicians in some areas. Legislatures in all states responded to what was perceived to be a crisis by enacting restrictive medical malpractice compensation schemes. Mr. Learner argues that utilizing an intermediate level of scrutiny to examine due process and equal protection claims challenging the constitutional integrity of these statutes is inadequate and proposes that courts apply a quid pro quo *standard instead. This analysis, similar to that used in workmen's compensation and nuclear accident compensation schemes, requires that medical malpractice compensation statutes provide a "reasonably just substitute" for a patient's abridged common-law remedies.*

INTRODUCTION: THE MEDICAL MALPRACTICE INSURANCE "CRISIS" AND ITS CAUSES

A. *The Purported "Crisis"*

The tumultuous medical malpractice insurance "crisis" came to public attention in 1975 amidst unprecedented news media coverage of striking physicians and anxious hospital administrators. Continually escalating malpractice insurance premiums had culminated in drastically soaring increases that purportedly made it difficult for some individual practitioners to obtain adequate coverage at reasonable rates.[1] Panic was further inflamed

---

* Howard Alan Learner, A.B., University of Michigan, 1976; J.D., Harvard Law School, 1980; Staff Attorney, Business and Professional People for the Public Interest, Chicago, Illinois.

1 Between 1960 and 1970 malpractice premiums rose 540.8 percent for physicians other than surgeons and 949.2 percent for surgeons. Blaut, *The Medical Malpractice Crisis — Its Causes and Future*, 44 INS. COUNSEL J. 114 n.3 (1977). *See* U.S. DEP'T OF HEALTH, EDUCATION & WELFARE, PUB NO. (OS) 73-88, REPORT OF THE SECRETARY'S

by alleged threats of some insurance companies to withdraw from the medical malpractice liability market or to discontinue coverage of certain high-risk medical specialities.[2] The macroscopic impact of increasing medical malpractice insurance rates is obvious: to the extent that physicians relocate to areas with lower insurance rates or avoid high-risk specialities, the public is deprived of critical medical treatment; furthermore, as physicians pay higher rates, the added cost will be passed on to their patients.

Numerous commentators have argued that the purported "crisis" was grossly exaggerated[3] and may have been attributable to the insurance companies' need for excess profits to cover their stock market losses.[4] Regardless of whether the underlying circumstances were truly severe, the perceptions of a panicked public as well as ferocious lobbying by the medical profession and insurance industry[5] generated intense pressure on state legislatures to enact remedial legislation.

B. *Suggested Causes of the "Crisis"*

The list of suggested causes of the medical malpractice insurance crisis is endless, but certain explanations have received prominent attention. First, the increasing number of medical

---

COMMISSION ON MEDICAL MALPRACTICE 13, 38–40 (1973) (hereinafter cited as HEW REPORT) noting that in 1962, a physician paid ½ percent of his gross income for professional liability insurance. By 1970, this figure had more than tripled to 1.8 percent.

2 On the national level, there had once been approximately 85 insurance carriers writing malpractice insurance. By 1975, that number declined to 5 insurance carriers. Oregon Medical Ass'n v. Rawls, No. 421-496, *slip op.* at 1 (Ore. Cir. Ct. May 4, 1976), *rev'd on other grounds,* 276 Or. 1101, 557 P.2d 664 (1976) (remanded for further factual findings), cited in Redish, *Legislative Response to the Medical Malpractice Insurance Crisis: Constitutional Implications,* 55 TEX. L. REV. 759, 759 n.4 (1977); Note, *California's Medical Injury Compensation Reform Act,* 52 S. CAL. L. REV. 829, 847–48, n.111 (1979); *but cf.* HEW REPORT, *supra* note 1, at 38–39 (finding medical malpractice insurance generally available to physicians under group plans.

3 D. LOUISELL & H. WILLIAMS, MEDICAL MALPRACTICE § 20.07, n.55 (Supp. 1979) (hereinafter cited as MEDICAL MALPRACTICE).

4 Aitken, *Medical Malpractice: The Alleged 'Crisis' in Perspective,* 637 INS. L.J. 90, 96 (1976); Oster, *Medical Malpractice Insurance,* 45 INS. COUNSEL J. 228, 231 (1978).

5 *See, e.g.,* T. LOMBARDI, JR., MEDICAL MALPRACTICE INSURANCE, *Foreword* at xi, 92–93 (1978).

malpractice claims filed is obviously important.[6] Commentators have described several factors which may be pertinent to such increases: the breakdown in patient trust and admiration for the physician;[7] the general increase in public litigiousness as citizens have become more aware of their legal rights;[8] media coverage of high jury awards in medical malpractice cases;[9] and lawyers' contingency fee arrangements.[10]

Second, insurance companies have argued that erratic and exceedingly high damage awards have made accurate rate prediction impossible.[11] The added impact of the "long-tail" problem[12] in most insurance coverage schemes has forced many conservative insurance companies to impose artificially high present premiums to protect against even higher future damage awards.

Third, several critics of the purported crisis have alleged that insurance companies have derived unwarranted profits from the imposition of unduly high premiums[13] necessitated, at least in part, by poor investment practice.[14]

Fourth, there seems to be an emerging consensus among physicians, lawyers, insurance providers, and lawmakers that

---

6 Redish, *Legislative Response to the Medical Malpractice Insurance Crisis: Constitutional Implications*, 55 TEX. L. REV. 759, 760-61 (1977).

7 *See* MEDICAL MALPRACTICE, *supra* note 3, § 5.02, at 137-38 (1977); Peterson, *Consumers' Knowledge Of and Attitudes Toward Medical Malpractice*, in HEW REPORT, *supra* note 1, at 655 app.

8 *See* MEDICAL MALPRACTICE, note 3 *supra*.

9 *See* HEW REPORT, *supra* note 1, at 6-12.

10 *Id.* at 32. Public hearings demonstrated that "many doctors are convinced that the contingent fee system is at the very root of today's malpractice problem and any number of them have proposed its outright abolition as the most effective way to solve the [malpractice] problem". *Id.* In contrast, the Commission study did not support this belief.

11 *Id.* at 10.

12 The "long-tail" effect refers to the protracted time period in which the insurer remains vulnerable to claimants who may discover their injury, and therefore is unable to accurately assess his claim experience for a particular time frame. To some extent, this infirmity can be alleviated by shifting from "occurrence" to "claims-made" policy coverage. In the "occurrence" policy, the insurance company must cover all malpractice acts committed in a given time period regardless of when the claim is made. In the "claims-made" policy, the insurer covers all claims made in a given period regardless of when the malpractice act transpired. Comment, *The 'Claims-Made' Dilemma in Professional Liability Insurance*, 22 U.C.L.A. L. REV. 925, 926 (1975).

13 *See* MEDICAL MALPRACTICE, *supra* note 3, § 20.07, at n.56; Koskoff, *Physician Insure Thyself*, TRIAL, Dec., 1979, at 4.

14 *See* note 4 *supra*.

the high incidence of medical malpractice occurring throughout the county is attributable to some physicians' poor intake diagnostic practices[15] and to increased risks of error accompanying the use of complex medical technology.[16]

The cause of the medical malpractice insurance crisis is multifaceted, and the absence of an accurate, comprehensive, and reliable statistical information base has impaired the passage of carefully sculpted legislation.[17] Despite data limitations, however, political pressure prompted swift passage of legislation intended to alleviate the crisis.[18] In some cases, these measures were directed at the medical profession by stiffening licensing controls,[19] or at the insurance industry by authorizing joint underwriting associations and physician-owned mutual insurance associations.[20] The predominant legislative response, however, has been the enactment of substantive reforms in the common-law tort system of patient-injury compensation. These changes include: (1) limiting the amount of the plaintiff's recovery or the health care provider's liability; (2) creating medical malpractice screening panels; (3) shortening the applicable statute of limitations; (4) abrogating the collateral source rule; and (5) establishing various arbitration mechanisms. This Note examines damage recovery limitations and medical malpractice screening panels as the legislative modifications most denigrative of individual liberties.[21]

15 Zachey, *Diagnosing Medical Negligence: Three Recent Claims Studies*, TRIAL, July, 1979, at 12 (citing recent HEW, Nat'l Ass'n of Ins. Comm'rs, and St. Paul Fire and Marine Ins. Co. studies showing that over two-thirds of medical malpractice claims were derived from diagnostic errors resulting from inadequate physical examinations and taking of case histories).

16 T. LOMBARDI, *supra* note 5, at 24.

17 *See* HEW REPORT, *supra* note 1, at 45, 118.

18 *See, e.g.*, Wright v. Central Du Page Hosp. Ass'n, 63 Ill. 2d 313, 321, 347 N.E.2d 736, 739 (1976). The Illinois medical malpractice legislative scheme was later found to violate a state constitutional prohibition against "special legislation" as it unduly favored the medical profession. *See* text accompanying notes 69 to 77 *infra*.

19 Grossman, *The Medical Malpractice Crisis: State Legislative Activities in 1975*, *reprinted in Hearing on Examination of the Continuing Medical Malpractice Insurance Crisis Before the Subcomm. on Health of the Senate Comm. on Labor and Public Welfare*, 94th Cong., 1st Sess. 22 (1975) (hereinafter cited as *Hearing*).

20 *E.g.*, CAL INS. CODE §§ 11890–11916 (Supp. 1979); Grossman, *supra* note 19, at 21–22.

21 For discussions of the constitutional permissibility of other legislative modifications of common-law medical malpractice tort remedies, *see* Note, *California's Medical Injury Compensation Reform Act: An Equal Protection Challenge*, 52 S. CAL. L. REV. 829, 942–69 (1979); Witherspoon, *Constitutionality of the Texas Statute Limiting Li-*

* * *

While no state legislature has explicitly abrogated a medical malpractice victim's common-law tort remedy *in toto,* the cumulative impact of the legislative modifications may well amount to functional abrogation of such a remedy. To the extent that the dollar amount necessary to fully compensate a seriously-injured medical malpractice victim exceeds the statutory recovery limitation, or the unfortunate victim reasonably fails to discover his injury until after the abbreviated statute of limitations has run, or the excessive delay and additional legal costs of a mandatory medical malpractice screening panel effectively preclude the assertion of the victim's meritorious claim, the common-law tort remedy has been effectively abrogated without a corresponding benefit provided to the victim.

These efforts to alleviate the medical malpractice insurance crisis by revamping the tort compensation scheme have placed the brunt of such reform on those least able to bear its burdens — future medical malpractice victims. Most state courts have upheld medical malpractice legislative schemes that unilaterally strip individuals of their protective tort remedies. In sustaining these legislative programs against federal due process and equal protection challenges, the courts have stressed the rational relationship between the legislative means and the social goal of lowering insurance premiums and medical treatment costs. Unable to locate traditional "suspect classes" or "fundamental rights" implicated by the legislative schemes, the courts have been without adequate jurisprudential tools to fulfill their most important institutional function: safeguarding enduring individual liberties against precipitous majoritarian usurpation. This Note seeks to develop a principled technique of judicial review which would effectively protect the rights of medical malpractice victims against legislative subjugation.

Section I of this Note briefly describes medical malpractice recovery limitations and screening panels. Section II summarily discusses the state court decisions reviewing medical malpractice compensation schemes, and some of the states' recent experiences with the screening panels. Section III examines the

---

ability for Medical Malpractice, 10 TEX. TECH L. REV. 419 (1979); Redish, *supra* note 2, at 784–800.

theoretical social justifications of compensation and deterrence underlying the common law tort remedy and explores the constitutional due process requirement of a *quid pro quo*. This concept, which has emerged from the early workman's compensation cases and the recent nuclear power accident compensation case, specifies that a *quid pro quo* (i.e., "reasonably just substitute")[22] be provided to the individual in return for the state's abrogation of his common-law remedy. Section IV argues that statutory medical malpractice compensation schemes are directed at a powerless, "semi-suspect" class and implicate constitutionally-protected "quasi-fundamental" rights. Section V examines the weakness of present "intermediate scrutiny" judicial review techniques with respect to protecting threatened individual liberties, and describes the jurisprudential attributes of the *quid pro quo* standard. Finally, in Section VI, the *quid pro quo* standard is applied to existing medical malpractice compensation schemes to determine whether a statutory "reasonably just substitute" is indeed provided to protect individual liberties.

I. State Legislative Responses

Faced with the perceived medical malpractice insurance crisis, all fifty states enacted some form of responsive legislation. In most cases, this legislation was an integrated scheme embracing multiple modifications of a plaintiff's tort remedy.[23] The statutory premise was that a reduction in the number and dollar amount of malpractice awards would enable the insurance industry to predict recoveries more accurately and, therefore, to reduce premiums to physicians. Not only would stable insurance rates assure the availability of practicing physicians in the state, but the insurance savings, in theory, would be passed on to consumers in the form of lower health care costs. Unfortunately, post-reform experience has been contrary to these expectations. Despite the widespread legislative enactments, insurance premiums have continued to increase.[24]

---

22 Throughout this note the terms "*quid pro quo*" and "reasonably just substitute" will be used interchangeably.

23 *See, e.g.*, FLA. STAT. § 768 (Supp. 1979).

24 Redish, *supra* note 2, at 761-62; Chapman, *Is Another Crisis Looming?*, Nat'l L.J., Feb. 4, 1980, at 34, col. 1.

A. *Legislative Limits on Recovery and Liability*

Seventeen states[25] have limited the dollar amount recoverable by medical malpractice victims.[26] The applicable ceilings on recovery range from $150,000[27] to $750,000.[28] While these amounts appear to be wholly inadequate to fully compensate a grievously-injured medical malpractice victim, the vast majority of claims fall well below these ceilings.[29] The ultimate effect is that the recovery ceiling is unlikely to have much systemic impact in reducing insurance premiums and health costs, but is likely to impose a severe hardship on the most seriously-injured plaintiffs — those least able to bear the burden.[30]

A few states have established "Patients' Compensation Funds" to limit an individual physician's potential liability.[31] In return for obtaining a specified level of malpractice insurance and paying a surcharge into the fund, the physician is protected against very high damage awards. The legislation spreads the cost of the damage award among many health care providers and purports to ensure that the victim will receive adequate compensation. The value of this benefit to the victim is spurious. At least one court has observed that there was no legislative finding or evidence to indicate that there actually have been unsatisfied judgments.[32]

---

25 American Insurance Association, Summary of Medical Legislation, Oct. 15, 1977 (on file with the author).

26 Recoveries are further limited in states which have restricted the collateral source rule. The effect is to weaken the deterrent force derived from a tort damage award, and to penalize those plaintiffs with sufficient foresight to purchase health insurance coverage.

27 IDAHO CODE §§ 39–4204, -4205 (1977).

28 VA. CODE § 8.01–581.15 (1977); *see also* CAL CIV. CODE § 3333.2(b) (1975) (limiting only pain and suffering and other non-economic losses to $250,000).

29 In Jones v. State Bd. of Medicine, 97 Idaho 859, 555 P.2d 399 (1976), *cert. denied,* 431 U.S. 914 (1977), the Idaho Supreme Court took notice of an affidavit submitted by the Director of the State Department of Insurance that the $150,000/$300,000 limit on medical malpractice recoveries "would have covered all claims to date." *Id.* at 413.

While the Virginia statute imposes a recovery ceiling of $750,000, "[f]rom 1970 to 1975 no claim for medical malpractice in Virginia was reported in excess of $500,000 and only one claim was settled for $250,000." NOTE, *Alternatives to the Medical Malpractice Phenomenon: Damage Limitations, Malpractice Review Panels, and Countersuits,* 34 WASH. LEE L. REV. 1179, 1185 n.43 (1977).

30 Some state constitutions contain explicit provisions prohibiting limitations on the amount of recoverable damages in personal injury cases. *E.g.,* KY. CONST. § 54.

31 *E.g.,* IND. CODE ANN. § 16–9.5–4 (1976).

32 McCoy v. Commonwealth, 37 Pa. Commw. Ct. 530, 551, 391 A.2d 723, 733 (1978) (Mencer, J., dissenting) (quoting McGuffey v. Hall, 557 S.W.2d 401, 412 (Ky. 1977)).

B. *Medical Malpractice Screening Panels*

At least twenty-five states have statutorily created pre-trial medical malpractice screening panels.[33] In most of these states, the parties are required to submit their claims to the panels before filing suit. The panels typically include three to seven members and are comprised of physicians, lawyers, judges, and laypersons. The proceedings are adversarial in nature, but the evidentiary rules are often less strict than in formal civil trials.[34]

Although the panel decisions are not binding,[35] they are intended to encourage the settlement of meritorious claims without the added costs of a formal trial.[36] Furthermore, the screening panels are expected to discourage the pursuit of frivolous claims or those which, in the opinion of the panel, present insufficient evidence to support the charge of medical malpractice.

In requiring that all claims be submitted to medical malpractice screening panels, states are imposing an added layer of delay and legal costs upon the parties. In most cases, this prelitigation burden will fall more heavily on the injured plaintiff than on the well-heeled insurance company and physician defendants. To the extent that the panels impose long delays or high costs, the litigation of meritorious claims may effectively be deterred.[37]

Most states permit the panel's findings to be admitted at trial. This has raised substantial concern that the expert panel's decision might unduly influence the lay jury's independent eval-

---

33 These panels are conceptually derived from earlier private plans adopted by some local medical societies. While plaintiffs could not be compelled to submit their claims to these private panels, they were encouraged by the promise that an expert medical witness would be provided at trial if the panel found the plaintiff's claim to be meritorious. This benefit was intended to circumvent the physicians' "conspiracy of silence." In exchange, it was contemplated that a plaintiff with a claim deemed to be non-meritorious would abandon his claim, either voluntarily or by prior agreement. *See* C. Baird, G. Munsterman & J. Stevens, *Alternatives to Litigation I, reprinted in* HEW REPORT, *supra* note 1, app., at 214, 224–25.

34 *E.g.*, the Florida statute states that "strict adherence to the rules of procedure and evidence applicable in civil cases shall not be required." FLA. STAT. ANN. § 768.44(b) (1977).

35 Maryland is the only state in which the panel's findings are binding and converted into judgments. MD. CTS. & JUD. PROC. CODE ANN. §§ 3–2A01, -2A06 (1980).

36 In addition to lowering litigation expenses for the plaintiff, out-of-court settlements are desired because they reduce the adverse publicity that may damage the physician's reputation.

37 *See* notes 340 to 343 and accompanying text *infra*.

uation of the evidence.[38] The admissibility question bears an ironic relation to the statutory purpose of the panel: if the panel's findings are subsequently admitted at trial, the panel loses its mediative role and becomes merely an additional adversarial forum; on the other hand, if the findings are not admissible, then there is little deterrence to either party's proceeding to trial, and the panel is transformed into a costly rehearsal.

Since the added burden of the medical malpractice screening panel falls most heavily on an injured plaintiff's ability to assert his claim, constitutional objections will often focus on the efficiency, cost, and fairness of the particular state's screening panel process.[39]

II. State Court Review of Medical Malpractice Compensation Schemes

A. *Federal Due Process and Equal Protection Review Standards*

With the repudiation of substantive due process in the 1930's,[40] equal protection and due process analyses devolved into the now traditional two-tiered structure.[41] Courts generally evaluate legislation using either the "strict scrutiny" standard or the "minimum rationality" test.[42]

---

38 *See* Note, *Malpractice Panels*, 46 FORDHAM L. REV. 322, 334 (Nov. 1977); *but see* DiAntonio v. Northampton Accomack Mem. Hosp., 628 F.2d 287 (4th Cir. 1980) (upholding medical malpractice panel opinion in order to promote meaningful mediation and settlement); Prendergast v. Nelson, 199 Neb. 97, 109, 256 N.W.2d 657, 666 (1977) (jury can evaluate panel's findings with objectivity).

39 Medical malpractice screening panels have enjoyed vastly divergent levels of success and failure in different states. *See* Margolick, *Mediation Isn't Cure for Patients' Claims*, Nat'l L.J., Feb. 4, 1980, at 2, col. 1.

40 West Coast Hotel Co. v. Parrish, 300 U.S. 379 (1937).

41 *See* San Antonio Independent School Dist. v. Rodriquez, 411 U.S. 1 (1973); United States v. Carolene Products Co., 304 U.S. 144, 152 n.4 (1938).

42 This Note will later discuss the "reasonably just substitute" test as an element of constitutional due process, rather than of equal protection, because that has been its historical association. New York Central R.R. Co. v. White, 243 U.S. 188 (1917). Furthermore, this Note will emphasize concerns of safeguarding individual liberty interests when reviewing medical malpractice legislation.

> 'Due process' emphasizes fairness between the State and the individual dealing with the State, regardless of how other individuals in the same situation may

Where a statute threatens a "fundamental right" or discriminates against a "suspect class," it is subject to strict scrutiny and is upheld only if it furthers a "compelling state interest." As has been aptly noted by many commentators, application of strict scrutiny is almost always "'strict' in theory and fatal in fact."[43] Though hardly self-defining, fundamental rights status has been extended by the Supreme Court to those individual liberty[44] and political process rights[45] "explicitly or implicitly guaranteed by the Constitution."[46] Classifications are treated as suspect where the "class is . . . saddled with such disabilities, or subjected to such a history of purposeful unequal treatment, or relegated to such a position of political powerlessness as to command extraordinary protection from the majoritarian political process."[47]

Most other classifications traditionally have been subjected to the minimum rationality test, which seeks only to establish that the legislature's chosen means are rationally related to its acknowledged ends.[48] In the context of economic regulation and social welfare legislation not implicating fundamental rights or a suspect class, this standard has resulted in extreme judicial deference to the legislature's majoritarian political process:

> be treated. 'Equal protection,' on the other hand, emphasizes disparity in treatment by a State between classes of individuals whose situations are arguably indistinguishable.

Ross v. Moffitt, 417 U.S. 600, 609 (1974). This *dicta* appears in opposition to the often repeated maxim presented in Mr. Justice Jackson's concurring opinion in Railway Express Agency, Inc. v. New York, 336 U.S. 106, 112 (1949):

> Invalidation of a statute or an ordinance on due process grounds leaves ungoverned and ungovernable conduct which many people find objectionable. Invocation of the equal protection clause, on the other hand, does not disable any governmental body from dealing with the subject at hand.

However, *dicta* set forth in the Supreme Court's recent decision in Duke Power v. Carolina Environmental Group, Inc. 438 U.S. 59, 93 (1978), indicates a merger of due process and equal protection review.

43 Gunther, *The Supreme Court, 1971 Term — Forward: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection*, 86 HARV. L. REV. 1, 8 (1972).

44 *See, e.g.*, Shapiro v. Thompson, 394 U.S. 618 (1969); Skinner v. Oklahoma, 316 U.S. 535 (1942).

45 *See, e.g.*, Reynolds v. Sims, 377 U.S. 533 (1964).

46 San Antonio Independent School Dist. v. Rodriguez, 411 U.S. 1, 33–34 (1973).

47 *Id.* at 28.

48 *See, e.g.*, Dandridge v. Williams, 397 U.S. 471 (1970); McGowan v. Maryland, 366 U.S. 420 (1961).

> The constitutional safeguard (of equal protection) is offended only if the classification rests on grounds wholly irrelevant to the achievement of the State's objective. State legislatures are presumed to have acted within their constitutional power despite the fact that, in practice, their laws result in some inequality. A statutory discrimination will not be set aside if any state of facts reasonably may be conceived to justify it.[49]

In recent years, the Supreme Court has begun to diverge from the rigid stratification of this "all or nothing" two-tiered analysis. In apparent reaction to the extreme judicial abdication inherent in the rubber-stamp minimum rationality standard of review, the Court has moved occasionally to a more flexible form of "intermediate scrutiny."[50] While the precise contours of this review standard have remained ill-defined, some general trends can be perceived.

Initially, commentators described the new test as a "means-oriented scrutiny."[51] The Court would no longer abstractly hypothesize a conceivable legislative purpose to justify the propounded ends,[52] but rather would require that a "substantial relationship" be established between the means and ends of the challenged legislation.[53] Several lower federal and state courts subsequently relied on this enhanced means-oriented analysis to more carefully scrutinize legislation.[54] The understated assumption of this formulation was that the court would examine the substantiality of the asserted means toward promoting the legislative ends, yet would not intervene directly to scrutinize the ends themselves.

It soon became evident that the Supreme Court's focus would extend to the legitimacy of legislative ends.[55] Faced with chang-

---

49 McGowan v. Maryland, 366 U.S. at 425–26.

50 L. TRIBE, AMERICAN CONSTITUTIONAL LAW, at 1057–97 (1978); Gunther, *supra* note 43, at 20–24. For a discussion of those circumstances triggering intermediate scrutiny, *see* notes 55 to 58 and accompanying text *infra*.

51 Gunther, *supra* note 43, at 20–24.

52 *See, e.g.*, Kotch v. Bd. of River Port Pilot Comm'rs, 330 U.S. 552 (1947).

53 Gunther, *supra* note 43; *see* Reed v. Reed, 404 U.S. 71 (1971).

54 *See, e.g.*, Green v. Waterford Bd. of Educ., 473 F.2d 629 (2d Cir. 1973); Jones v. State Bd. of Medicine, 97 Idaho 859, 555 P.2d 399 (1976), *cert. denied*, 431 U.S. 914 (1977).

55 *See, e.g.*, Craig v. Boren, 429 U.S. 190 (1976); Cleveland Bd. of Educ. v. LaFleur,

ing social norms and constitutional challenges to legislation predicated on arguably antiquated social mores,[56] the Court grappled to formulate a principled standard of review. While lower courts appeared reluctant to expand judicial intrusion into the policy judgments of democratically representative legislatures by enlarging the range of statutes subject to strict scrutiny, the extreme deference of minimum rationality review seemed equally untenable.

The Court's qualified intervention, therefore, emerged in "twilight zone" cases — where significant burdens were imposed on "quasi-fundamental" rights or "semi-suspect" classes. These quasi-fundamental rights have included basic human liberties and those goods comprising the "necessities of life."[57] Enhanced judicial review of semi-suspect classifications[58] recognizes that historic stereotyping and prejudice, though perhaps not so severe as that confronted by racial minorities, provokes a set of fixed assumptions in shaping legislation that has unduly penalized certain groups, and that the contemporary American political process resembles more a pluralist compromise of forcefully asserted interest-group demands,[59] than the ancient ideal of the democratic polis.[60] For a variety of sociological and political reasons, certain individuals sharing mutual interests will inevitably fail to coalesce and, therefore, will be continually subjugated in the political bartering process. Where this political impotence has been readily apparent, courts have displayed special solicitude in remedying burdens cast on unorganized groups by a majoritarian legislative or administrative process from which those groups have been effectively excluded.[61]

---

414 U.S. 632 (1974); United States Dep't of Agriculture v. Moreno, 413 U.S. 528 (1973); Eisenstadt v. Baird, 405 U.S. 438 (1972).

56 *See, e.g.*, Cleveland Bd. of Educ. v. LaFleur, 414 U.S. 632, 641 n.9 (1974); TRIBE, *supra* note 50, § 16-24, at 1060-61, § 16-30, at 1084-85.

57 *See, e.g.*, Memorial Hosp. v. Maricopa County, 415 U.S. 250 (1974) (non-emergency medical care); Turner v. Department of Employment Security of Utah, 423 U.S. 44 (1975) (unemployment compensation); United States Dep't of Agriculture v. Moreno, 413 U.S. 528 (1973) (food stamps).

58 *See, e.g.*, Trimble v. Gordon, 430 U.S. 762 (1977) (illegitimates); Hampton v. Mow Sun Wong, 426 U.S. 88 (1976) (aliens); Weinberger v. Wiesenfeld, 420 U.S. 636 (1975) (women); Stanley v. Illinois, 405 U.S. 645 (1972) (unwed fathers).

59 *See, e.g.*, R. DAHL, PLURALIST DEMOCRACY IN THE UNITED STATES (1967).

60 *See, e.g.*, ARISTOTLE, THE POLITICS, BOOK III at 97-107, BOOK VII at 279-89 (Oxford ed., E. Barker ed. and trans. 1974).

61 *See* note 58 *supra*.

Later in this Note, it will be argued that the individual's interest in bodily integrity and personal security from unconsented assault is a quasi-fundamental right abridged by certain medical malpractice legislative schemes.[62] Furthermore, it will be suggested that the class of medical malpractice victims shares traits of political powerlessness similar to those of recognized semi-suspect classes.[63] As a result of these infirmities, whether considered separately or in combination, this Note concludes that courts are warranted to engage in a more careful scrutiny of the social and economic justifications for medical malpractice legislation,[64] and should intervene on behalf of victimized individuals when statutory compensation schemes sacrifice their basic liberties to a purported collective welfare.[65]

B. *State Court Review of Medical Malpractice Legislation*

This section will address the general trends in state court adjudication of medical malpractice compensation schemes as background for this Note's broader constitutional analysis. Numerous other articles have examined the various decisions in detail.[66]

1. The First Round

The early state court decisions[67] displayed skepticism towards medical malpractice compensation schemes that harshly affected individual liberties. Plaintiffs' sweeping attacks generally alleged that federal due process and equal protection rights had been impaired by the statutes classifying medical malpractice

---

62 *See* notes 272 to 297 and accompanying text *infra.*

63 *See* notes 249 to 271 and accompanying text *infra.*

64 *See* notes 298 to 328 and accompanying text *infra.*

65 *See* notes 329 to 349 and accompanying text *infra.*

66 *E.g.*, Comment, *Constitutional Challenges to Medical Malpractice Review Boards*, 46 TENN. L. REV. 607, 607-48 (1979); Note, *supra* note 2, at 868-91.

67 This discussion will examine judicial review of medical malpractice victims' claims that the legislative schemes are unconstitutional. This Note will not discuss physicians' constitutional challenges to the legislation mandating compulsory medical malpractice insurance as a condition of licensing. *E.g.*, McCoy v. Commonwealth, 37 Pa. Commw. Ct. 530, 391 A.2d 723 (1978).

victims as distinct from other tort victims, and that the intrusive means employed by the legislative schemes were insufficiently related to the purported ends. Furthermore, plaintiffs asserted violations of state constitutional provisions that protected rights of access to the courts[68] and to jury trials, as well as those which prohibited "special legislation" and the usurpation of judicial functions.

Essentially, these courts invoked federal constitutional grounds to analyze both the purported basis for the scheme and its structural impact on the individual. The Illinois Supreme Court was the first to invalidate medical malpractice legislation. In *Wright v. Central Du Page Hospital Association*,[69] the court initially determined that medical malpractice screening panels violated state constitutional provisions which vested all judicial power in the courts[70] and which protected the right to a jury trial.[71] The court then examined the constitutional integrity of the statutory classification which established a $500,000 recovery limitation solely for medical malpractice tort victims.

The Illinois court's equal protection analysis was oblique. Defendants cited the state's Workmen's Compensation Act[72] as precedent for the limitation on recoveries. That act established a *quid pro quo* insofar as the employer was obligated to waive certain defenses in exchange for the limitation on the employee's common-law tort action. The court found the analogy inappropriate since the malpractice legislation did not extract a comparable *quid pro quo* from the medical profession.[73] The court also explicitly rejected defendant's "societal" *quid pro quo* argument — "that the loss of recovery potential to some malpractice victims is offset by 'lower insurance premiums and

68 For a short time, it appeared that the Supreme Court might establish a Fourteenth Amendment right of access to the courts. Boddie v. Connecticut, 401 U.S. 371 (1971) (state's imposition of court fees denies access to indigents seeking divorce and, therefore, violates due process). This holding, however, was soon drastically restricted to cases where a fundamental right was at stake. Ortwein v. Schwab, 410 U.S. 656 (1973) (state can constitutionally restrict welfare recipient's court access by requiring filing fee for appeal from adverse welfare decision); United States v. Kras, 409 U.S. 434 (1973) (required payment of filing fee prior to discharge in bankruptcy was valid).

69 63 Ill. 2d 313, 347 N.E.2d 736 (1976).

70 63 Ill. 2d at 322, 347 N.E.2d at 739-40.

71 63 Ill. 2d at 322-24, 347 N.E.2d at 740-41.

72 *See* notes 163 to 168 and accompanying text *infra*.

73 63 Ill. 2d. at 328, 347 N.E.2d at 742.

lower medical care costs for all recipients of medical care'"[74] — because these benefits did not extend significantly to the seriously-injured malpractice victim.[75] After finding the legislative classification to be arbitrary, the court refused to determine conclusively whether a *quid pro quo* was essential, but invalidated the legislation because "to the extent that recovery is permitted or denied on an arbitrary basis a special privilege is granted in violation of the Illinois Constitution."[76] Apparently, the court either harshly applied the minimum rationality review standard or invoked an elevated basis of scrutiny to protect individual rights impaired by the medical malpractice legislation.[77]

Soon after *Wright* was decided, the Idaho Supreme Court examined its state's legislation in *Jones v. State Board of Medicine*.[78] The Idaho Hospital-Medical Liability Act limited health care provider liability to $150,000 per claim or $300,000 per occurrence, and abrogated the collateral source rule in medical malpractice cases. Plaintiff's due process and equal protection attacks alleged that the provision arbitrarily discriminated against malpractice victims entitled to awards exceeding the statutory limit. Since the appeal record was devoid of facts supporting the existence of a medical malpractice insurance "crisis" in Idaho,[79] or establishing a relation between claims recovery and Idaho health care costs,[80] the court remanded the case to the trial court for specific fact-finding.

Though it found no fundamental right or suspect classification present,[81] the Idaho Supreme Court's obvious suspicion of the actuality of a crisis,[82] and the state's claims recovery experience,[83] clearly shaped its review standard. Asserting that the legislative classification was "discriminatory . . . on its face"[84] the court determined that the "means-focused" intermediate scrutiny, as enunciated by the United States Supreme Court in

---

74 *Id.*
75 *Id.*
76 63 Ill. 2d at 329, 347 N.E.2d at 743.
77 *See* Note, *supra* note 2, at 872 n.258.
78 97 Idaho 859, 555 P.2d 399 (1976), *cert. denied*, 431 U.S. 914 (1977).
79 97 Idaho at 872, 876, 555 P.2d at 412, 416.
80 97 Idaho at 872, 555 P.2d at 412.
81 97 Idaho at 870, 555 P.2d at 410.
82 97 Idaho at 872, 876, 555 P.2d at 412, 416.
83 97 Idaho at 873, 555 P.2d at 413.
84 97 Idaho at 871, 555 P.2d at 411.

Reed v. Reed.[85] was applicable.[86] The texture of the court's opinion[87] makes quite clear that this elevated review standard is intended to protect medical malpractice victims by providing a sterner test of the legislative classification and recovery limitations than would the deferential minimum rationality standard.[88]

2. The Second Round

The Florida Supreme Court's tentative step in upholding the constitutionality of a medical malpractice screening panel in *Carter v. Sparkman*[89] opened the floodgates for an effusive flow of judicial deference. In that case, plaintiff's due process and equal protection claims focused on the burden on his access to the courts, as well as on a perceived procedural inequality that required plaintiff to submit to mediation, but seemingly would have permitted the defendant to refrain from participation. After creatively construing the statute to constrain the defendant by permitting plaintiff, at trial, to present the fact of the physician's non-participation in the administrative hearing,[90] the court turned its attention to plaintiff's substantive access claim.

The court first acknowledged its general predilection against burdens on access to the courts, but stressed the importance of the state's ability to invoke its police power to respond to the malpractice crisis.[91] Ultimately, the court concluded that while "the pre-litigation burden cast upon the claimant reaches the outer limits of constitutional tolerance,"[92] the statute would

85 404 U.S. 71 (1971).

86 The *Jones* court rejected a plaintiff hospital's contention that the *quid pro quo* doctrine required that it be provided with specific substitute benefits in return for certain statutory burdens. 97 Idaho at 866, 870-71, 555 P.2d at 406, 410-11.

87 97 Idaho at 866-67, 870-71, 555 P.2d at 406-07, 410-11.

88 In Simon v. St. Elizabeth Medical Center, 3 Ohio Op. 3d 164, 355 N.E.2d 903 (1976), and Graley v. Satayatham, 74 Ohio Op. 2d 316, 343 N.E.2d 832 (1976), the Ohio trial courts apparently applied strict scrutiny review to invalidate the Ohio Medical Malpractice Act which created mandatory arbitration, imposed a $200,000 limit on recovery, and abolished the collateral source rule. No higher court has overturned these decisions.

89 335 So. 2d 802 (Fla. 1976), *cert. denied*, 429 U.S. 1041 (1977).

90 335 So. 2d at 805.

91 *Id.* at 806.

92 *Id.*

survive. In a concurring opinion joined by three other justices, Justice England compared the screening panel to an acceptable "required pre-trial settlement conference,"[93] but nevertheless expressed deep concern with the harsh burden of the "expense of two full trials"[94] that was placed on plaintiff.[95]

Subsequent courts were far less reticent in upholding medical malpractice legislative schemes. Instead of focusing on the scheme's structural effect upon the individual, they adopted a much more atomized approach. Typically, the courts rejected plaintiff's due process and equal protection claims by finding neither fundamental rights nor a suspect class at stake, and then referred directly to the legislative findings to demonstrate rationality to a legitimate social end.[96]

The Wisconsin Supreme Court's decision in *State ex rel. Strykowski v. Wilkie*[97] is representative. The state statute required that claims be submitted to the medical malpractice screening panel, and imposed a recovery limitation of $500,000 in certain cases. After observing that neither fundamental rights nor a suspect class were implicated by the legislative scheme, the court concluded that only "a reasonable basis upon which the legislature might have acted" need be shown.[98] The court dismissed plaintiff's claim that the medical malpractice crisis was illusory[99] by citing the suppositions set forth in the statute's preamble[100] as propounding a rational basis for the screening panel. Furthermore, the court applied an equal protection classification analysis derived from an earlier Wisconsin case to

93 *Id.* at 807.

94 *Id.*

95 History would prove the added cost and delay to be sufficiently burdensome so that four years later, the Florida Supreme Court overruled *Carter* and invalidated the screening panels as violative of due process and rights of access to the courts. Aldana v. Holub, 381 So. 2d 231 (Fla. 1980).

96 *See* cases cited at notes 102 to 110 *infra*. It is interesting to note that many of the courts upholding these statutes have prominently cited Professor Redish's article in support of their reasoning. That article, which takes a dim view of both federal and state constitutional challenges to these statutes, "*was prepared for and funded by the American Hospital Association,* but the conclusions are solely those of the author." *See* Redish, *supra* note 2, at 759 n.* (emphasis supplied).

97 81 Wis. 2d 491, 261 N.W.2d 434 (1978).

98 81 Wis. 2d at 507, 261 N.W.2d at 442.

99 *Id.*

100 *Id.*

demonstrate that the statute's treatment of medical malpractice actions as "substantially distinct" from other tort actions was reasonable.[101] The *Strykowski* rationale was essentially repeated in similar federal and state constitutional adjudications of medical malpractice legislative schemes in Arizona,[102] Florida,[103] Indiana,[104] Louisiana,[105] Nebraska,[106] Maryland,[107] Massachusetts,[108] New York,[109] and Pennsylvania.[110]

Since these courts directed the two-tier due process and equal protection inquiries solely towards ascertaining the general rationality of the legislative schemes for achieving the intended broad public goals, the courts were forced to turn to state constitutional provisions to assess the impairment of individual rights "sacrificed for the social good." Yet, these courts subsequently refused to invalidate any statutes on the basis of denial of access to the courts or the right to trial by jury.

It is noteworthy that no state supreme court has explicitly upheld a dollar limitation on medical malpractice damage recoveries. While the early state court decisions found these recovery limits suspect under constitutional standards,[111] this second set of state court decisions steadfastly side-stepped the issue amidst otherwise deferential approval.[112] This avoidance reflects a significant structural gap in the form of constitutional review applied to these legislative schemes.[113]

---

101 *Id.* at 442 n.8.

102 Eastin v. Broomfield, 116 Ariz. 576, 570 P.2d 744 (1977) (en banc).

103 Woods v. Holy Cross Hosp., 591 F.2d 1164 (5th Cir. 1979); Carter v. Sparkman, 335 So. 2d 802 (Fla. 1976), *cert. denied*, 429 U.S. 1041 (1977); *but see* Aldana v. Holub, 381 So. 2d 231 (Fla. 1980).

104 Hines v. Elkhart General Hosp., 465 F. Supp. 421 (N.D. Ind. 1979).

105 Seoane v. Ortho Pharmaceuticals, Inc., 472 F. Supp. 468 (E.D. La. 1979); Everett v. Goldman, 359 So. 2d 1256 (La. 1978).

106 Prendergast v. Nelson, 199 Neb. 97, 256 N.W.2d 657 (1977).

107 Attorney General v. Johnson, 282 Md. 274, 385 A.2d 57 (1978).

108 Paro v. Longwood Hospital, 373 Mass. 645, 369 N.E.2d 985 (1977).

109 Comiskey v. Arlen, 55 App. Div. 2d 304, 390 N.Y.S.2d 122 (1976).

110 Parker v. Children's Hosp. of Philadelphia, 483 Pa. 106, 394 A.2d 932 (1978).

111 *See, e.g.*, Wright v. Central Du Page Hosp. Ass'n, 63 Ill. 2d 313, 329, 347 N.E.2d 736, 743 (1976).

112 *See, e.g.*, Strykowski v. Willkie, 81 Wis. 2d 491, 511, 261 N.W.2d 434, 443-44 (1978); *cf.* Prendergast v. Nelson, 199 Neb. 97, 116, 256 N.W.2d 657, 669 (1977) ($500,000 recovery limitation upheld only by three judge plurality; furthermore, the court strongly implied that plaintiff had the election of not pursuing his remedy under the act.).

113 *See* notes 344 to 349 and accompanying text *infra*.

3. The Recent Round of Critical Decisions

As the panicked days of the medical malpractice insurance "crisis" have become more distant, and the operation of screening panels, in some states, has demonstrated a history of inefficiency,[114] some courts have begun to examine the individual rights impaired by the social schemes more carefully.

Arneson v. Olson[115] was the only departure from the pattern of judicial approval in the "second round" of decisions.[116] The North Dakota statute[117] imposed a $300,000 recovery limitation, virtually eliminated the collateral source rule and the application of res ipsa loquitor in malpractice actions, and established a screening panel.[118]

Initially, the court announced that it would apply an intermediate scrutiny standard of review that would look for a "close relationship" between legislative means and ends.[119] Furthermore, the court stated "that while there need not always be a *quid pro quo,* any limitation or elimination of a preexisting right may not be arbitrarily imposed."[120] This confluence of review standards seemed designed to balance, rather than sacrifice, individual rights with the social scheme.

The court determined that some malpractice vicitms may have expenses well in excess of the $300,000 recovery limit. Therefore, it was necessary to closely examine whether the limit was arbitrary, thereby violating due process, or whether there was

---

114 *See* notes 130 to 145 and accompanying text *infra.*

115 270 N.W.2d 125 (N.D. 1978).

116 Since the statute required most physicians to obtain malpractice insurance, four physicians sued to have the Act declared invalid. In its review of their federal and state constitutional due process and equal protection challenges, the court broadened its scope to include the provisions of the act impairing the rights of medical malpractice victims.

117 A patient might choose not to be bound by the act, but then a physician could refuse to provide services. 270 N.W.2d at 127, 133. The court viewed this provision as hollow since:

> The only choices available to the patient who is refused care apparently are to suffer or die of his ailment or to travel outside the State to obtain medical attention. . . . [W]e will consider the apparent harshness . . . in determining whether due process and equal protection are violated.

Id. at 134.

118 The court avoided adjudicating the constitutionality of the screening panels since "[n]one of the parties before us challenges [them] . . ." *Id.* at 131.

119 *Id.* at 132-33.

120 *Id.* at 135.

a sufficiently "close correspondence between the statutory classification and legislative goals so as not to violate the equal protection requirements of the State and Federal Constitutions."[121] The court determined:

> Certainly the limitation of recovery does not provide adequate compensation to patients with meritorious claims; on the contrary, it does just the opposite for the most seriously injured claimants. It does nothing toward the elimination of nonmeritorious claims. Restrictions on recovery may encourage physicians to enter into practice and remain in practice, but do so only at the expense of claimants with meritorious claims.[122]

After affirming the trial court's finding that there was no insurance cost or availability crisis in the state,[123] the court invalidated the recovery limit on federal and state constitutional grounds. In functionally applying a *quid pro quo* form of analysis, the court concluded that "the cumulative effect [of the tort remedy restrictions] of the Act . . . violate[s] the right of medical patients in this State to due process of law."[124]

Soon thereafter, the Missouri Supreme Court held that its state's medical malpractice screening panel statute violated a state constitutional "fundamental right" of access to the courts, even though a panel's decision was not admissible at trial.[125]

A California appellate court declared unconstitutional a provision of its medical malpractice compensation scheme[126] that permitted damage awards in excess of $50,000 to be paid on a periodic basis.[127] The court determined that the legislative restriction that withheld the normal lump-sum payments was impermissibly underinclusive in singling out severely injured medical malpractice victims — as compared to the classes of the health care-consuming general public, all tort victims, and all medical malpractice victims — to bear the statutory bur-

---

121 *Id.*
122 *Id.* at 135-36.
123 *Id.* at 136.
124 *Id.*
125 Cardinal Glennon Mem. Hosp. for Children v. Gaertner, 583 S.W.2d 107 (Mo. 1979).
126 CAL. CIV. PROC. CODE § 667.7 (1980).
127 American Bank and Trust v. Community Hosp., 104 Cal. App. 3d. 219, 163 Cal. Rptr. 513 (1980).

den.[128] The court concluded that the legislature by "tax[ing] an impermissible special class for the purported benefits to be enjoyed by the general public" acted with sufficient "arbitrar[iness] as to deny equal protection".[129]

At about the same time that some health care and insurance analysts were suggesting that the crisis had passed,[130] the grossly inefficient operations of some states' medical malpractice screening panels had become apparent. High costs and excessive delays were presumably discouraging the litigation of meritorious claims.[131] For example, in Maryland, 315 of the 350 claims filed since July, 1976 were still open at the end of 1979, and only ten hearings had actually been held.[132] Similarly, the sorry record of Pennsylvania's mandatory arbitration panels is that of 2,466 claims filed in 3 1/2 years, only nine reached the necessary pre-trial hearing stage. "Less than ten percent of the claims have been settled, discontinued, and ended by conciliation conferences and 507 were dismissed or discontinued without conciliation. . . . It describes a system that, though theoretically sound, is actually a resounding flop."[133]

The most startling judicial shifts have been the Florida Supreme Court's reversal of *Carter v. Sparkman*[134] and the Pennsylvania Supreme Court's recent invalidation[135] of the state's medical malpractice compensation scheme after an earlier determination that it did not violate state constitutional guaranties.[136] In *Aldana v. Holub,*[137] the defendant physician was deprived of a screening panel hearing when, through no fault of his own, Florida's strict ten-month statutory time period

---

128 *Id.* at 518-19, 521.

129 *Id.* at 521.

130 "Donald Clifford, vice president for underwriting of St. Paul Fire and Marine Insurance Co., the nation's largest malpractice insurer, says 'the malpractice crisis has been over for a year.' " Margolick, *supra* note 39, at 35. To some extent, the rise in rates has been checked by the increasing number of physician self-insurance plans.

131 The Justice Department has reportedly dropped a plan to standardize pre-trial screening panels in all 50 states. *Id.* at 34.

132 *Id.*

133 Edelson v. Soricelli, 610 F.2d 131, 135-36 (3rd Cir. 1979), *citing,* [1979] ANN. REP. ADMIN. FOR ARB. PANELS FOR HEALTH CARE, 17 App. (Statistics through Aug. 31, 1979).

134 335 So. 2d 802 (Fla. 1976), *cert. denied,* 429 U.S. 1041 (1977).

135 Mattos v. Thompson, 49 U.S.L.W. 2249 (Pa. Sup. Ct., Sept. 22, 1980).

136 Parker v. Children's Hosp. of Philadelphia, 483 Pa. 106, 394 A.2d 932 (1978).

137 Aldana v. Holub, 381 So. 2d. 231, 234 (Fla. 1980).

elapsed before adjudication. Since the judicial referee had acted unilaterally to set the hearing date after the statutory time period, defendant argued that he had been deprived of a legal right without due process of law.[138] Plaintiff challenged the entire medical malpractice act as denying due process and equal protection[139] — grounds that previously had been rejected in *Carter*.

The court readily determined that the strict time limitation was jurisdictional, and found that extensions were neither authorized by the legislature, nor permissible in light of state constitutional guarantees of "speedy access to the courts."[140] Since the rigidity of the statute irrevocably deprived defendant of a right to mediation solely because of the judge's crowded docket, the court found it arbitrary and violative of due process.[141]

In its decision, the court discussed its examination of more than seventy similar cases in which "over fifty percent of the time, a valuable legal right [had] arbitrarily evaporated through no fault of either party."[142] The court determined that the "medical mediation statute has proven unworkable and inequitable in practical operation."[143]

In *Carter*, the court rejected plaintiff's argument, but had recognized that the serious burden of expense and delay from the screening panel limited the plaintiff malpractice victim's right of access to the courts.[144] In *Aldana*, the court's dilemma was apparent: to the extent that it construed an extended statutory time period to preserve its constitutionality, "[t]o now increase the prelitigation burden cast upon the claimant by permitting continuances and exceptions . . . would transcend those outer limits of constitutional tolerance. What was originally contemplated as an inexpensive summary procedure would now extend to twelve, fourteen, or possibly even sixteen months or

---

138 *Id.*
139 *Id.* at 234-35.
140 *Id.* at 235.
141 *Id.* at 236.
142 *Id.* at 237.
143 *Id.*
144 "Although the Court rejected [plaintiff's] argument, we explicitly recognized that 'the pre-litigation burden cast upon the claimant reaches the outer limits of constitutional tolerance." *Id.* at 238, *citing* Carter v. Sparkman, 335 So. 2d 802, 805 (Fla. 1976), *cert. denied*, 429 U.S. 1041 (1977).

more, thereby effectively depriving one's access to the courts."[145] Faced by this intractable deficiency, the court declared that the entire medical mediation act violated the due process clauses of the United States and Florida Constitutions.[146]

Essentially, the court decided that the performance of the screening panels was so erratic and arbitrary as to warrant invalidation under the minimum rationality strand of due process analysis. The court did not relate the panel's purpose to a legislative goal, but apparently concluded that the attendant delay and cost so impaired individual liberties that it presumptively violated due process considerations of fairness.

In *Mattos v. Thompson*,[147] the Pennsylvania Supreme Court recently reassessed its prior theoretical approval[148] of the state's Health Care Services Malpractice Act. The court held that the lengthy delays engendered by the arbitration system did in fact burden the right to a jury trial with "onerous conditions, restrictions, or regulations which . . . make the right practically unavailable."[149]

The court extensively evaluated the poor statistical record of the arbitration panels in providing prompt adjudication of medical malpractice claims, and determined that the history of delays was unconscionable and irreparably damaging to public confidence in the efficiency and effectiveness of the judicial system. The court concluded that the legislative scheme, which gave the panels "original exclusive jurisdiction," was incapable of achieving its stated purpose and impermissibly burdened plaintiff's right to a jury trial as guaranteed by the Pennsylvania Constitution.

Despite these renewed indications of judicial concern in a few states, medical malpractice compensation schemes that restrain individuals' exercise of their protective tort remedies are still

---

145 *Id.*

146 *Id.* The court noted, "It should be emphasized that today's decision is not premised on a reevaluation of the wisdom of the *Carter* decision. Rather it is based on the unfortunate fact that the medical mediation statute has proven unworkable and inequitable in practical operation." *Id.*

147 49 U.S.L.W. 2249 (Pa. S. Ct., Sept. 22, 1980).

148 Parker v. Children's Hosp. of Philadelphia, 483 Pa. 106, 394 A.2d 932 (1978).

149 Mattos v. Thompson, 49 U.S.L.W. 2249, *quoting Parker*, 394 A.2d 932.

firmly entrenched in the vast majority of states. The prevailing judicial reaction has been highly deferential.

III. Tort Law Theory and the Emergence of the "Quid Pro Quo"

A. *Tort Remedies and the "Social Contract"*

Early American legal development was posited upon natural rights principles of personal liberty and private property that inured to the individual by virtue of his citizenship. The exercise of governmental power was to be reconciled with the preservation of these inherent values which comprised the central tenet of the social contract upon which consensual government was based.[150]

For individual liberties, the state's adjudication of private tort actions helped fulfill the premise by which individuals agree to cede power to the state: that collectivity provides greater protection for personal security against private assault and encroachment.[151] Common-law tort remedies were predicated upon the need to compensate injured persons for unreasonable interference with their interests.[152] Contemporary tort theory also recognizes the importance of deterrence. Social order is theoretically maintained by the symbiosis of these dual functions — by judicially imposing a sufficiently high level of compensatory and punitive damages, private individuals can effectively be deterred from intentionally or unreasonably inflicting injury on others.[153]

As a tool of economic and social policy, tort actions promote a scheme of market-based accountability that closely parallels "free-market" economic theory of production and distribution determined by the aggregate of individual consumer choice.

---

150 *See* J. Rousseau, Social Contract (E. Barker ed. 1962); J. Locke, The Second Treatise of Civil Government (P. Laslett ed. 1960); E. Corwin, Liberty Against Government (1948).

151 *See* T. Hobbes, Leviathan, at 139-43, 170-72 (Bobbs-Merrill ed. 1958); *see generally* M. Olson, The Logic of Collective Action (1971).

152 *See* W. Prosser, Torts 6 (4th ed. 1971).

153 *Id.* at 9.

When a strict liability standard is imposed on hazardous business activities, the burden is equalized over the entire industry and the risk is transferred to the most efficient loss-spreader.[154]

As the unfettered avariciousness of the free-market economy in the United States became socially intolerable, the state's regulatory[155] and welfare[156] interventionist roles increased. The influence of public law emerged where previously only private law had governed.

At the same time that it was woodenly invoking the substantive due process doctrine to strike down legislation impinging on individuals' perceived liberty and property *rights,*[157] the Supreme Court was formulating flexible balancing standards to treat the common-law *remedies* privately employed to enforce these rights. To freeze common-law remedies as vested property rights would frustrate the state's ability to employ innovative remedial schemes for the collective welfare.[158] However, to allow the state to facilely abrogate private law tort remedies — where deterrence and compensation principles were central to protecting individual liberty interests — would permit the pernicious harm focused on one individual to be subsumed by a mere assertion of collective expediency. The Court's aversion to the inherent political socialism of sweeping collective benefit justifications that functionally abrogate individual remedies can best be seen in the early eminent domain and workmen's compensation cases.

It suffices here to observe that the Supreme Court has held that property rights may be invaded for the public benefit, but where regulation so severely focuses harm upon one individual or totally undermines a right, it constitutes a "taking."[159] Such

---

154 *See id.* at 531.

155 *See, e.g.,* Sherman Antitrust Act, 15 U.S.C. §§ 1-7 (1978).

156 For example, child labor laws enacted by many states.

157 *See, e.g.,* Lochner v. New York, 198 U.S. 45 (1905).

158 *See* Silver v. Silver, 280 U.S. 117, 122 (1929); Munn v. Illinois, 94 U.S. 113, 134 (1877).

159 Justice Holmes posited his often repeated "balancing test" in Pennsylvania Coal Co. v. Mahon, 260 U.S. 393, 415-16 (1922):

> The general rule at least is, that while property may be regulated to a certain extent, if regulation goes too far it will be recognized as a taking. . . . We are in danger of forgetting that a strong public desire to improve the public condition is not enough to warrant achieving the desire by a shorter cut than the constitutional way of paying for the change.

regulation is functionally equivalent to condemnation by eminent domain, and, therefore, the government must provide reasonable compensation.[160]

When the individual believes that governmental regulation, or authorization of private action, has intruded or drastically diminished his property value without providing "just compensation," he retains a residual cause of action to claim a taking or proceed by inverse condemnation.[161]

B. *The "Quid Pro Quo" in Workmen's Compensation Schemes*

1. Theory and Function

State workmen's compensation acts represented the first major shift from private common-law tort remedies based on negligence principles to a strict liability regulatory framework.

Compensation systems may be divided into those in which the purpose is to deter aberrant, avoidable future conduct, and those in which the primary function is to efficiently provide quick compensation to a class of injured persons without noticeably imposing a burden upon those who will ultimately absorb the loss.[162] The negligence-based tort compensation system promotes deterrence. A strict liability system, in providing compensation for a broad range of assumed inevitable injuries, does not address deterrence.

For a lucid theoretical discussion of the generally confused "takings" doctrine, *see* Michelman, *Property, Utility, & Fairness: Comments on the Ethical Foundations of 'Just Compensation' Law*, 80 HARV. L. REV. 1165 (1967).

160 *Id.* The Supreme Court's recent decision in Penn Central Transp. Co. v. New York City, 438 U.S. 104 (1978), is wholly consistent with the approaches discussed in this Note. In upholding New York City's Landmark Preservation Law against a "takings" challenge, the Court placed great weight on the law's assurance that Penn Central was entitled to a "reasonable return," *id.* at 110-15 & n.13, had received "valuable" benefits of transferable development rights in exchange for property use restrictions, *id.* at 122, 129, 137, and had been able to actively participate in an individualized administrative process determining the extent of the limitations, *id.* at 110-19.

161 *E.g.*, United States v. Causby, 328 U.S. 256 (1946).

162 *See* Chittenden, *The Designable Compensable Event in Medical Malpractice*, 12 FORUM 919 (1977); Keeton, *Conditional Fault in the Law of Torts*, 72 HARV. L. REV. 401 (1959).

Workmen's compensation laws are premised on the unavoidability of injuries and death in the course of employment.[163]

> Certainly, the operation of industrial establishments that in the ordinary course of things frequently and inevitably produce disabling or mortal injury to the human beings employed is not a matter of wholly private concern." (*citing Mountain Timber Co. v. Washington,* 243 U.S. 219, 239 (1917))
>
> Hence recognizing that injuries to workmen constitute a part of the unavoidable cost of hazardous industries, we will require that it be assumed by the one in control of the industry as employer, just as he pays other items of cost. . . .[164]

Within the rapidly industrializing American society, an appalling number of workers were being killed and maimed in workplace accidents.[165] The negligence-based tort system confronted the injured worker with costly and tedious litigation. The employers' assertion of the common-law defenses of contributory negligence, assumption of risk, and the fellow-servant rule often proved to be insurmountable obstacles. The worker and his family were forced to bear the unsustainable loss and were often overcome by poverty.[166] The workmen's compensation scheme addressed this pattern of uncompensated death and hardship by establishing countervailing sacrifices upon the employee and employer; in return for the loss of his common-law tort remedy, the worker was to receive a speedy and guaranteed recovery for his injury. The system minimized transaction costs by imposing strict liability and by waiving the employer's potential common-law defenses.[167]

In short, this new scheme assumed an amoral probabilistic form of causation and shifted from a view of individuals in society to that of a mass society capable of effective loss-spreading. The imposition of strict liability on the employer was the *quid pro quo* for administrative restraints imposed on the em-

---

163 *See* Arizona Employers' Liability Cases, 250 U.S. 400, 422-23, 425 (1919); New York Central R. R. Co. v. White, 243 U.S. 188, 203 (1917).

164 Arizona Employers' Liability Cases, 250 U.S. 400, 425 (1919).

165 *See* C. EASTMAN, WORK-ACCIDENTS AND THE LAW, (1910); J WEINSTEIN, THE CORPORATE IDEAL IN THE LIBERAL STATE: 1900-1918, at 40 (1968).

166 New York Central R. R. Co. v. White, 243 U.S. 188, 197 (1917); C. EASTMAN, *supra* note 165, at 119-24.

167 *See* PROSSER, *supra* note 152, at 531.

ployee's recovery. Efficient compensation delivery was emphasized in view of the implausibility of deterrence.

2. Judicial Implication of a "Reasonably Just Substitute"

In upholding the New York workmen's compensation laws against employers' constitutional challenges in *New York Central Railroad Co. v. White,*[168] the Supreme Court stressed the public benefits of the scheme and asserted that "no person has a vested right in any rule of law entitling him to insist that it shall remain unchanged for his benefit."[169] Although the Court did not expressly require that a "reasonably just substitute" be provided, it strongly intimated that the statutory *quid pro quo* was critical:

> Nor is it necessary, for the purposes of the present case, to say that a State might, without violence to the constitutional guaranty of 'due process of law,' suddenly set aside all common-law rules respecting liability between employer and employee, without providing a reasonably just substitute . . . it perhaps may be doubted whether the State could abolish all rights of action on the one hand, or all defenses on the other, without setting up something adequate in their stead. No such question is here presented, and we intimate no opinion on it. The statute under consideration sets aside one body of rules only to establish another system in its place. If the employee is no longer able to recover as much as before in case of being injured through the employer's negligence, he is entitled to moderate compensation in all cases of injury, and has a certain and speedy remedy without the difficulty and expense of establishing negligence or proving the amount of damages. . . . On the other hand, the employer is left without defense respecting the question of fault . . .[170]

The court ultimately determined that the New York laws, in fact, provided the employees with a sufficient *quid pro quo* in that benefits created by the statute were exchanged for abridged common-law benefits.[171]

---

168 243 U.S. 188 (1917).

169 *Id.* at 198.

170 New York Central R. R. Co. v. White, 243 U.S. 188, 201 (1917).

171 A similar theory of a "reasonably just substitute" was articulated by the court in Crowell v. Benson, 285 U.S. 22 (1932) (upholding the Longshoremen's and Harbor Workers' Compensation Act):

It is useful to examine this *quid pro quo* in the context of the balance of economic power confronted by the *New York Central R. R. Co.* court: the business enterprise was regarded as an efficient loss-spreader, able to adjust wages and prices to allocate the economic benefits and burdens imposed by the statutory scheme;[172] to the contrary, the individual worker — a highly inefficient loss-spreader — was sheltered against focused harm and undue burden.[173] Therefore, while perhaps not necessary to protect the economic interests of the business enterprise, a reasonably just substitute was essential to safeguard individual liberty and personal security. Otherwise, the individual could be subject to a collusive scheme between business and the state, justified by "public benefit" concerns, that would leave him with even less structural protection than his prior weak common-law tort remedy.

This intimation that a *quid pro quo* doctrine would emerge as an element of constitutional due process has fluttered ambiguously since its first airing in *New York Central R. R. Co.* Though seemingly discredited in its formal application, it has never been explicitly rejected despite numerous opportunities.

Soon after *New York Central R. R. Co.*, the Supreme Court faced another employer challenge to a state workmen's compensation scheme for hazardous industries.[174] In response to the *employers'* assertion that the statute unconstitutionally failed to provide them with a sufficient benefit in return for the imposition of strict liability, a divided Court[175] repeated the *New York Central R. R. Co.* declaration that "no opinion was intimated" pertaining to the necessity for such a substitute, and stated that "[w]e cannot, however, regard this statute as anything else other than a substitute for the law as it previously stood; whether it be a proper substitute was for the people to determine . . ."[176]

> In view of the difficulties which inhere in the ascertainment of actual damages, the Congress was entitled to provide for the payment of amounts which would reasonably approximate the probable damage. *Id.* at 41.

172 New York Central R. R. Co. v. White, 243 U.S. 188, 201-04 (1917); *see* Arizona Employers' Liability Cases, 250 U.S. 400, 424, 427 (1919).

173 New York Central R. R. Co., 243 U.S. at 203; *see* Arizona Employers' Liability Cases, 250 U.S. 400, 423 (1919).

174 Arizona Employers' Liability Cases, 250 U.S. 400 (1919).

175 The Court split 5 to 4 in its decision. Three of the justices in the majority, including Justices Holmes and Brandeis, filed separate concurring opinions.

176 Arizona Employers' Liability Cases, 250 U.S. 400, 427 (1919).

This erosion of a potential *quid pro quo* requirement was followed by the Supreme Court's clear statement in *Silver v. Silver*[177] as it upheld Connecticut's automobile guest statute[178] against a due process attack: "[the] Constitution does not forbid the creation of new rights, or the abolition of old ones recognized by the common law, . . . to attain a permissible legislative object."[179]

Only two years later, however, the Supreme Court rekindled this flickering flame when it reviewed the constitutionality of the Longshoremen's and Harbor Workers' Compensation Act[180] in *Crowell v. Benson.*[181] In upholding a limitation on recovery as a proper *quid pro quo* for the imposition of strict liability, the Court affirmed: "In view of the difficulties which inhere in the ascertainment of actual damages, the Congress was entitled to provide for the payment of *amounts which would reasonably approximate the probable damages.*"[182]

Since this decision, the development of a *quid pro quo* doctrine has remained dormant and ambiguous. Application by state courts and lower federal courts has been erratic. In general, courts have tended either to presume the viability of the doctrine and enunciate the requisite *quid pro quo,*[183] or to assert the

---

177 280 U.S. 117 (1929).

178 Automobile guest statutes forbid non-paying passengers to sue drivers for injury or death occurring from a collision. These statutes have been widely criticized as pernicious and blatant special interest legislation attributable to baneful lobbying by insurance companies. *See, e.g.,* Gibson, *Guest Passenger Discrimination,* 6 ALBERTA L. REV. 211 (1968); Note, *supra* note 21, at 934 n.623. The majority of state courts have held these laws to be unconstitutional. *See* Note, *supra* note 21, at 895-96 n.400 (collecting cases).

The alleged purpose of such legislation is to prevent collusion and fraudulent insurance claims. It is posited that there is no "less restrictive alternative" to the drastic step of abrogating the injured passengers' common-law tort remedy. To the extent that fraud is the chief concern, there is no legitimate interest in affording a cause of action to the dishonest claimant. In the non-fraudulent case, the passenger's potential tort suit is not necessary to deter hazardous driving; presumably, the risk of self-injury to the driver is the more effective, natural deterrent.

It should also be noted that the burden cast by this legislation falls equally on all individuals — theoretically, we are all equally likely to be drivers or passengers. The medical malpractice legislative schemes, however, specially benefit and protect only those individuals who are physicians while burdening all individuals who suffer injury.

179 Silver v. Silver, 280 U.S. 117, 122 (1929).

180 33 U.S.C. § 901 *et seq.*

181 285 U.S. 22 (1931).

182 *Id.* at 41 (emphasis added).

183 *E.g.,* Ritholz v. March, 105 F.2d 937, 939 (D.C. Cir. 1939) (Federal Trade Commission procedural changes); Lasky v. State Farm Ins. Co., 296 So. 2d 9, 14 (Fla.

doctrine's invalidity, but carefully prove that such a requirement has nevertheless been met.[184] In particular, state courts have considered the constitutional validity of "no-fault" automobile insurance statutes by examining whether a reasonably just substitute remedy provides plaintiff with a prompt and less costly procedure for the guaranteed recovery of a significant portion of his losses.[185]

C. *Duke Power and the "Quid Pro Quo"*

Almost 50 years after *Silver* and *Crowell* left the *quid pro quo* in doctrinal limbo, the Supreme Court once again confronted the issue in *Duke Power v. Carolina Environmental Study Group*.[186] The litigation involved a constitutional attack on the Price-Anderson Act,[187] which places a $560 million ceiling on liability resulting from a nuclear power plant accident. In return, injured citizens are presumedly guaranteed a quick recovery under a strict liability system that waives common-law defenses. The statutory scheme was enacted by Congress in response to the inability of the private nuclear power companies to obtain adequate insurance coverage and their threat "that they would be forced to withdraw from the field if their liability were not limited by appropriate legislation."[188]

Plaintiffs were two citizens groups and forty individuals who lived near a planned nuclear power plant. In addition to concern over environmental and health damage from the operation of the nuclear power plant, plaintiffs alleged that in the event of

---

1974) (automobile no-fault statute); Grace v. Howlett, 51 Ill. 2d 478, 283 N.E.2d 474 (1972) (automobile no-fault statute).

184 *See, e.g.*, Carr v. United States, 422 F.2d 1007, 1011 (4th Cir. 1970) (Federal Drivers' Act); Sparks v. Wyeth Laboratories, Inc., 431 F. Supp. 411, 416-17 (W.D. Okla. 1977) (Federal Swine Flu Act); Montgomery v. Daniels, 38 N.Y.2d 41, 56, 340 N.E.2d 444, 453 (1975) (automobile no-fault statute).

185 *See, e.g.*, Lasky v. State Farm Ins. Co., 296 So. 2d 9, 15 (Fla. 1974); Pinnick v. Cleary, 360 Mass. 1, 271 N.E.2d 592 (1971).

186 438 U.S. 59 (1978).

187 71 Stat. 576, 42 U.S.C. § 2210 (1973).

188 *Duke Power*, 438 U.S. at 64 (citing *Hearings Before the Joint Committee on Atomic Energy on Government Indemnity for Private Licensees and AEC Contractors Against Reactor Hazards*, 84th Cong., 2d Sess. 9, 109-10, 115, 120, 136-37, 148, 181, 195, 240 (1956)).

a nuclear accident, their property would be taken without an assurance of just compensation.[189] Plaintiffs contended that "the Price-Anderson Act — which both creates the source of the underlying injury and limits the recovery therefor — constitutes such arbitrary action . . . and is the instrument of the taking since on this record, without it, there would be no power plants and no possibility of an accident."[190]

The district court had held that the Price-Anderson Act contravened both the due process and equal protection components of the Fifth Amendment.[191] The Act violated the due process clause because "[t]he amount of recovery is not rationally related to the potential losses; [because the Act] tends to encourage irresponsibility in matters of safety and environmental protection; [and because] there is no *quid pro quo*" for the liability limitations.[192] The Act was found to violate the equal protection clause because it "places the cost of [nuclear power] on an arbitrarily chosen segment of society, those injured by nuclear catastrophe."[193]

The United States Supreme Court reversed the district court on both the due process and equal protection grounds.[194] In arriving at its holding, the Court utilized a minimum rationality test for the purportedly economic regulation,[195] and then tentatively applied the *quid pro quo* doctrine[196] first intimated in *New York Central R. R. Co. v. White.*[197]

---

189 *Duke Power,* 438 U.S. at 69.

190 *Id.* at 69.

191 Carolina Environmental Study Group v. United States, 431 F.Supp. 203 (W.D.N.C. 1977), *rev'd sub nom,* Duke Power v. Carolina Environmental Study Group, 438 U.S. 59 (1978).

192 431 F. Supp. at 222-23.

193 *Id.* at 225.

194 The Supreme Court found "it unnecessary to resolve the claim that such an accident would constitute a 'taking' as that term has been construed in our precedents since our reading the Price-Anderson Act does not withdraw the existing Tucker Act remedy, 28 U.S.C. §1491." *Duke Power,* 438 U.S. at 94 n.39.

195 *Id.* at 83.

196 *Id.* at 87-88. The Court noted, "Initially, it is not at all clear that the Due Process Clause in fact requires that a legislatively enacted compensation scheme either duplicate the recovery at common law or provide a reasonable substitute remedy. However we need not resolve this question here since the Price-Anderson Act does, in our view, provide a reasonably just substitute for the common-law or state tort law remedies it replaces." *Id.*

197 243 U.S. 188 (1917).

1. The Minimum Rationality Test

Plaintiffs argued that the court should apply "a more elevated standard of review on the ground that the interests jeopardized by the Price-Anderson Act 'are far more important than those in the economic due process and business-oriented cases' where the traditional rationality standard has been invoked."[198] Once the Supreme Court rejected this intermediate scrutiny standard, the constitutionality of the Price-Anderson Act as to its social purposes was virtually assured.

The district court had focused its scrutiny on the health and welfare hazards to private individuals which were generated by the Act.[199] The drastic state-authorized intrusion on individual liberties stood in stark opposition to traditional economic regulation of "business practices." The Supreme Court, however, viewed the Act as simply a scheme

> to stimulate the private development of electric energy by nuclear power while simultaneously providing the public compensation in the event of a catastrophic nuclear incident. (citation omitted) The liability-limitation provision thus emerges as a classic example of economic regulation — a legislative effort to accomodate 'the burdens and benefits of economic life.'[200]

In *Usery v. Turner Elkhorn Mining Co.*,[201] the Court reviewed a mine operator's challenge to legislation which expanded remedies available to miners with job-related disabilities.[202] The Court accorded deference to this "business practice" regulation as a legislative attempt to structure economic burdens and benefits. As with the workmen's compensation schemes, the legislation protected individual workers' liberties, and imposed the costs on the business enterprise — the most-efficient loss-spreader.[203]

---

198 *Duke Power,* 438 U.S. at 83.

199 It was assumed that nuclear power plants, the source of the injury, would not have been constructed absent the Price-Anderson Act. *Id.* at 64.

200 *Id.* at 83 (quoting Usery v. Turner Elkhorn Mining Co., 428 U.S. 1 (1976)).

201 428 U.S. 1 (1976).

202 Title IV of the Federal Coal Mine Health and Safety Act of 1969, 83 Stat. 792, as amended by the Black Lung Benefits Act of 1972, 86 Stat. 150, 30 U.S.C. §901 *et seq.* (1980), provides benefits to coal miners suffering from "black lung disease" (pneumoconiosis) and to their survivors.

203 *See* notes 172 to 174 and accompanying text *supra.*

Duke Power extended the holding, but not the underlying rationale, of *Usery* to the Price-Anderson Act. Here, the risk of economic burden from a nuclear accident has been shifted from nuclear plant owners to the individuals who reside near the plants.[204] This extension of judicial deference from legislation simply regulating business practices to that limiting tort liability is troubling. As one commentator has noted:

> Any law which removes a burden from A and places it on B adjusts the benefits and burdens of economic life regardless of whether it alters a business practice of A. If one substitutes 'corporations' for A and 'product liability claimants' for B, the vast scope of this new formulation is evident.[205]

This fear of widespread legislative abrogation of private common-law remedies is tempered by two significant strands of the *Duke Power* decision: (1) The uniqueness of the nuclear power industry; and (2) The Court's emphatic determination that the Act does, in fact, provide a *quid pro quo* that sufficiently protects individual rights.

2. Application of the Minimum Rationality Test to Promotion of Nuclear Power

To withstand a due process challenge, statutory economic regulation need only be rationally related to a permissible legislative objective.[206] In order to prevail, plaintiffs must "establish that the legislature acted in an arbitrary or irrational way."[207]

The Supreme Court's initial premise in *Duke Power* was that nuclear power is a unique form of energy production, requiring

---

204 For a nuclear catastrophe with damages of greater than $560 million, the costs are externalized to those least able to bear the loss. An alternative scheme could either force the nuclear industry to internalize the costs, or spread the risk to the general public by having the federal government assume the cost burden.

205 Dickerson, *Limited Liability for Nuclear Accidents: Duke Power Co. v. Carolina Environmental Study Group, Inc.*, 8 Ecology L. Q. 163, 170-71 (1979). For a suggestion that such extraordinary protection be afforded to corporations faced with products liability suits, *see* White & McKenna, *Constitutionality of Recent Malpractice Legislation*, 13 Forum 312, 330 (1977).

206 *See, e.g*, Ferguson v. Skrupa, 372 U.S. 726 (1963).

207 *Duke Power*, 438 U.S. at 83 (citing Usery v. Turner Elkhorn Mining Co., 428 U.S. 1, 15 (1976)).

extraordinary solicitude.[208] Private development had been explicitly encouraged by Congress,[209] but was hindered by extraneous economic forces. Though "the risk of a major nuclear accident was extremely remote,"[210] "the very uniqueness of nuclear power meant that the possibility remained and the potential liability dwarfed the ability of the industry and private insurance companies to absorb the risk."[211]

Unlike the medical malpractice or products liability fields, which operate in the competitive market system, the capital-intensive nuclear industry represents a unique public-private partnership. Initially, Congress intended to develop nuclear power as a government monopoly.[212] In the 1950's, however, the government ceded responsibility for atomic energy generation to private industry operating as a natural monopoly within the public utility regulation model. Furthermore, plant construction, operation, and ownership are subject to an elaborate licensing and safety framework supervised by the federal Nuclear Regulatory Commission.[213]

In light of the Court's implicit endorsement of the legitimacy of the government's promotion of nuclear power, the economic due process standard required only that the means chosen not be arbitrary or irrational. The Court transformed the plaintiffs' claim that the liability limitation was not rationally related to the aggregation of individuals' potential losses[214] into a more

---

208 *Id.* at 64.

209 The Court stated: "Congress concluded that the national interest would be best served if the Government encouraged the private sector to become involved in the development of atomic energy for peaceful purposes under a program of federal regulation and licensing." *Id.* at 63 (citation omitted).

210 *Id.* at 64. At the time of the *Duke Power* decision, there was general public confidence in the safety of nuclear power and the extreme remoteness of a major accident. While not necessarily determinative, these factors obviously impressed the Supreme Court. *See id.* at 64, 85, 86 n.30, 91. One can only speculate how the Court's decision would have been affected if the case had been argued *after* the accidental release of radioactive materials from the Three Mile Island nuclear plant on March 28, 1979.

211 *Id.*

212 *Id.* at 63; *see also* J. FULLER, WE ALMOST LOST DETROIT (1975).

213 *Id.* at 63, 87. The Nuclear Regulatory Commission replaced the Atomic Energy Commission as the licensing and regulatory agency in 1974.

214 *Id.* at 84; *see* Carolina Environmental Study Group v. United States, 431 F. Supp. at 222. "The remaining contention that recovery is uncertain because of the aggregate rather than individualized nature of the liability ceiling is but a thinly disguised version of the contention that the $560 million figure is inadequate which we have already rejected." 438 U.S. at 92.

general question of whether the limit itself was rationally related to the promotion of nuclear power.[215] Since a statutory limit on recovery of $10 would absurdly satisfy the legislative goal of encouraging the private development of nuclear power, the Court examined only the "arbitrariness of the *particular figure* of $560 million."[216] Essentially, this is a test for a "societal *quid pro quo*" — where the legislation grants extraordinary favor to a particular private industry, the aggregate citizen recovery in return for the collective abrogation of state common-law tort remedies cannot be wholly irrational.

The Court acknowledged that "whatever ceiling figure is selected will, of necessity, be arbitrary,"[217] but strained to find this liability ceiling constitutionally permissible because:

> [t]he legislative history clearly indicates that the $560 million figure was not arrived at on the supposition that it alone would necessarily be sufficient to guarantee full compensation in the event of a nuclear incident. Instead, it was conceived of as a 'starting point' or working hypothesis.[218]
>
> . . . When appraised in terms of both the extremely remote possibility of an accident where liability would exceed the limitation and Congress' now statutory commitment to 'take whatever action is deemed necessary and appropriate to protect the public from the consequences of' any such disaster . . . [the ceiling is] within permissible limits.[219]

It seems highly unlikely that a court reviewing a medical malpractice compensation scheme would reasonably determine that a medical malpractice accident was similarly such an "extremely remote possibility." Furthermore, existing state regimes do not make an express commitment to take further steps should the liability ceiling provide inadequate compensation for an injured individual.[220]

---

215 *Id.* at 84, 94.

216 *Id.*

217 *Id.* at 85-86 n.29.

218 *Id.* at 85 (footnote omitted).

219 *Id.* at 86 (footnote omitted); *see also* note 210 *supra.*

220 The North Dakota Supreme Court is the only state court to have expressly considered the applicability of *Duke Power* in evaluating the constitutionality of a state scheme limiting medical malpractice liability. The court concluded that "Duke Power Co., . . . is of no assistance [because] [h]ere, we have a strong possibility of damages above the limitation and no Legislative commitment beyond the limitation." Arneson v. Olson, 270 N.W.2d 125, 135 n.6 (1978).

In overturning the district court's finding that the Price-Anderson Act contravened the due process clause because it "tends to encourage irresponsibility . . . on the part of builders and owners" of nuclear power plants,[221] the Court referred to the multitude of safety regulations guiding construction and operation.[222] By asserting that "nothing in the liability limitation provision undermines or alters in any respect the rigor and integrity of that process,"[223] the Court tacitly conceded that the Act had no effect on business practices. Furthermore, the Court stressed that "the risk of financial loss and possible bankruptcy to the utility is in itself no small incentive to avoid the kind of irresponsible and cavalier conduct attributed to the licensees. . ."[224]

The Court's assumptions are consistent with the predicates of the workmen's compensation and automobile no-fault statutes. In workmen's compensation schemes, it is assumed that workplace injuries are inevitable but may be controlled by regulation;[225] similarly, the Court concluded that federal regulation enhances safety in the nuclear industry. In automobile no-fault schemes, deterrence is posited not from tort remedies, but by the natural assumption that a tortfeasor would not choose to drive recklessly because his own life would be endangered in an accident; similarly, the Court concluded that a nuclear plant owner would not operate recklessly because he would be bankrupted by a tragic accident.

In contrast, the economic incentives guiding medical treatment quality are far different from the economic premises underlying the workman's compensation, automobile no-fault, and nuclear accident compensation schemes. The private physician-patient relationship is not easily shaped by prospective regulation.[226] Medical malpractice injuries are not considered to be inevitable, but may be preventable. Thus, compensation through

---

221 431 F. Supp. at 222.

222 *See also* Vermont Yankee v. Natural Resources Defense Council, 435 U.S. 519, 526-27 (1978).

223 *Duke Power*, 438 U.S. at 87.

224 *Id.*

225 *See, e.g.*, Occupational Safety and Health Act, 84 Stat. 1590, 29 U.S.C. §651 *et. seq.* (1975).

226 Professional Standard Review Organizations — "PSROs" — may provide some elements of retrospective evaluation and control.

a private tort action for medical malpractice not only aids the victim, but is justified as an effective economic counterforce to deter irresponsible conduct by physicians.[227]

3. The "Quid Pro Quo" of a "Reasonably Just Substitute"

The Supreme Court explicitly refused to decide whether the due process clause requires the legislature to provide a *quid pro quo* when it abrogates common-law tort remedies.[228] Despite avoiding adjudication of this underlying constitutional issue,[229] the Court went to great lengths to demonstrate that, in its view, the Price-Anderson Act's compensation scheme did provide a reasonably just substitute.[230]

The district court had determined that the Act failed to provide an adequate substitute of " 'reasonable, certain and adequate provision for obtaining compensation' which due process of law requires."[231] Clearly, the Supreme Court could have reversed the lower court by simply denying the validity of the *quid pro quo* test altogether. Instead, the Court chose to cast the Price-Anderson Act into the mold established by *New York Central R. R. Co. v. White*[232] and *Crowell v. Benson*.[233] By

227 The adverse publicity from a high malpractice award is a further deterrent to irresponsible conduct by physicians.

228 *Duke Power*, 438 U.S. at 87-88, 93 & nn.32, 33.

229 The Court indicated that prior cases had "clearly established" that a person has "no vested interest, in any rule of the common law." *Id.* at 88 n.32. The refusal to reject the requirement of a reasonably just substitute is not necessarily in conflict with this proposition; it simply affirms that while the individual cannot rely on the perpetual continuance of a precise common-law remedy, he is entitled to a concomitant substitute. As the Court expressly stated: "The logic of *New York Central* [i.e., suggesting the necessity of a *quid pro quo*] would seem to apply with renewed force in the context of this challenge to the Price-Anderson Act." *Id.* at 93. If the court was simply reaffirming that the minimum rationality standard for equal protection is as deferential as that in due process, then this *dicta* still leaves the nuclear accident victim with the possible protection of a reasonably just substitute.

To the extent that an independent equal protection analysis is being asserted, based solely on the general rationality of the legislation itself rather than the general rationality of the classification to the legislatve goal, the Court is charting an entirely new equal protection course in this brief *dicta*. This seems highly unlikely.

230 *Id.* at 87-93.

231 431 F. Supp. at 224 (*quoting* Regional Rail Reorganization Act Cases, 419 U.S. 102, 124-25 (1974); Cherokee Nation v. Southern Kansas Railroad Co., 135 U.S. 641, 659 (1890)).

232 243 U.S. 188 (1917).

233 285 U.S. 22 (1932). *See also Duke Power*, 438 U.S. at 88.

weighing the benefits and burdens of the Act differently than the lower court did, the Supreme Court concluded that a satisfactory *quid pro quo* was provided. That the Court felt compelled to demonstrate this necessary balance is significant and explainable.

Because the Court restricted its due process analysis to the liability limitation's general rational relationship to the legislative purpose of promoting nuclear power,[234] it could have held it unnecessary to provide any compensation at all to the nuclear accident victims. But forcing a random class of citizens — those unlucky enough to reside near a nuclear plant — to bear the entire costs of a nuclear accident, would have raised substantial "taking" and equal protection questions which the Court seemed anxious to avoid.[235]

Traditional equal protection analysis requires that the classification be rationally related to the goal of the legislation.[236] Furthermore, there must be "some rationality in the nature of the class singled out."[237] In satisfying this second requirement, it would appear more legitimate to either impose the burden on the general public by having the federal government act as a cost-spreading insurer or force the industry to internalize its costs.[238] Imposing the entire burden on a random class of citizens who live near plants and become individual accident victims — those least able to bear the loss — could be logically justified only if they were to be provided with a reasonably just substitute for their existing common-law tort remedy.[239] Requiring, or finding the existence of, a sufficient *quid pro quo* was imperative

---

234 438 U.S. at 93.

235 *Id.* at 93-94 (avoiding equal protection analysis), & n.39 (avoiding a "taking" claim).

236 *See* L. TRIBE, *supra* note 50, § 16-2, at 995. For a critical discussion of the *Duke Power* Court's treatment of the equal protection claim, *see* Dickerson, *supra* note 205, at 179-81.

237 Rinaldi v. Yeager, 384 U.S. 305, 308-09 (1966).

238 The inability of the nuclear power industry to obtain adequate insurance rendered this second solution impossible.

239 At the end of its opinion, the Court stated that "[t]he general rationality of the Price-Anderson Act liability limitations — particularly with reference to . . . [promoting] nuclear energy — is ample justification for the difference in treatment between those injured in nuclear accidents and those whose injuries are derived from other causes." 438 U.S. at 93. This statement followed the Court's assertion that "the equal protection arguments largely track and duplicate those made in support of the due process claim." *Id.* at 93.

to avoid adjudication of the difficult equal protection problems posed by this Act.

Plaintiffs argued that the $560 million compensation fund would be inadequate in the event of a major nuclear accident, that the Act's waiver of common-law defenses was superfluous since strict liability already applied to the ultra-hazardous operation of a nuclear plant, and that the Act's claims administration process would not guarantee prompt recovery. The Supreme Court disagreed and found sufficient benefits for nuclear accident victims on all three counts.

First, the Court sought to establish that the compensation fund was sufficient because it provided a guaranteed recovery,[240] while a victim relying on private tort law might come away empty-handed from a bankrupt nuclear power operation.[241] To presume the reasonableness of this liability ceiling and assure that there was, indeed, a reasonably just substitute, the Court found it necessary to assert *seven* times that "the Act contains an explicit congressional commitment to take further action to aid victims of a nuclear accident in the event that the $560 million ceiling on liability is exceeded."[242] Such an interpretation of the ambiguous legislative history[243] amounts to extraordinary prospective appropriation by the judiciary.

Second, the Court found that the Act's waiver of defenses benefited the plaintiffs because there was no assurance that a strict liability standard would be applied to a unique, future nuclear accident, and because common-law exceptions for acts of God or of third parties might otherwise be asserted. The Court affirmed that "the . . . waiver of defenses establishes at the threshold the right of injured parties to compensation without proof of fault and eliminates the burden of delay and uncertainty which would follow from the need to litigate after an accident."[244] The Court then quickly disposed of plaintiff's contention that the delay in compensation under the Act would exceed that inherent in common-law tort litigation.[245]

---

240 *Id.* at 92, 93. For a critique of the adequacy of the compensation fund, *see* Dickerson, *supra* note 205, at 175-77.
241 *Id.* at 91 & n.36.
242 *Id.* at 93; *see id.* at 66-67, 85, 86, 87 n.31, 90-91.
243 *See id.* at 66.
244 *Id.* at 91.
245 *Id.* at 92.

After citing and implicitly endorsing the logic of the *quid pro quo* compensation scheme set forth in *New York Central R. R. Co.*,[246] the Supreme Court concluded:

> The Price-Anderson Act not only provides a reasonable, prompt and equitable mechanism for compensating victims of a catastrophic nuclear incident, it also guarantees a level of net compensation generally exceeding that recoverable in private litigation. Moreover, the Act contains an explicit congressional commitment to take further action to aid victims of a nuclear accident in the event that the $560 million ceiling liability is exceeded. This *panoply of remedies* and guarantees is at the least a *reasonably just substitute* for the common-law rights replaced by the Price-Anderson Act.[247]

IV. Medical Malpractice Victims as a "Semi-Suspect" Class with "Quasi-Fundamental" Rights

While *Duke Power* did not resolve the ambiguous constitutional status of the *quid pro quo* doctrine, the Court's reasoning reinforced the contours of the "reasonably just substitute" standard defined in the workmen's compensation cases. Furthermore, the rationale for establishing a *quid pro quo* to protect individual rights against total social domination and expedience was logically compelled by the Court's strict adherence to the rigidly stratified two-tier due process analysis in evaluating the social legitimacy of the statutory scheme. The doctrinal deference of the minimum rationality test may function to assure that a tort compensation scheme satisfies the "societal *quid pro quo*" — that the legislation is not simply a cleverly concealed scheme for special interest protectionism wholly unrelated to fulfilling the asserted public purpose.[248] However, absent the

246 *Id.* at 93. *See* notes 228 to 230 and accompanying text *supra.*

247 *Duke Power,* 438 U.S. at 93 (emphasis added).

248 *See id.* at 82-87; *but cf.* City of New Orleans v. Dukes, 427 U.S. 297 (1976) (upholding ordinance permitting two long-established vendors to continue operations in the French Quarter while forbidding new businesses; for the purpose of promoting the area's "tourist-oriented charm" so as to aid the city economy). Although the ordinance at issue in *City of New Orleans* clearly is special interest legislation, the scope of the protection is readily apparent and may easily generate a challenge through the political process. By contrast, the public, as potential medical malpractice victims, has more to fear from an oblique tort compensation scheme that avoids direct political attack by appearing to provide a significant benefit that is, in reality, a chimera. *See* Smith v. Cahoon, 283 U.S. 553 (1931); L. Tribe, *supra* note 50, § 16-4, at 998.

requisite of the reasonably just substitute — the "individual *quid pro quo*" — the judicial tools are impotent to protect individual liberties. The following portions of this Note will specifically apply this doctrinal concept to state statutes restricting recovery for medical malpractice.

A. *Medical Malpractice Victims as a "Semi-Suspect" Class*

1. "Semi-Suspect" Classes in the Pluralist Legislative Process

Due process and equal protection judicial scrutiny of legislation have evolved as counter-majoritarian safeguards to protect enduring constitutional premises of a democratic social order against the whims of a temporary or vindictive majority. "Strict scrutiny" is applied when "fundamental constitutional rights"[249] are affected, or where the majoritarian legislative process has unduly burdened classes of individuals that historically have suffered invidious discrimination or are particularly impotent in the political process. Such "suspect classifications"[250] traditionally sought to protect racial minorities against discriminatory legislation steeped in prejudice and developed in a political arena from which they have been excluded. In recent years, courts have become increasingly suspicious of legislation directed against other "semi-suspect" classes of individuals[251] who may also be "relegated to such a position of political powerlessness as to command extraordinary protection from the majoritarian political processes."[252]

While these groups — minorities, women, illegitimates, and aliens — are certainly victimized by the legislative process because of their identifiable status or certain immutable characteristics, the fundamental justification for judicial protection is their institutional political powerlessness.[253] The realities of the

---

249 *See* notes 272 to 297 and accompanying text *infra*.

250 "Suspect classifications" have included race, Reitman v. Mulkey, 387 U.S. 369 (1967), and alienage, Graham v. Richardson, 403 U.S. 365 (1971).

251 "Semi-suspect classifications" are generally considered to include gender, Stanton v. Stanton, 421 U.S. 7 (1975), and illegitimacy, Weber v. Aetna Casuality & Surety Co., 406 U.S. 164 (1972).

252 San Antonio Independent School Dist. v. Rodriguez, 411 U.S. 1, 28 (1973).

253 L. TRIBE, *supra* note 50, § 16-29, at 1077; *cf.* Sugarman v. Dougall, 413 U.S.

modern pluralistic political process are contrary to the traditional democratic assumption that the law-making process aggregates, and then balances, all individual citizens' concerns and ideals.[254] "[I]n a society whose legislative and administrative processes of value-formation and conflict-resolution resemble less the ancient ideal of the polis than the contemporary notion of pluralist compromise, any suggestion that bartered rules are necessarily expressions of substantive consensus seems almost impossible to maintain."[255] The basic strand of pluralist theory is that no independent "public interest" exists apart from the aggregate of private interest-group representation in the political process.[256] Taken to its extreme, this theory is antithetical to those enduring values of individual autonomy which the Constitution is designed to preserve.

The pluralist mode of political functioning penalizes unorganized classes of individuals. Since any group may claim to have "inadequate coalition formation and electoral representation,"[257] it is essential to determine whether its political powerlessness is structural; arising either from a history of pervasive prejudice or from clearly apparent socio-political barriers to interest-group formation.

2. The Political Powerlessness of Medical Malpractice Victims

The political powerlessness of the class of future medical malpractice victims is extreme and is functionally equivalent to that of other suspect and semi-suspect classes precisely because the members of this class share no common or immutable traits. Almost all citizens obtain medical treatment which may result in an injury from malpractice. In theory, the risk of being a future medical malpractice victim extends equally to all individuals and classes in society. Potential medical malpractice

---

634, 657 (1973) (Rehnquist, J., dissenting) ("It would hardly take extraordinary ingenuity for a lawyer to find 'insular and discrete' minorities at every turn in the road.")

254 *See generally* ARISTOTLE, THE POLITICS, BOOK III and BOOK VII (Oxford ed., E. Barker ed. and trans. 1974).

255 L. TRIBE, *supra* note 50, § 16-31, at 1091

256 *See generally* R. DAHL, PLURALIST DEMOCRACY IN THE UNITED STATES (1967); R. DAHL, WHO GOVERNS? (1961).

257 L. TRIBE, *supra* note 50, § 16-29, at 1077.

victims possess no readily identifiable personal characteristics that facilitate their political organization in opposition to legislation restricting malpractice recovery.

Without these external characteristics to enhance interest-group cohesiveness, organization would only occur if the individuals were to share a similar perception of their common danger. Yet, it seems highly unlikely that many individuals actively contemplate the relatively remote risk that they may become malpractice victims. The number of actual victims is not large enough to generate widespread public concern for personal safety, nor is notice of the restrictive legislation sufficiently prominent to draw the attention of individuals who may rationally assume that they continue to possess an effective judicial remedy.[258] There is no apparent impetus to trigger the intense concerns that galvanize individuals to coalesce into political organizations that participate actively in the legislative process. These characteristics justify treating future medical malpractice victims similarly to other politically powerless, semi-suspect classes who receive judicial protection through an enhanced scrutiny of legislation critically affecting their individual rights. While this analysis does not bind the legislature to perpetually preserve the common-law medical malpractice tort action, a substitute remedy arguably must be "reasonable" and "just."[259]

State courts examining medical malpractice compensation schemes have focused on the absence of any traditional suspect classes or fundamental rights.[260] Therefore, they have examined the statutory schemes' classification of medical malpractice victims[261] solely in terms of their rational relationship to the avowed legislative purpose.[262] Only in *Wright v. Central Du Page Hos-*

---

258 *See* Note, *California's Medical Injury Compensation Reform Act*, 52 S. CAL. L. REV. 829, 861 n.209, 937-38 (1979).

259 *See* notes 162 to 247 and accompanying text *supra*.

260 *E.g.*, Jones v. State Board of Medicine, 97 Idaho 859, 555 P.2d 399 (1976), *cert. denied*, 431 U.S. 914 (1977) (limits on recovery for medical malpractice do not affect a suspect class or infringe a fundamental right).

261 Separate equal protection concerns may be raised by the classification within the scheme between those injured plaintiffs seeking damage sums below the liability limit and those whose claims exceed the limitation. *See* American Bank and Trust v. Community Hosp., 104 Cal. App. 3d 219, 163 Cal. Rptr. 513, 518, 521 (1980)

262 *See* Arneson v. Olson, 270 N.W. 2d 125, 135 (N.D. 1978) ("We proceed to a determination of whether . . . there is a sufficiently close correspondence between statutory classification and legislative goals . . "); *see also* cases cited in notes 97 and

pital Association[263] did the state court analogize to the workmen's compensation schemes and consider whether there was a similar *quid pro quo* to protect individual rights. The general rationality focus, if used alone, gives excessive leeway to the legislature to enact statutes favoring physicians and insurance companies over unrepresented medical malpractice victims so long as pure arbitrariness is not found.[264]

Where the general public interest in having an effective medical malpractice remedy is so diffuse, it easily falls prey to the lobbying of highly-organized special interest groups. Political accountability is drastically skewed as the remote risk of malpractice injury for each individual produces a collective impotence that allows a politically powerful medical establishment to exploit and manipulate the political process. By persuading the legislature to shift the burden of liability to the medical malpractice victim — an arbitrary minority probably least able to bear the loss — the medical establishment has effectively "capitalized the benefits and socialized the risks" of medical treatment. The tragic impact of such action on the legitimacy of the legislative process is well-described by Mr. Justice Jackson. "[N]othing opens the door to arbitrary action so effectively as to allow . . . officials to pick and choose only a few to whom they will apply legislation and thus escape the political retribution that might be visited upon them if large numbers were affected."[265]

The intensity of political lobbying during the perceived malpractice crisis[266] by the medical profession and the medical malpractice insurance industry, and the absence of effective representation of the interests of potential victims,[267] has been

---

102 to 110 *supra; but cf.* Brown v. Merlo, 8 Cal. 3d 855, 506 P.2d 212, 106 Cal. Rptr. 388 (1973) (legislative classification abrogating injured automobile guest's common-law tort remedy violates equal protection).

263 63 Ill. 2d 313, 347 N.E.2d 736 (1976).

264 Redish, *supra* note 2, at 77; *contra,* Witherspoon, *supra* note 21, at 452-53.

265 Railway Express Agency, Inc. v. New York, 336 U.S. 106, 112-13 (1949) (Jackson, J., concurring).

266 *See* note 1 *supra.*

267 The organized bar, particularly the American Trial Lawyers Association, has participated in the medical malpractice legislative process. It may fairly be assumed that their interest centers more around the legislative impact on contingency fees, than on a deeply-held concern for individual liberties *per se*. The prospective malpractice victim, however, still was not adequately represented in the legislative process by this surrogate lobbying group.

strongly alleged.[268] Where the democratic political processes are so prone to distortion and abuse, reviewing courts should be wary of extending broad deference to asserted legislative policy judgments.[269]

The opportunity for abuse of individual rights in a medical malpractice compensation scheme is far more severe than in the workmen's compensation[270] or nuclear power[271] cases. The absence of a sufficiently salient personal threat results in little individual motivation towards organizing a political coalition of potential malpractice victims that could be an effective countervailing force in the legislative process. By contrast, workers are effectively represented in the political arena by their labor unions. Their voices were amply heard by legislators in the bartering and compromising process that shaped the various workmen's compensation schemes. Similarly, nuclear power opponents and those residing near nuclear plants who want better compensation benefits confront a readily apparent danger; as a result, they can coalesce into a political lobbying group capable of pushing for legislative revisions, or abrogation, of the Price-Anderson Act, or they can pressure the Nuclear Regulatory Commission to impose stricter licensing and operation safety requirements.

The qualitative nature of the relationships involved is also quite different. The employment relationship and workplace conditions have been radically transformed in the 20th century along with the transition in this area from private contractual law to public labor law. Energy development is similarly broadly governed by public law. Additionally, both the worker, by

---

It might also be suggested that the Blue Cross/Blue Shield health insurance groups may have an economic incentive to serve as *ad hoc* patients' representatives in the legislative process. To the extent that recovery limitations fail to provide full and necessary coverage, this primary insurer may be forced to cover additional health care costs. However, there has been no reported instance of active Blue Cross/Blue Shield lobbying on medical malpractice legislative schemes on behalf of the potential victims. One can only speculate as to the reasons for their non-participation, but the high degree of control of the Blue Cross/Blue Shield Boards of Directors and managements by physicians and hospital administrators may be pertinent. *See generally* S. LAW, BLUE CROSS — WHAT WENT WRONG? (2d ed. 1976)

268 *See* note 5 *supra*.

269 *See* Note, *California's Medical Injury Compensation Reform Act*, 52 S. CAL. L. REV. 829, 934-35 n.623 (1979).

270 *E.g.*, New York Central R. R. Co. v. White, 243 U.S. 188 (1917); Usery v. Turner Elkhorn Mining Co., 428 U.S. 1 (1976); Crowell v. Benson, 285 U.S. 22 (1932).

271 *E.g.*, *Duke Power*, 438 U.S. 59 (1978).

changing jobs, and the nuclear power plant neighbor, by relocating to a new residence, can ultimately choose to minimize their personal risk of injury. The physician-patient relationship, however, has remained essentially private. True, no matter whether the patient is receiving non-consensual emergency medical treatment or consensual elective surgery, he may be able to minimize his risks by choosing to avoid certain dangerous surgical procedures, but his ability to avoid malpractice is severely constrained.

When the legislative balancing process is unduly skewed by the structural inability of the burdened class to form active political coalitions, a court must be sensitive to its institutional role as a counter-majoritarian monitor of legislative legitimacy. The political powerlessness of future medical malpractice victims arguably justifies their status as a semi-suspect class entitled to judicial protection against majoritarian subjugation of individual rights.

B. *A "Quasi-Fundamental" Right to Bodily Freedom From Uncompensated Private Assualt*

1. Bodily Integrity

Strict judicial scrutiny is also triggered when legislation impairs those fundamental rights "explicitly or implicitly guaranteed by the Constitution."[272] The Supreme Court has included among these rights those central to political participation[273] and those basic to the exercise of intimate personal freedoms.[274] In addition, some commentators have described judicial trends postulating quasi-fundamental rights which trigger intermediate forms of scrutiny[275] when liberties and benefits comprising the basic necessities of life are at stake.[276]

---

272 San Antonio Independent School Dist. v. Rodriguez, 411 U.S. 1, 33-34 (1973).

273 *See, e.g.*, Boddie v. Connecticut, 401 U.S. 371 (1971) (litigating); Reynolds v. Sims, 377 U.S. 533 (1964) (voting).

274 *See, e.g.*, Zablocki v. Redhail, 434 U.S. 374 (1978) (marriage); Shapiro v. Thompson, 394 U.S. 618 (1969) (travel); Griswold v. Connecticut, 381 U.S. 479 (1965) (privacy); Skinner v. Oklahoma, 316 U.S. 535 (1942) (procreation).

275 L. TRIBE, *supra* note 50, § 15-9, at 919-20, § 16-31, at 1089-92.

276 *See, e.g.*, Turner v. Department of Employment Security of Utah, 423 U.S. 44

The Due Process Clause of the Fourteenth Amendment protects the individual against deprivation of those liberties "long recognized at common law as essential to the orderly pursuit of happiness by free men."[277]

> Among the historic liberties so protected was a right to be free from, and to obtain judicial relief for, unjustified intrusions on personal security.
>
> While the contours of this historic liberty interest in the context of our federal system of government have not been defined precisely, they have always been thought to encompass freedom from bodily restraint and punishment.[278]

Surely, no personal liberty can be more fundamental to the individual than freedom from unconsented assaults. Whatever the precise contours of the social contract undergirding the relationship between the individual and the democratic state,[279] it seems axiomatic that a central element is the expectation that a collectivity will better safeguard the individual against harm to his bodily integrity and personal security.

In general, judicially-enforced rules limiting legislative power are undesirable because of their anti-democratic tendencies and restraints upon flexible and expedient legislative action. The Framers of the Constitution, and the various United States Supreme Courts, have deemed intervention desirable, however, to protect exceptional enduring interests which may be undervalued by the legislative process in view of momentary expedience. The Framers may have considered interests in free speech and expression,[280] and freedom of association,[281] as particularly susceptible to legislative derogation, and, therefore, in need of explicit Constitutional protection. However, the Fra-

---

(1975) (employment); Memorial Hosp. v. Maricopa County, 415 U.S. 250 (1974) (nonemergency medical treatment); United States Dep't of Agriculture v. Moreno, 413 U.S. 528 (1973) (food stamps).

277 Meyer v. Nebraska, 262 U.S. 390, 399 (1923).

278 Ingraham v. Wright, 430 U.S. 651, 673-74 (1977) (footnote omitted); *see* Roe v. Wade, 410 U.S. 113 (1973); Rochin v. California 342 U.S. 165 (1952); Skinner v. Oklahoma, 316 U.S. 535 (1942).

279 *See* notes 150 to 151 *supra.*

280 *See, e.g.,* Linmark Associates v. Town of Willingboro, 431 U.S. 85 (1977) (municipal ordinance that prohibits individual's posting of house "For Sale" signs violates the First Amendment).

281 NAACP v. Alabama, 357 U.S. 449 (1958) (state's attempt to procure private organizations' membership lists violates the First Amendment).

mers may have quite reasonably assumed that legislatures would never undervalue or act adversely towards the natural rights of personal security and bodily integrity. These values were presumed to enjoy preeminent protection in the political process.[282]

The philosophical breakdown of a "natural rights" basis for modern society, though, left a curious jurisprudential void. A static Constitution would protect freedom of association against hostile or impulsive laws, but not protect personal autonomy. The Supreme Court logically reformed this incongruity by accommodating rights central to personal control and bodily autonomy. Rights of procreation[283] and privacy[284] were deemed to be implicitly protected by the Constitution. The rights to personal security and bodily integrity, and corresponding rights to relief from invasions thereof, are logically encompassed by this constitutional interpolation.[285]

In obtaining quality medical treatment,[286] the individual im-

---

282 This argument owes much to the writings of Professor Frank Michelman. *See* Michelman, *The Supreme Court and Litigation Access Fees: The Right to Protect One's Rights — Part I,* 1973 DUKE L.J. 1153, 1206-15 (1973).

283 Skinner v. Oklahoma, 316 U.S. 535 (1942).

284 *E.g.,* Roe v. Wade, 410 U.S. 113 (1973); Griswold v. Connecticut, 381 U.S. 479 (1965).

285 *See* Ingraham v. Wright, 430 U.S. 651 (1977).

286 Some commentators have interpreted the Supreme Court's "right to travel" cases as creating a substantive rights umbrella promoting a hierarchy of constitutionally protected interests. *See, e.g.,* TRIBE, *supra* note 50, § 16-8, at 1003-05. In a series of cases, durational residence requirements for receipt of certain important public benefits were invalidated by the Supreme Court. It has been suggested that the Court's assessment of the social value of the interests affected was the determinative factor, rather than the "right to travel." *Compare* Memorial Hosp. v. Maricopa County, 415 U.S. 250 (1974) (durational residence requirement affecting non-emergency medical treatment is unconstitutional) *and* Dunn v. Blumstein, 405 U.S. 330 (1972) (residence requirement affecting voting held unconstitutional); Shapiro v. Thompson, 394 U.S. 618 (1969) (residence requirement affecting welfare benefits held unconstitutional) *with* Sosna v. Iowa, 419 U.S. 393 (1975) (residence requirement affecting divorce petition held constitutional) *and* Starns v. Malkerson, 401 U.S. 985 (1971), *aff'g without opinion,* 326 F. Supp. 234 (D. Minn. 1970) (residence requirement affecting college tuition held constitutional). *But cf.* Vlandis v. Kline, 412 U.S. 441 (1973) (overturning as violative of due process the state's use of an irrebutable presumption of students' non-residence for college tuition purposes).

This argument suggests that under some circumstances medical treatment may be a fundamental benefit to the individual that warrants enhanced judicial protection. In Memorial Hosp. v. Maricopa County, the Supreme Court held unconstitutional a state statute which required fulfillment of a residency requirement before an indigent person could receive non-emergency medical treatment at county expense. The Court ruled that the medical treatment — a "necessity of life" — could not be conditioned upon an abrogation of the fundamental right to travel. The court gave considerable weight

plicitly consents not only to intrusion upon his body, but also to the inherent risks of treatment.[287] Such consent, however, does not extend to inept or shoddy treatment. Insofar as inappropriate treatment exceeds the limits of consent, malpractice constitutes private assault upon the individual. Historically, victims have relied upon their common-law medical malpractice tort action[288] as a sufficient safeguard against this invasion of bodily integrity.

2. The Correlative Tort Remedy to Protect Individual Liberty

The aggregation of private malpractice tort actions creates a scheme of economic incentives that help to shape physician conduct. As economic actors, physicians may be encouraged to operate to the margins of the balance implicit in this system.[289] Recent studies have generally concluded that the threat of high medical malpractice tort damage awards provides a deterrent to negligent and irresponsible physician conduct straying beyond these margins.[290] As discussed earlier, this deterrence factor more significantly affects medical treatment than it would

---

to the penalty inflicted by the restriction. "[Diseases] if untreated for a year, may become all but irreversible paths to pain, disability and even loss of life." 415 U.S. at 261; *but cf.* Maher v. Roe, 432 U.S. 464 (1977) (no fundamental right to choose abortion as medical treatment).

287 Though obtaining emergency medical treatment may be formally non-consensual, it can be reasonably assumed that the grievously-ill individual would be willing to consent to treatment.

288 *See* Note, *Medical Malpractice — Constitutionality of Limits on Liability*, 78 W. VA. L. REV., 381, 389 (1976).

289 The traditional conception of tort law has been that of shifting losses from the injured plaintiff to the tortfeasor. W. PROSSER, TORTS 6 (4th ed. 1971). Many contemporary commentators, however, have approached tort law from the perspective of economic efficiency. This approach emphasizes the maximization of net social wealth through deterrence of non-optimal actions, rather than compensation of the victim. *See* G. CALABRESI, THE COST OF ACCIDENTS (1970); Calabresi and Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089 (1972); Posner, A *Theory of Negligence*, 1 J. LEGAL STUD. 29 (1972).

290 *See, e.g*, W. SCHWARTZ & N. KOMESAR, DOCTORS, DAMAGES AND DETERRENCE: AN ECONOMIC VIEW OF MEDICAL MALPRACTICE (The Rand Corp. 1978); AMERICAN BAR ASSOCIATION REPORT OF THE COMMISSION ON MEDICAL PROFESSIONAL LIABILITY, at 57 (1977); Shore, *Civil Liability Expands in State But Insurance Costs Keeps Up Pace*, L. A. Daily J., Sept. 11, 1978, at 9, col. 3 (describing Rand Corporation study).

The full impact of economic deterrence is diluted by the haphazard manner in which random claims are brought, high transaction costs, and inadequate systemic feedback of claims information to hospital administrators and medical insurers. *See* Chittenden, *supra* note 162, at 921 (1977) (discussing American Bar Association Commission report).

conduct in the workplace, nuclear power, or automobile no-fault situations.[291]

Where a tortious private assault has occurred, the state's political role is legitimated by its temporizing ability to facilitate, through the judicial system, an individual's full recovery of compensation for his injuries, and to impose such further punitive damages as necessary to deter future assaults.[292] The democratic state's continued authority is based upon maintaining this delicate equilibrium of correlative common-law tort remedies and property rights to assure protection of fundamental individual liberties.

State laws that impose limitations on recovery for medical malpractice, and require additional costly and time-consuming review panel procedures that may deter the assertion of legit-

---

291 *See* notes 225 to 228 and accompanying text *supra.*

292 The Ninth Circuit recently reversed a district court decision that had found the California rule prohibiting the award of punitive damages in wrongful death cases to violate the Equal Protection Clause of the Forteenth Amendment. The district court had concluded that "punitive damages emanate from the public's right to be free from oppressive, fraudulent or malicious conduct. . . . That protection is personal and substantially within the protections accorded fundamental rights found in . . . the United States Constitution." In re Paris Air Crash of March 3, 1974, 427 F. Supp. 701, 708 (C.D. Cal. 1977), *reversed,* 622 F.2d 1315 (9th Cir. 1980). The Court of Appeals applied a deferential rational relationship test in reviewing the limitation on punitive damages in the wrongful death situation becuase:

> [t]he plaintiffs, appellees here, are not without a remedy for personal losses sustained in fact. . . . Plaintiffs seek in addition, and as a matter of constitutional right, the opportunity to act as private attorneys general to effect the deterrence and retribution functions. . . . So far is this opportunity from being a fundamental personal right that it is an interest not truly personal in nature at all. . . . This distinction between compensatory and punitive damages explains why a court will in general be less likely to hold unconstitutional a limit on the latter. . . . The case is therefore quite different from one in which the plaintiff has suffered a serious loss, be it a personal injury . . . and is left with no effective remedy and thus no compensation for a grievious deprivation. In such cases, the courts have adopted remedial innovations through common-law power, or through application of equal protection principles. . . . In this case, as we have noted, fair compensation is provided. . . . This point also serves to distinguish the recent decision of the California Court of Appeal in *American Bank & Trust Co. v. Community Hospital,* 104 Cal App. 3d 219, 163 Cal. Rptr. 513 (1980), in which the court invalidated a medical malpractice recovery limitation.

Id. at 1319-20 & n.7.

The appellate court's specific exception of the medical malpractice legislation reflects its concerns that (1) fair compensation was not provided to the medical malpractice victim, *id.* at 1320 n.7; (2) "the effect of the law was somewhat discriminatory, in that one particular class of tortfeasors — those guilty of medical malpractice — was benefited at the expense of a narrower class of plaintiffs — those injured by malpractice, *id.;* (3) medical malpractice injuries are truly personal in nature, *id.* at 1320; and (4) the medical malpractice cause of action is grounded in the common-law, as contrasted with the statutory wrongful death action, *id.* at 1320 n.9.

imate claims, shift the economic incentives governing quality medical treatment. Where legislation furthers private assaults on bodily integrity, intrusiveness into individual liberties is at its most extreme.[293] The individual's common-law tort remedy is the sole countervailing force to assure that an alleged public purpose is not wrongfully extended to serve merely as a shield for the private infliction of harm. It acts as a residual individual right that tempers a balance between individual liberty and the reach of intrusive state action. When the state has disrupted this equilibrium by partially stripping the individual of his protective remedies, the legislation impinges on a quasi-fundamental right of bodily integrity, and the reviewing court should scrutinize the legislation more carefully than a mere rationality standard would suggest.

Legislative schemes restricting medical malpractice liability intrude much more deeply into the realm of intimate, private affairs than similar schemes in the areas of workmen's compensation or nuclear power development.[294] In those areas of public law, the imposition of strict liability assumes a sense of social injury, not private harm. The medical malpractice compensation schemes, however, depart from the strict liability model and retain the basic negligence standards typically applied to private conduct.

Imposing a stricter standard of judicial review over violations of this asserted quasi-fundamental right to bodily integrity against uncompensated private assault may be innovative, but hardly remarkable.[295] Professor Tribe describes such an expan-

293 *But cf.* Martinez v. State of California, 100 S. Ct. 553 (1980) (state statute providing a defense to potential state tort law liability, though having an attenuated incremental impact in causing death, is not state action depriving plaintiff of due process).

294 *See* notes 270 to 271 and accompanying text *supra.*

295 To some extent, the argument proposed herein — that an apparent quasi-fundamental right to bodily integrity should trigger enhanced judicial review of state tort remedial schemes — is at odds with the Supreme Court's expressed desire to avoid superimposing "a body of general federal tort law" upon systems administered by the states. Paul v. Davis, 424 U.S. 693, 701 (1976). Subsequently, in Ingraham v. Wright, 430 U.S. 651 (1977), the Supreme Court held that students, who had been paddled by a teacher, had been deprived of a protected constitutional liberty interest in bodily integrity. The five-man majority held that there was no due process violation, however, because plaintiffs possessed an adequate state tort law remedy to vindicate their rights. The four dissenting justices concluded that the state tort law remedies were wholly insufficient, and therefore determined that constitutional due process was violated.

In both *Paul* and *Ingraham,* plaintiffs brought § 1983 actions against state tortfeasors. In effect, plaintiffs were explicitly attempting to transform a garden-variety, common-

sion following from *Skinner v. Oklahoma*:[296]

> [E]qual protection analysis demands strict scrutiny not only of classifications that penalize rights already established as fundamental for reasons unrelated to equality, and of classifications that unequally distribute access to established rights whose very fundamentality resides in a norm of equal availability, *but also of classifications unequally distributing access to choices that ought to be placed beyond government's reach — and in that sense be deemed fundamental — because, in government's hands, control over those choices would pose too great a danger of majoritarian oppression or enduring subjugation.*[297]

V. THE "QUID PRO QUO" AS A PREFERRED TECHNIQUE FOR "INTERMEDIATE SCRUTINY"

A. *"Intermediate Scrutiny" Techniques*

Legislation implicating semi-suspect classes, quasi-fundamental rights, or both, triggers "intermediate scrutiny".[298] Al-

---

law state tort action into a federalized tort action. *Cf.* Martinez v. State of California, 100 S. Ct. 553 (1980) (state interest in fashioning own tort law rules paramount to federal interest except to protect individual citizen from arbitrary state action).

When constitutional rights are predicated on the sufficiency of state remedial scheme (*viz., Ingraham*), there is an implied federal remedy to prevent the state from parochially construing its own laws to nullify federal constitutional rights. *See* Anderson v. Brand, 303 U.S. 95 (1938); Iowa-Des Moines Nat'l Bank v. Bennett, 284 U.S. 239 (1931), Ward v. Love County, 253 U.S. 17 (1920); General Oil v. Crain, 209 U.S. 211 (1908).

In the past, the Supreme Court has either forced the state to provide the necessary remedy, *see Ward,* 253 U.S. 17; *General Oil,* 209 U.S. 21, or has overridden the state's substantive determination of its own laws, *see Indiana ex rel. Anderson,* 303 U.S. 95; *Iowa-Des Moines Nat'l Bank,* 284 U.S. 239.

Despite the Court's avowed efforts to avoid creating a federalized tort scheme, *Paul,* 424 U.S. at 701, the *Ingraham* decision indicates that § 1983 creates a residual federal common-law tort remedy drawing substance from the Fourteenth Amendment when the state tort remedy, as in the medical malpractice case, is insufficient.

Such a drastic remedy can be avoided by the Court's utilization of the *quid pro quo* standard to review medical malpractice compensation schemes. The state legislature is free to alter the common-law medical malpractice tort action; the reviewing court employs the due process safeguards to assure only that the state has provided a reasonable and just substitute remedy. There is no need to impose an inflexible federal standard or preclude states from legislating in this area. While the *quid pro quo* requirement outlines the parameters of the substitute, the state is free to design the specifics of its own scheme. Essentially, the reviewing court has established a dialogue in which its appropriate role is supervisory, not preemptive.

296 316 U.S. 535 (1942).

297 L. TRIBE, *supra* note 50, § 16-12, at 1011 (emphasis added).

298 *See* notes 50 to 58 and accompanying text *supra.*

though the precise parameters of this elevated standard of review are ill-defined, some general patterns are apparent. Early commentators described this standard as directed solely to the closeness of the relationship between legislative means and ends.[299] Though trying to escape the strict deference of minimum rationality review, courts appeared chary of excessive judicial intervention in examining legislative ends. The futility of retaining an artificial means-ends dichotomy while pressing for enhanced intermediate scrutiny is readily apparent.[300] Since it was ultimately unable to sidestep the review of legislative ends,[301] the Supreme Court has treaded cautiously, structuring narrow remedies designed to cure constitutional imperfections rather than wholly invalidating flawed legislation.[302]

Professor Tribe has identified five techniques of intermediate review applied by the courts:[303] (1) assessing the importance of the legislative objective;[304] (2) demanding a close fit between legislative means and ends;[305] (3) requiring current articulation of the underlying legislative rationale;[306] (4) limiting "afterthought" justifications;[307] and (5) permitting rebuttal of "conclusive presumptions" in individual cases even if the entire scheme is not invalidated.[308]

B. *The Inadequacy of Common "Intermediate Scrutiny" Techniques in Reviewing Medical Malpractice Compensation Schemes*

The first four of these techniques assess the general legitimacy of the legislatively-enacted social program, but avoid examining

299 Gunther, note 43 *supra*.
300 *See* Redish, *supra* note 2, at 777.
301 *See, e.g.*, Eisenstadt v. Baird, 405 U.S. 438, 442-43, 447-52 (1972); Reed v. Reed, 404 U.S. 71, 76 (1971).
302 *See, e.g.*, Cleveland Bd. of Educ. v. LaFleur, 414 U.S. 632 (1974).
303 TRIBE, *supra* note 50, § 16-30, at 1082-89.
304 *See, e.g.*, Reed v. Reed, 404 U.S. 71 (1971); TRIBE, *supra* note 50, § 16-30, at 1082-83.
305 *See, e.g.*, Craig v. Boren, 428 U.S. 190 (1976); TRIBE, *supra* note 50, § 16-30, at 1083.
306 *See, e.g.*, Cleveland Bd. of Educ. v. LaFleur, 414 U.S. 632 (1974), TRIBE, *supra* note 50, § 16-30, at 1083-85
307 *See, e.g.*, Eisenstadt v. Baird, 405 U.S. 438 (1972); TRIBE, *supra* note 50, § 16-30, at 1085-88.
308 *See, e.g.*, Stanley v. Illinois, 405 U.S. 645 (1972); TRIBE, supra note 50, § 16-30, at 1088-89.

the particular impacts on individual rights. This myopic focus on the validity of the social program provides the individual medical malpractice victim with scant more protection than that offered by the *Duke Power* court sans the *quid pro quo* analysis. State legislation restricting private medical malpractice tort actions purports to promote important legislative objectives of reducing health care costs. Even if the legislation does little to actually reduce health costs, but rather redistributes the cost burden to medical malpractice tort victims,[309] the legislative objective of assuring the availability of malpractice insurance may be achieved. Though the closeness of the fit is statistically suspect,[310] the means of recovery limitations and screening panels are clearly directed to promoting the asserted legislative ends. The legislature's rationale is currently articulated, rather than antiquated, and so long as the legislative findings are carefully stated, no creative "afterthought" justifications are necessary.

One technique frequently employed to protect individual rights forbids the use of legislative irrebuttable presumptions. A hearing is required to allow the affected individual an opportunity to rebut the personal application of conclusive presumptions contained in legislation.[311] Though some commentators have suggested that this doctrine is simply aimed at impermissibly overinclusive legislation,[312] others argue that this restrained form of judicial intervention sensitively protects certain individual liberties against inappropriate legislative impairments without entirely invalidating the legislation.[313]

It may be argued that a state medical malpractice remedial scheme may so limit the plaintiff's potential dollar recovery,[314] and so burden his ability to assert his claim,[315] that the legislation

309 *See* American Bank and Trust v. Community Hosp., 104 Cal. App. 3d 219, 163 Cal. Rptr. 513, 521 (1980).

310 *See* Arneson v. Olson, 270 N.W.2d 125, 135-36 (N.D. 1978); Jones v. State Bd. of Medicine, 97 Idaho 859, 871-72, 555 P.2d 399, 411-12 (1976), *cert. denied,* 431 U.S. 914 (1977). *See* note 342 *infra.*

311 *See, e.g.,* Cleveland Bd. of Educ. v. LaFleur, 414 U.S. 632, 645-49 (1974).

312 *See, e.g.,* Note, *The Irrebuttable Presumption Doctrine in the Supreme Court,* 87 HARV. L. REV. 1534 (1974).

313 L. TRIBE, *supra* note 50, §16-32, at 1095.

314 For example, the Idaho Hospital-Medical Liability Act creates a damage ceiling of only $150,000 per malpractice claim. IDAHO CODE § 39-4204 (1977).

315 Out of 2,466 claims submitted, from April, 1976, to August, 1979, to the condition precedent Pennsylvania Arbitration Panels for Health Care, *only nine* reached the hearing stage. *See* Edelson v. Soricelli, 610 F.2d 131, 135-36 (3rd Cir. 1979).

can be deemed to conclusively presume that all claims are non-meritorious or excessive.[316] Therefore, the legislation is impermissibly applied to a grieviously-injured malpractice victim asserting a legitimate claim. Under traditional doctrine, however, the sole remedy would be a hearing to determine whether the victim should be relieved of the statutory burdens. The impact would be directly contradictory to a central legislative purpose: to reduce health care costs, restrictive recovery limits must reduce dollar judgments which would otherwise be awarded. "Ironically, the higher the limit set, the lower the likelihood that the limit will accomplish its goal."[317] Furthermore, an additional hearing would simply create another burdensome layer of pre-trial delay and cost to the medical malpractice victim — the precise gravamen of his complaint.

C. *Jurisprudential Constraints*

Since the prevalent remedial tools available under intermediate scrutiny appear inapplicable for critical adjudication of medical malpractice compensation schemes enacted by many states, skeptical reviewing courts have tended to employ an *ad hoc* enhanced scrutiny to analyze the general rationality of the legislation.[318] Essentially, this form of review tends to graft judicial conceptions of irrationality and arbitrariness onto an examination of the assumptions underlying the social program. This mode not only perpetrates "judicial legislating" of social policy, but it may particularly strain the exercise of judicial competence. Furthermore, by focusing on the sagacity of the social program, the courts often fail to fulfill their more appropriate function of redressing individual rights eroded by majoritarian excesses.

Some state courts have independently analyzed the evidence purportedly supporting the legislature's assumption that a medical malpractice insurance crisis exists.[319] Where courts have

316 *See* Witherspoon, *supra* note 21, at 440-41.

317 Redish, *supra* note 2, at 781 n.134.

318 *See* Note, *supra* note 2, at 907, *passim* (endorsing judicial approach of non-strict, but elevated, scrutiny of medical malpractice legislation to offer a "genuine review of legislation").

319 *See* Jones v. State Bd. of Medicine, 97 Idaho 859, 869-70, 872-76, 555 P.2d 399, 409-10, 412-16 (1976), *cert. denied,* 431 U.S. 914 (1977).

rejected the veracity of the premise underlying the medical malpractice schemes, the structural rationality has tottered.

Unfortunately, as many commentators have forcefully argued,[320] this is the precise form of intrusion into the legislative process that courts may be least able to handle competently. That a purported crisis may be illusory, temporary, or cunningly engineered is unquestioned,[321] but an effective determination of its legitimacy necessarily entails a detailed statistical analysis of claims histories, and of local and national medical malpractice insurance rate structures. A legislature, blessed with direct access to economists, insurance experts, and health care specialists participating in the continuing dialogue of the legislative process, has important resources available in making its programmatic decision — even if it is normatively incorrect due to extraneous influences.[322] A court, which has been presented solely with static legal briefs and a one-shot oral argument, is far less competent to defensibly determine the legitimacy of this particular crisis.

Though a court's blind deference[323] to a legislature's bare assertion of what may be a particularly specious "crisis"[324] is troublesome, in all but the most patently egregious cases, this seems an ill-advised occasion for a court to act as a "super-legislature," adjudicating the wisdom of social and economic policy choices.[325] While the prevalent intermediate scrutiny tech-

---

320 *E.g.*, Linde, *Due Process of Lawmaking*, 55 NEB. L. REV. 197 (1976).

321 *See* Witherspoon, *supra* note 21, at 445-52 (1979).

322 Some medical malpractice schemes favoring physicians and other medical treatment providers have later been found to violate state constitutional prohibitions against special legislation. *E.g.*, Wright v. Central Du Page Hosp. Ass'n, 63 Ill. 2d 313, 329, 347 N.E.2d 736, 743 (1976).

323 *See, e.g.*, McGuffey v. Hall, 557 S.W. 2d 401, 413 (Ky. 1977).

324 *See* Witherspoon, *supra* note 21, at 445-52.

325 General judicial deference to a legislature's economic analysis, however, need not imply that the court should abandon its institutional role of protecting individuals' constitutional rights against majoritarian usurpation.

As the combination of social and economic forces that once led to the purported medical malpractice crisis no longer exists, the legislature's findings may seem suspiciously antiquated, A legislature, which is preoccupied with the newest emergency, may well be institutionally inert to re-evaluate outmoded medical malpractice legislation. A court concerned with protecting individual rights has several tools at its disposal to force the legislature to re-examine its prior assessment of the determination of the appropriate balance inherent in the medical malpractice compensation scheme.

The court may find that the scheme has operated so inefficiently that it deprives the individual of constitutional rights. *E.g.*, Aldana v. Holub, 381 So. 2d 231, 236-38 (Fla. 1980) (Florida Medical Malpractice Reform Act violates the due process clauses of the

niques may provide inefficient tools for protecting individual liberties sacrificed by medical malpractice compensation schemes, the tendency of skeptical courts to utilize stricter review provides added justification for imposition of a stringent "reasonably just substitute" requirement.

D. *The "Quid Pro Quo" As A Jurisprudentially Preferable Review Technique*

Although the Supreme Court has not expressly required that a reasonably just substitute be present in a statute abrogating common-law tort remedies, the discussions in *Duke Power* and *New York Central R. R. Co.* have strongly endorsed substitute statutory benefits. The rationales of these cases logically dictate that this *quid pro quo* should be present even to satisfy the minimum rationality standard of review. Because the semi-suspect class and quasi-fundamental rights threatened by medical malpractice compensation schemes support a stricter review standard, the policy arguments supporting inclusion of a reasonably just substitute are especially applicable.

Whereas the common intermediate scrutiny techniques may strain the court's competence and inject it into economic policy formulation, the *quid pro quo* analysis readily comports with the traditional judicial function of assessing the relative values of legal benefits and burdens affecting the individual. This principled form of judicial balancing fits comfortably within the court's intellectual and institutional expertise.

The flexible *quid pro quo* standard enables the court to examine whether the medical malpractice compensation scheme structurally derogates individual liberties.[326] Not only does this

---

United States and Florida Constitutions, and the Florida constitutional right of access to courts), *overruling* Carter v. Sparkman, 335 So. 2d 802 (Fla. 1976) (Florida Medical Malpractice Reform Act "reaches the outer limits of constitutional tolerance"). *See* notes 137 to 146 and accompanying text *supra*.

The court may assess, with scrutiny techniques — intermediate or otherwise — the contemporaneous validity of the importance of the prior legislative objective of ensuring the availability of medical malpractice insurance, or it may require the current articulation of a rational basis for such sweeping legislation. In essence, this course amounts to the "legislative remand" suggested by some constitutional scholars. *See, e.g.*, A. BICKEL, THE LEAST DANGEROUS BRANCH, 111-98 (1962).

326 In Arneson v. Olson, 270 N.W.2d 125 (N.D. 1978), the North Dakota Supreme

programmatic review provide more meaningful protection to the individual, but it allows the legislature more freedom to shape and fine-tune its regime innovatively.[327] The legislature is not barred from modifying common-law tort remedies to accommodate the exigencies of contemporary society. Judicial review enforces due process safeguards only to assure that the state has provided a reasonable and just substitute.[328]

The *quid pro quo* evaluation does not preclude states from legislating in an area, but simply prescribes the structural conditions within which the state can freely design the particular characteristics of its scheme. A legitimate dialogue is established between the judiciary and the legislature in assessing the dimensions of an appropriate reasonably just substitute. The reviewing court's role is supervisory, rather than presumptive, in preventing important individual rights from being wholly sacrificed to the popular will.

VI. THE "QUID PRO QUO" AS APPLIED TO MEDICAL MALPRACTICE COMPENSATION SCHEMES RESTRICTING COMMON-LAW TORT ACTIONS

In substance, if not form, a few of the state courts that have invalidated medical malpractice compensation schemes have responded to *quid pro quo* principles.[329] While the *Arneson v.*

Court refused to declare that a *quid pro quo* was formally required. *Id.* at 135. Nonetheless, the court noted that any limitation on a pre-existing right may not arbitrarily be imposed. The court then analyzed the medical malpractice compensation scheme by focusing on its individual components and determined that the aggregate of the burdens, without accompanying benefits, resulted in a constitutionally infirm structural scheme. *Id.* at 137.

327 The *quid pro quo* review test might appear more difficult to apply if the legislature attempted to abrogate or alter the individual's common-law tort remedy on an incremental basis, rather than by a single, packaged scheme. The individual, though, has less to fear from incremental changes than from sweeping structural changes. First, it is more likely that the medical, hospital, and insurance lobbies will coalesce and prioritize this interest for a single legislative session's one-shot scheme, than for incremental changes in a series of legislative sessions. Second, a powerful interest-group coalition is better able to capitalize on a temporary "crisis" to promote a broad, precipitous legislative scheme. Third, the sum of incremental changes is less likely to have as great an effect as is a broad legislative scheme because the legislature will be able to observe the impact of each small step before enacting an unnecessary further step.

328 *See also* note 295 *supra*.

329 *See* Arneson v. Olson, 270 N.W.2d 125 (N.D. 1978); Wright v. Central Du Page Hosp. Ass'n, 63 Ill. 2d 313, 347 N.E.2d 736 (1976).

Olson[330] court specifically analyzed the constitutional validity of each component of the legislative scheme, it also addressed the structural burdens, absent any corresponding benefits, imposed on malpractice victims. The court concluded that "the cumulative effect [of the statutory burdens] violate[s] the rights of medical patients in this state to due process of law."[331]

The workmen's compensation and nuclear accident compensation schemes provide guidance as to the elements of a constitutionally acceptable reasonably just substitute for the private common-law tort remedy. Countervailing sacrifices are demanded from each party. In exchange for the limitations on the workmen's or the nuclear victim's remedy, a strict liability system guarantees a recovery by superseding the plaintiff's prior obligation to prove negligent conduct.[332] Furthermore, the administrative claims procedure is presumed to significantly benefit the plaintiff by providing a prompt and efficient mechanism for damage recovery, thereby obviating the costs and long delays of private litigation.

These elements of reciprocity are totally absent in the medical malpractice compensation schemes. Though some respected commentators have suggested that a form of strict liability be systemically applied to medical malpractice injury compensation,[333] the present legislative schemes universally have retained the requirement that the plaintiff establish negligence. This "fault" principle is in accordance with the general presumption that medical malpractice injuries, unlike those occurring in the workplace or nuclear accident setting, are preventable and, therefore, amenable to deterrence forces.[334] Yet these medical

---

330 270 N.W.2d 125. *See* notes 115 to 124 and accompanying text *supra*.

331 *Id.* at 137.

332 Similarly, automobile no-fault statutes assure immediate and certain compensation for both parties without regard to fault.

333 The "designable compensable event" proposal, which would pre-define compensable outcomes by reference to fixed criteria, has received considerable attention. This approach steers a middle ground between negligence and strict liability compensation systems. *See* Havighurst, *'Medical Adversity Insurance' — Has Its Time Come?*, 1975 DUKE L. J. 1233 (1975); Havighurst & Tancredi, *Medical Adversity Insurance — A No-Fault Approach to Medical Malpractice and Quality Assurance*, 613 INS. L. J. 69 (1974); *see also* Chittenden, note 162 *supra*. For an excellent discussion of alternative systemic techniques for handling medical malpractice claims, *see A Center Report: A Discussion of Alternative Compensation and Quality Control Systems*, THE CENTER MAGAZINE, July/Aug., 1975, at 25-56.

334 *See* note 290 *supra;* and notes 225 to 227 and accompanying text *supra*.

malpractice compensation schemes functionally deprive the plaintiff-victim of his common-law tort remedy for negligent injury by imposing burdensome pre-trial screening panels, recovery limitations, abbreviated statutes of limitations, and other constraints. The schemes have combined the sacrifices of the strict liability system with the difficulties of proving negligence. The collective burdens are imposed on the plaintiff without the corresponding benefits.

Medical malpractice screening panels impose an added layer of tort litigation on plaintiffs. Early defenders of the panels suggested that the process might screen out non-meritorious claims and encourage settlements, thereby promoting a quicker and more efficient disposition of medical malpractice claims than at the purely adversarial trial.[335] Presumably, "[a]s a result of savings in litigation, insurance costs would be steadied or lowered."[336] At least some courts apparently believed that the actual burden would be slight and analogized the panels to "pre-trial settlement conferences"[337] or yet another "hearsay exception."[338] The Supreme Court has historically tolerated such limited pre-trial burdens.[339]

Many courts and commentators, however, have expressed considerable suspicion towards this attempt to transform a substantial burden into an innocuous benefit.[340] In many states, the inefficient operation of the panels has resulted in long delays preceding adjudication.[341] No reported studies have indicated that panels actually have deterred the litigation of non-meritorious claims. Indeed, the scanty evidence tends to indicate that the filing of meritorious claims has been deterred, and that

---

335 *See, e.g.*, Note, *supra* note 38, at 323.

336 *Id.* at 323.

337 Carter v. Sparkman, 335 So. 2d 802, 807 (1976), *cert. denied*, 429 U.S. 1041 (1977).

338 Comiskey v. Arlen, 55 App. Div. 2d 304, 309, 390 N.Y.S.2d 122, 126 (1976).

339 *See, e.g.*, Cohen v. Beneficial Indus. Loan Corp., 337 U.S. 541 (1949) (in a shareholders' derivative suit brought by less than five percent of the shareholders, a litigation bond may be required); Jones v. Union Guano Co., 264 U.S. 171 (1924) (plaintiff may be required to submit fertilizer to state agency for chemical analysis before suing a fertilizer supplier for crop damage).

340 *See* Aldana v. Holub, 381 So. 2d 231 (Fla. 1980); Carter v. Sparkman, 335 So. 2d 802, 807 (Fla. 1976), *cert. denied*, 429 U.S. 1041 (1977); Note, *The Constitutional Considerations of Medical Malpractice Screening Panels*, 27 AM. U. L. REV. 161, 167-77 (1977).

341 *See* notes 130 to 133 and accompanying text *supra*.

the added front-end time and costs of the screening panel are forcing plaintiffs toward early settlement of their claims.[342]

The harsh pre-trial litigation burden falls disproportionately[343] on the medical malpractice victim. To the extent that the panels fail to deter plaintiffs from proceeding to trial, the net effect is that the system has delayed justice and doubled its costs. Furthermore, it is difficult to fathom how the imposition of screening panels constitutes a benefit facilitating the medical malpractice victims' reasonably just substitute. All but the most solicitous interpretations recognize the panel as an added burden upon the victim.

The impact of recovery limitations and abbreviated statutes of limitation for medical malpractice claims is readily apparent. Significantly, no court has expressly upheld a dollar recovery limitation. The state and federal courts that have rejected constitutional attacks on the medical malpractice compensation schemes consistently have dodged this issue.[344] Given a case of medical malpractice that has left a twenty-year old man severely crippled or paralyzed, it appears obvious that the $500,000 maximum recovery would be grossly inadequate to satisfy the victim's medical expenses and lifetime wage losses, let alone provide compensation for pain and suffering, and other noneconomic damages. When a court is eventually forced to confront this situation, it clearly must conclude that this limitation severely burdens the grievously-injured medical malpractice victim.[345]

---

342 Until recently, physicians and their insurance companies had won approximately 80 percent of the medical malpractice cases brought to court. However, physicians have won 92 percent of all screening panel decisions. *See* Margolick, *supra* note 39, at 34, col. 4. While this statistical discrepancy may reflect a variety of factors, a reasonable interpretation is that a number of meritorious claims are not being litigated, and, as compared to the courts, the panels tend to tilt their outlook in favor of the physician.

343 Since many states have imposed restraints on medical malpractice attorneys' contingency fees, the added cost and delay may economically deter effective representation and litigation. Furthermore, it is fair to assume that a sole private plaintiff is far less able than the physician and insurance company defendants to bear the costs of expert witnesses at both the panel and the trial.

344. *E.g.*, Seoane v. Ortho Pharmaceuticals, Inc., 472 F. Supp. 468, 472 (E.D.La. 1979); Hines v. Elkhart General Hosp., 465 F. Supp. 421, 431 (N.D. Ind. 1979); State ex rel. Strykowski v. Wilkie, 81 Wis. 2d 491, 511, 261 N.W.2d 434, 443-44 (1978).

345 The medical malpractice compensation schemes fail to include any provisions comparable to Congress' "express statutory commitment", as construed by the *Duke Power* court, to appropriate additional funds in the event of a catastrophic nuclear accident that exhausts the initial statutory recovery limit. *See* note 242 and accompanying text *supra*.

Some commentators and courts tentatively, and unpersuasively, have attempted to construe associated benefits to the plaintiffs. In *Wright v. Central Du Page Hospital Association*,[346] the court emphatically rejected the defendant's contention that the plaintiff benefitted from a "societal *quid pro quo*" of purportedly lower insurance premiums that would result in reduced medical costs and greater availability of services. The court concluded that the presumed benefit extended to all members of society, whereas only malpractice victims were forced to bear the costs. The plaintiff's alleged benefit relative to his burden was so disproportionately minute as to be virtually irrelevant. Furthermore, since insurance rates are national in scope, large judgments in other states might keep one state's premiums artificially high.

It has been suggested that the existence of a Patients' Compensation Fund provides the victim with the benefit of a "guaranteed recovery."[347] Not only does this assertion distort the meaning of a "guaranteed recovery" in the workmen's compensation context — a strict liability system "guarantees" compensation versus the attendant difficulties of proving negligence — but the threat of an "empty pockets" physician or insurance company, that it purports to resolve, is illusory.[348] No more persuasive is the suggested benefit that a favorable panel decision will aid plaintiff at trial. This element merely supports the underlying merit of plaintiff's claim after much added expense and delay. The plaintiff still must present his full case, including all witnesses, once again at trial.

Conclusion

The medical malpractice compensation schemes fail to provide injured patients with sufficient reciprocal benefits to constitute a "reasonably just substitute" comparable to those approved by the Supreme Court in the workmen's compensation

---

346 63 Ill. 2d 313, 328, 347 N.E.2d 736, 742 (1976).

347 Many of the states that have legislated limitations on recovery do not have a Patients' Compensation Fund. *See* note 25 *supra*.

348 *See* McCoy v. Commonwealth, 37 Pa. Commw. Ct. 530, 551, 391 A.2d 723, 733 (1978) (Mencer, J., dissenting) and note 32 and text accompanying *supra*.

and nuclear accident compensation cases. A court applying the *quid pro quo* review standard must examine the salient features of the state's particular medical malpractice compensation scheme to evaluate whether the legislation, as written, provides a theoretically suitable reasonably just substitute. Especially where a state regime imposes a meager dollar recovery limitation or an unduly burdensome malpractice screening panel plagued with delays, the legislation appears unconstitutional. Furthermore, a scheme that initially passes constitutional muster on a theoretical level may prove to be dysfunctional in its operation. A promising and balanced arbitration procedure, for example, may be demonstrably unworkable or inefficient in practice.[349] A carefully directed judicial inquiry into the continued validity of the prima facie basis for the state's medical malpractice insurance "crisis" may also be pertinent.

The greatest value of the *quid pro quo* standard is that it is flexible, rather than preemptive, and is well suited to the courts' intellectual and institutional expertise. In employing this innovative technique that relies on traditional jurisprudential balancing of benefits and burdens to construe due process safeguards, courts can offer sculpted guidance — instead of an outright veto — to legislatures as they protect fundamental individual liberties.

---

349 See Mattos v. Thompson, 49 U.S.L.W. 2249 (Pa. Sup. Ct., Aug. 22, 1980) (decision that the actual inefficient operation of the state's medical malpractice screening panels violated state constitutional guaranties reverses earlier decision that the panels theoretically satisfied state constitutional guaranties).

FEDERAL PROTECTION OF EMPLOYMENT RECORD PRIVACY

THOMAS E. REINERT, JR.*

In this article, Mr. Reinert argues that federal legislation is necessary to remedy the growing number of abuses of privacy in employment recordkeeping. In the first section, the author identifies and evaluates the employment record privacy problem by examining the concept of informational privacy, the nature and use of employment records, the typical abuses of employee privacy, the available legal responses to privacy violations, and social attitudes towards employment record privacy. He goes on, in section two, to advocate that federal legislation is the best method for employment record privacy protection. In the final section, Mr. Reinert propounds a federal system of employment record privacy protection and argues that such federal regulation can be efficient and politically feasible if implemented simultaneously with reform of the employment data-gathering practices of the federal government.

I. THE EMPLOYMENT RECORD PRIVACY PROBLEM

A growing social concern for informational privacy has focused recently on recordkeeping in the employment relationship, and has raised the prospect of federal legal protection of employment-record privacy. The Privacy Protection Study Commission's (Privacy Commission) extensive investigation of employment record privacy[1] and the recent Department of Labor Hearings on Workplace Privacy (Workplace Privacy Hearings)[2] indicate both public and governmental interest in the pro-

---

* B.A., Harvard College, 1975; J.D., Harvard Law School, 1980; Associate with Morgan, Lewis & Bockius, Washington, D.C. The author wishes to thank Professor Paul C. Weiler, David Williams, and the staff of the Harvard Legislative Research Bureau for their assistance in the preparation of this Note.

1 PRIVACY PROTECTION STUDY COMMISSION, THE REPORT OF THE PRIVACY PROTECTION STUDY COMMISSION (1977). The report consists of a main volume, PERSONAL PRIVACY IN AN INFORMATION SOCIETY [hereinafter cited as PERSONAL PRIVACY], and five appendices. Chapter six of the main volume focuses on the employment relationship, but employee privacy issues are considered throughout the report. Appendix 3, EMPLOYMENT RECORDS [hereinafter cited as EMPLOYMENT RECORDS], provides a more comprehensive version of chapter six.

2 The Department of Labor Hearings on Workplace Privacy [hereinafter cited as *Workplace Privacy Hearings*] were conducted in cooperation with the National Telecommunications and Information Administration of the Department of Commerce during the first three months of 1980 in Washington, D.C., New York, Chicago, and San Francisco. 44 Fed. Reg. 57537 (1979), 44 Fed. Reg. 75755 (1979). Transcripts of the

tection of employee privacy. Surveys of employees find strong support for legal guarantees of employee privacy,[3] and labor unions increasingly raise privacy issues in collective bargaining and legislative lobbying.[4] Many large corporate employers, in accordance with the Privacy Commission's recommendations and in anticipation of federal legislation, are adopting voluntary privacy protection programs.[5] Perhaps most significantly, eight states have enacted statutory protection of employment record privacy.[6] In the opinion of many observers, these developments suggest that federal legislation protecting employment record privacy is likely, if not inevitable.[7]

A. *Informational Privacy*

1. The Privacy Revolution

Although a social and legal interest in an individual's right of privacy has been recognized in the United States for many years, during the last fifteen years privacy concerns have become acute, provoking a social response that has been termed a "privacy revolution."[8]

Two factors have caused the privacy revolution: technological

---

hearings and related written submissions are available at the Department of Labor, Washington, D.C. 45 Fed. Reg. 8780 (1980).

3 *See* LOUIS HARRIS & ASSOCIATES, INC., THE DIMENSIONS OF PRIVACY: A NATIONAL OPINION RESEARCH SURVEY OF ATTITUDES TOWARD PRIVACY (1979) (by commission of Sentry Insurance Co.) [hereinafter cited as SENTRY STUDY]; Westin, *What Should Be Done About Employee Privacy?* 25 PERSONNEL ADMINISTRATOR 27 (1980); text accompanying notes 98 to 125 *infra*.

4 *See Workplace Privacy Hearings*, note 2 *supra;* text accompanying notes 98 to 125, 163 to 174 *infra*.

5 *See Workplace Privacy Hearings*, note 2 *supra;* Westin, note 3 *supra;* text accompanying notes 142 to 162 *infra*.

6 The states are California, Connecticut, Maine, Michigan, Ohio, Oregon, Pennsylvania, and Wisconsin. *See* Westin, *supra* note 3, at 27; text accompanying notes 175 to 198 *infra*.

7 Significantly, this view is shared by both advocates and critics of employment record privacy legislation. *See* Murg & Maledon, *Privacy Legislation — Implications for Employers*, 3 EMPLOYEE REL. L.J. 168 (1977); Schien, *Privacy and Personnel: A Time for Action*, 55 PERSONNEL J. 604 (1976); N.Y. Times, July 6, 1979, at D3, col. 2 (statement of Perry Bullard); *Workplace Privacy Hearings*, note 2 *supra* (statement of Fern Stimpson, Manufacturers Life Insurance Co.).

8 Miller, *The Privacy Revolution: A Report from the Barricades*, 19 WASHBURN L.J. 1 (1979).

advances in electronic data processing and distrust of social institutions. Since the 1960's, recordkeeping systems in many branches of society have become increasingly automated.[9] Computer technology has aided the collection, maintenance, use, and dissemination of greater quantities of information at lower cost,[10] thus threatening privacy by enhancing the power of institutions to keep and use information about individuals. Automatic data-processing has also facilitated decision-making based on recorded information, which may be inaccurate, untimely, or misleadingly taken out of context,[11] resulting in unfairness to the individuals concerned. The impersonal decision-making facilitated by computer technology and necessitated by social complexity has diminished individual trust in institutional decision-makers. Aggravated by revelations of governmental invasion of individual privacy during the 1970's,[12] this distrust has lead to widespread demands for greater accountability of public and private institutions.[13]

The unusual aspect of the privacy revolution is how rapidly a heightened social sensitivity to informational privacy has received a governmental response through legislation and investigation.[14] Significant federal legislation protecting informational privacy has been enacted in the areas of consumer credit records,[15] educational records,[16] and governmental records.[17] Many states have followed the federal government's lead by enacting privacy legislation, most commonly concerning state governmental records.[18] The federal government has given careful and

---

9 U.S. DEP'T OF HEALTH, EDUCATION & WELFARE, SECRETARY'S ADVISORY COMMITTEE ON AUTOMATED PERSONAL DATA SYSTEMS, RECORDS, COMPUTERS, AND THE RIGHTS OF CITIZENS 8 (1973) [hereinafter cited as ADVISORY COMMITTEE REPORT].

10 *See* A. MILLER, THE ASSAULT ON PRIVACY (1971).

11 Miller, *supra* note 8, at 11, 12.

12 *Id.* at 16.

13 Note, *Let Industry Beware: A Survey of Privacy Legislation and its Potential Impact on Business,* 11 TULSA L.J. 68, 70 (1975) [hereinafter cited as *Let Industry Beware*].

14 Miller, *supra* note 8, at 1.

15 The Fair Credit Reporting Act, 15 U.S.C.A. § 1681 (1974).

16 The Family Educational Rights and Privacy Act of 1974, 20 U.S.C.A. § 1232g (1978).

17 The Privacy Act of 1974, 5 U.S.C.A. § 552a (1977).

18 *See* PRIVACY PROTECTION STUDY COMMISSION, PRIVACY LAW IN THE STATES, *in* THE REPORT OF THE PRIVACY PROTECTION STUDY COMMISSION app. 1 (1977) [hereinafter cited as PRIVACY LAW].

extensive study to the problem of informational privacy. The Department of Health, Education and Welfare, through the Secretary's Advisory Committee on Automated Personal Data Systems (Advisory Committee),[19] drew attention to informational privacy issues resulting in the enactment of the Privacy Act of 1974.[20] In addition to improving privacy protection in governmental recordkeeping, the Privacy Act established the Privacy Commission with a charge to investigate the application of Privacy Act protection to the private sector.[21] The Privacy Commission's thorough investigation of privacy and its specific recommendations for legal reform have formed a basis for all subsequent discussion concerning informational privacy.

2. Principles and Policies

Legal theorists since Warren and Brandeis have struggled to develop a general definition of the right of privacy as used in common and constitutional law.[22] Central to the many definitions of privacy that have been proffered is the preservation of individual autonomy and dignity against the intruding and conforming pressure of social organization.

A definition of informational privacy based in individual control over personal information is the starting point for most recent examinations of privacy and recordkeeping.[23] Professor Fried's definition of privacy is elegantly succinct: "Privacy is not simply an absence of information about us in the minds of others; rather it is the *control* we have over information about ourselves."[24] The social value of informational privacy lies in

---

19 ADVISORY COMMITTEE REPORT, note 9 *supra*.

20 5 U.S.C.A. § 552a (1977).

21 PERSONAL PRIVACY, *supra* note 1, at xv.

22 *See* Warren & Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193 (1890); Beaney, *The Right to Privacy and American Law*, 31 LAW & CONTEMP. PROB. 253 (1966); Fried, *Privacy*, 77 YALE L.J. 475 (1968); A. MILLER, note 10 *supra;* W. PROSSER, THE LAW OF TORTS 802-18 (4th ed. 1971). Since this paper deals with constraints on private action, it is primarily concerned with the common law right of privacy.

23 *See* R. K. GREENAWALT, LEGAL PROTECTIONS OF PRIVACY: FINAL REPORT TO THE OFFICE OF TELECOMMUNICATIONS POLICY, EXECUTIVE OFFICE OF THE PRESIDENT 3 (1975); ADVISORY COMMITTEE REPORT, *supra* note 9, at 38. The Privacy Commission did not adopt a definition of informational privacy, but built upon the conceptual structure developed by the Advisory Committee. PERSONAL PRIVACY, *supra* note 1, at 15.

24 Fried, *supra* note 22, at 482 (original emphasis). A similar definition has been

the maintenance of a sense of separateness and individuality, and the prevention of personal embarrassment and economic harm.[25]

The breadth of the informational privacy interest has not been difficult to judge: the individual's interest in information control is usually apparent, and it is simply assumed. The difficulty instead arises in defining the conflicting social interests in privacy intrusion, and in weighing the competing interests. Informational privacy is not an absolute social goal; there are social and institutional needs for information to enable efficient decision-making. Indeed, the individual often has an interest in information disclosure when it promotes his own benefit.[26] Presented with the difficulty of reconciling individual privacy interests and institutional informational interests, the Advisory Committee adopted an approach that has been widely followed in informational privacy policy and legislation: the substitution of procedural protections of informational privacy for substantive determinations among competing social interests. The Advisory Committee adopted a set of "principles of fair information practice," providing individual influence on recordkeeping through procedural rights of notice of the existence of records, access to records with power to amend, and control of information usage.[27]

While maintaining the Advisory Committee's procedural ap-

---

offered by Alan F. Westin: "Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others." A. F. WESTIN, PRIVACY AND FREEDOM 7 (1967).

25 R. K. GREENAWALT, *supra* note 23, at 9.

26 PERSONAL PRIVACY, *supra* note 1, at 5.

27 The principles of fair information practice are:

(1) There must be no personal-data record-keeping systems whose very existence is secret.

(2) There must be a way for an individual to find out what information about him is in a record and how it is used.

(3) There must be a way for an individual to prevent information about him obtained for one purpose from being used or made available for other purposes without his consent.

(4) There must be a way for an individual to correct or amend a record of identifiable information about him.

(5) Any organization creating, maintaining, using, or disseminating records of identifiable personal data must assure the reliability of the data for their intended use and must take reasonable precautions to prevent misuse of the data.

ADVISORY COMMITTEE REPORT, *supra* note 9, at 40-41.

proach and principles of fair information practice,[28] the Privacy Commission developed a more general theoretical structure for formulating informational privacy policy. The Privacy Commission established three general policy goals: (1) minimizing intrusiveness, (2) maximizing fairness, and (3) creating legitimate expectations of confidentiality.[29] These goals correspond roughly to three stages of the recordkeeping process: (1) collection, (2) use, and (3) dissemination of information.

The Privacy Commission defined the goal of minimizing intrusiveness as "creat[ing] a proper balance between what an individual is expected to divulge to a recordkeeping organization and what he seeks in return."[30] The principal component of this goal is giving an individual notice of the need for and potential use of information at the time of solicitation, to enable him to make an informed decision whether to disclose. But the Privacy Commission's goal of minimum intrusiveness goes beyond the procedural approach of the Advisory Committee to a substantive determination that requests for some types of information or certain methods of solicitation may be so intrusive as to warrant general prohibition.[31]

Maximizing fairness entails "open[ing] up record-keeping operations in ways that will minimize the extent to which recorded information about an individual is itself a source of unfairness in any decision about him made on the basis of it."[32] Fairness in information use requires that a record be "accurate, timely, complete and relevant."[33] Procedural rights of access and correction, as well as some limits on collection and dissemination, were viewed by the Privacy Commission as necessary to insure fairness. Fairness is the broadest and most important of the Privacy Commission's goals, effectively encompassing all the Advisory Committee's procedural recommendations. The breadth of the fairness goal is unclear. Some fairness in the use of information is necessary to insure privacy,

---

28 PERSONAL PRIVACY, *supra* note 1, at 15.

29 *Id.* at 14; Linowes, *Employee Rights to Privacy and Access to Personnel Records: A New Look,* 4 EMPLOYEE REL. L.J. 34, 39 (1978).

30 PERSONAL PRIVACY, *supra* note 1, at 14.

31 *Id.* at 16.

32 *Id.* at 14-15.

33 *Id.* at 17.

but the informational privacy concept is much less inclusive than a general notion of fairness. Also, it may be difficult to distinguish a goal of fairness in the use of information from general fairness in decision-making. For this reason, the Privacy Commission stressed that the fairness goal must be defined in terms of specific recordkeeping relationships and the rights and responsibilities inherent in such relationships.[34]

Creating legitimate expectations of confidentiality requires "creat[ing] and defin[ing] obligations with respect to the uses and disclosures that will be made of recorded information about an individual."[35] Confidentiality limits the recordkeeper's freedom to make voluntary disclosure about a subject, and gives a subject the right to resist attempts by third parties to compel disclosure by the recordkeeper.[36] The core of this goal is the imposition of a recordkeeper's privilege analogous to currently recognized legal duties of confidentiality.

The privacy revolution has produced a conceptual structure for analyzing informational privacy problems and formulating privacy protection policy based on the recognition of competing social interests, mediation through procedural approaches, and considerations of intrusiveness, fairness, and confidentiality.

B. *The Employment Context*

Although interest in employment record privacy has been present throughout the privacy revolution,[37] the Privacy Commission's 1977 report has been primarily responsible for directing attention to the issue. Coupled with the recent Workplace Privacy Hearings, the Privacy Commission has generated considerable information about the nature and extent of the employment record privacy problem.

---

34 *Id.* at 18.

35 *Id.* at 15.

36 *Id.* at 20.

37 Since 1970, the Fair Credit Reporting Act, 15 U.S.C.A. § 1681 (1974), has provided limited protections for employees who are the subjects of credit reports. Since 1974, the Privacy Act, 5 U.S.C.A. § 552a (1977), has protected federal employees' privacy. The first state employment record statute, ME. REV. STAT. ANN. tit. 26, § 631 (Supp. 1979), was enacted in 1975. Voluntary employer programs have been in effect since the early 1970's. Ewing & Lankenner, *IBM's Guidelines to Employee Privacy*, HARV. BUS. REV., Sept.-Oct., 1976, at 82.

1. Employment Records

The Privacy Commission found that employment records were the most diverse, perhaps the most important, and in many ways the most problematic of the recordkeeping systems it studied.[38] The importance and diversity of employment records are based, in part, on the central role of employment in most individuals' lives, and the variety of employment situations. Employment records are as diverse as the numerous management functions for which records can be used, including recruitment and selection, promotion, discipline, termination, security, provision of benefits, labor relations, occupational safety and health, and regulatory compliance.[39]

From the employee's perspective, much of the information in an employment record is sensitive.[40] Medical and health insurance records can contain highly personal information, sometimes even of a non-medical nature, that if disclosed can lead to personal embarassment or economic harm to the individual. Security records, if disseminated, can clearly harm an individual's reputation. Performance evaluations, personal references, psychological test results, and attitude assessments are inherently personal and sensitive. Even salary information is considered sensitive by many employees. In addition to the sensitivity of its individual components, the employment record is the most comprehensive data file maintained on most individuals,[41] and therefore poses the greatest potential risk to an individual's privacy.

Computerization of employment records, though currently not widespread, nevertheless poses a significant future privacy problem. While the Privacy Commission focused its investigation on large employers, in part because they were perceived

---

38 *Workplace Privacy Hearings,* note 2 *supra* (statement of Ronald L. Plesser, former General Counsel, Privacy Protection Study Commission).

39 PERSONAL PRIVACY, *supra* note 1, at 225-26.

40 In one survey the following were mentioned by employees as sensitive employment record items: on-the-job performance ratings, salary information, medical records, personality or psychological information, comments on attitude, references and recommendations, and personal histories. Westin, *Privacy and Personnel Records: A Look at Employee Attitudes,* 4 CIV. LIB. REV. 28, 31 (1978).

41 *Workplace Privacy Hearings,* note 2 *supra* (statement of Ray Marshall, Secretary of Labor).

as most likely to use computers in recordkeeping,[42] one study found that even the employment records of the largest employers, with the exception of payroll information, are usually not automated.[43] Still, the trend is toward computerization of employment records, sparked by the growth in the quantity and importance of employment information, and the availability of data-processing technology.[44] Though many of the employers who have adopted privacy protection programs have done so concurrently with the automation of their employment record systems,[45] the effect of computerization on employment record privacy is unclear. Some employers argue that computerization facilitates privacy through the physical securing of record systems and, to reduce storage costs, the elimination of extraneous information.[46] But computerization presents the general danger of facilitating the collection and use of more information about individuals, leading to impersonal decision-making. While computerization may not be a current threat to employment record privacy, it is a significant prospective concern.[47]

Recordkeeping in the employment relationship has grown substantially in recent years.[48] The Privacy Commission identified the causes of this phenomenon as the increase in the provision of benefits and the use of scientific management techniques.[49] Benefits such as pensions and health insurance, with their recordkeeping requirements, are now common. Many employers process health insurance claims for their employees, producing larger files containing sensitive information.[50] Rec-

---

42 PERSONAL PRIVACY, *supra* note 1, at 224.

43 *Workplace Privacy Hearings*, note 2 *supra* (statement of Naomi O. Seligman, McCaffrey, Seligman & Von Simson, Inc.).

44 *Workplace Privacy Hearings*, note 2 *supra* (statement of Alan F. Westin, Columbia University).

45 *See Workplace Privacy Hearings*, note 2 *supra* (statement of Edward S. Cabot, Equitable Life Assurance Society of the United States).

46 *Workplace Privacy Hearings*, note 2 *supra* (statement of Vico E. Henriques, Computer and Business Equipment Manufacturing Association).

47 *Workplace Privacy Hearings*, note 2 *supra* (statement of Prof. David F. Linowes, former Chairman, Privacy Protection Study Commission). Professor Linowes defended the Privacy Commission's focus on large employers, stating that large institutions with computerized data systems are the greatest prospective threat to individual dignity and freedom.

48 PERSONAL PRIVACY, *supra* note 1, at 227-30.

49 *Id.* at 229-30.

50 *Id.* at 229.

ords have multiplied as some employers have become actual providers of health[51] or counselling[52] services. Scientific management, with its emphasis on objective decision-making, increases reliance on personnel records, such as the results of formal systems of performance evaluation.[53] As employers grow in size and complexity, the extent and importance of personnel records increases. Thus employment records have increased their volume of job-related information, as they have simultaneously subsumed many types of non-employment information, including medical, insurance, educational, and credit records.[54]

Increased governmental regulation of the employment relationship is undoubtedly the most significant causal factor in the growth of employment recordkeeping.[55] Employers are currently subject to a broad array of employment legislation that imposes both explicit and implicit recordkeeping requirements.[56] The two statutes with the greatest impact are the Occupational Safety and Health Act (OSHA)[57] and the Equal Employment Opportunity Act (Title VII).[58] OSHA explicitly requires some employers to conduct medical surveillance of their employees,[59] as well as to maintain specific records concerning employee occupational health.[60] Title VII does not require specific employment records, except for an annual statement of racial,

---

51 *Id.* at 229, 258-60.

52 *Workplace Privacy Hearings,* note 2 *supra* (statement of Dr. Sheila Akabas, Columbia University School of Social Work).

53 EMPLOYMENT RECORDS, *supra* note 1, at 2-6.

54 *Workplace Privacy Hearings,* note 2 *supra* (statement of Henry Geller, Assistant Secretary of Commerce for Communication and Information).

55 PERSONAL PRIVACY, *supra* note 1, at 227; Murg & Maledon, note 7 *supra.*

56 Among the federal laws affecting employment recordkeeping are: Title VII of the Civil Rights Act of 1964, 42 U.S.C.A. § 2000e (1974); Executive Order 11246, 41 C.F.R. § 60 (1978); the Age Discrimination in Employment Act of 1967, 29 U.S.C.A. § 621 (1975); the Vocational Rehabilitation Act of 1973, 29 U.S.C.A. § 793 (1975); the Equal Pay Act of 1963, 29 U.S.C.A. § 206 (1978); the Occupational Safety and Health Act of 1972, 29 U.S.C.A. § 651 (1975); the Employee Retirement Income Security Act of 1974, 29 U.S.C.A. § 1001 (1975). *See* Murg & Maledon, *supra* note 7, at 169-70; W. B. CONNOLLY, A PRACTICAL GUIDE TO EQUAL EMPLOYMENT OPPORTUNITY vol. I, at 457 and vol. II, at 141 (1975).

57 29 U.S.C.A. § 651 (1975).

58 42 U.S.C.A. § 2000e (1974).

59 *See* PERSONAL PRIVACY, *supra* note 1, at 228; Note, *OSHA Records and Privacy: Competing Interests in the Workplace,* 27 AM. U.L. REV. 953 (1978) [hereinafter cited as *OSHA Records and Privacy*].

60 OSHA requires enployers to keep a log, a supplementary record, and an annual summary of occupational health problems. 29 C.F.R. § 1904 (1979); M. A. ROTHSTEIN, OCCUPATIONAL SAFETY AND HEALTH LAW 180-91 (1978).

ethnic, and sex composition of the employer's workforce,[61] but it has had a profound effect on employment recordkeeping.[62] At best, employment-discrimination legislation results in the imposition of a system of rational employment decision-making based on more extensive recordkeeping; at worst, it motivates employers, in anticipation of litigation defense, to produce extensive favorable documentation concerning adverse decisions regarding an employee. In either case the result is the recording of sensitive information: genuine performance evaluations, or documentation of an employee's most objectionable characteristics or behavior.

2. Privacy Abuses

There are many current, and often egregious abuses of employee informational privacy. These abuses can be divided into three categories: (1) intrusive data gathering, (2) unfair information use, and (3) breach of confidentiality.

Collection of information about an employee can be intrusive because of the method of solicitation or the type of information sought, and the absence of informed employee choice whether to disclose. Methods of intrusive data collection include truth verification devices such as the polygraph and personality test, which violate an individual's privacy by denying him the ability to control the release of information.[63] Intrusiveness can occur throughout the employment relationship, but it is most apparent at the application stage. Employment applications often require disclosure of information about arrest and criminal records, psychiatric history, familial relations, political and social activities, and credit worthiness.[64] Applicants are also often required to provide blank authorizations that enable employers to gather information from former employers, credit agencies, educational facilities, medical facilities, and social acquaintances.[65]

---

61 Form EEO-1, 29 C.F.R. § 1602.7 (1979).

62 PERSONAL PRIVACY, *supra* note 1, at 227; EMPLOYMENT RECORDS, *supra* note 1, at 26-28.

63 PERSONAL PRIVACY, *supra* note 1, at 239.

64 In the Sentry study, less than half the employees surveyed felt that these items were proper subjects for questions by employers to job applicants. Other types of information judged improper by many employees were pregnancy, marital status, drug and alcohol use, military discharge status, and medical history. SENTRY STUDY, *supra* note 3, at 33; Westin, *supra* note 3, at 28.

65 R. E. SMITH, PRIVACY: HOW TO PROTECT WHAT'S LEFT OF IT 69-72 (1980).

Intrusiveness is, in itself, a limited evil, although it can result in considerable personal embarassment to an individual. The major harm resulting from intrusiveness is that it leads to the presence of sensitive data in employment records, which can subsequently be misused by the employer or disseminated to third parties.

Broadly defined, unfair use of information occurs whenever personal information is used as the basis for an adverse decision about an individual in a manner to which the individual did not consent when disclosing. Unfair use can occur at the application stage, when personal information such as political beliefs or psychiatric history are used to discriminate against the individual.[66] Currently, the most troublesome form of such discrimination involves medical records, which are used to screen out applicants who are physically capable of performing jobs, but because of a history of exposure to toxic substances carry risks of future health problems and employer liability.[67] Another form of unfairness is the collateral use for management decisions of information such as medical, insurance, or counselling records created in the course of providing benefits.[68] When information is inaccurate or irrelevant, its use in decision-making is also considered unfair, although such unfairness transcends privacy concerns.

Breach of confidentiality consists of disclosure of employee information to a third party by the employer without notice to and consent of the employee.[69] Dissemination of employee information may be the most serious form of privacy abuse, because it violates such expectations of privacy and can lead to personal or economic harm to the employee. Disclosures to credit agencies[70] and other employers are often questionable because they can lead to direct economic damage to the indi-

---

66 It is difficult to distinguish unfair use from discrimination based on highly personal criteria. But an employer is privileged to discriminate on the basis of personal criteria such as psychiatric history. The privacy abuse occurs in discriminating after violating an employee's control over the personal information.

67 *OSHA Records and Privacy*, *supra* note 59, at 971-72.

68 PERSONAL PRIVACY, *supra* note 1, at 229, 258, 265.

69 *See* Mironi, *The Confidentiality of Personnel Records: A Legal and Ethical View*, 25 LAB. L.J. 270 (1974).

70 Employers sometimes establish relationships with credit agencies, whereby the employer provides information to the agency in return for other information or compensation. R. E. SMITH, *supra* note 65, at 73.

vidual. Other disclosures, such as to collective bargaining agents[71] or the government,[72] are more problematic because they can both harm and benefit the individual. Although some dissemination of employee information is socially necessary, unrestricted disclosure of personal information from employment records can be a significant privacy problem.

The most egregious privacy abuses usually combine intrusive data gathering, unfair information use, and breach of confidentiality. For example, in *Peller v. Retail Credit Co.*[73] an adolescent who applied for a job in a retail store was asked to undergo a polygraph examination. The polygraph examination indicated falsely that the individual had both used and sold illegal drugs. The retail store employer denied the applicant employment, and then sold the polygraph results to a credit reporting agency. Subsequently, the plaintiff obtained employment at an accounting firm, but was terminated a few weeks after he was hired, when the employer obtained the polygraph results through a credit check. For several years the individual was unable to secure employment. This anecdote illustrates intrusiveness in the use of the polygraph, the credit investigation, and the questions about drug use; unfairness in both employers' use of inaccurate and highly personal information without the individual's knowledge; and a breach of confidentiality by the first employer in sale of the information. Compounded, these privacy abuses lead to grave economic consequences for the individual.

3. Legal Protections

Current legal mechanisms are inadequate for the protection of employment record privacy.[74] Neither common law actions

---

71 Under section 8(a)(5) of the National Labor Relations Act, an employer is under a legal obligation to provide employee information to the collective bargaining agent. *See* Bartosic & Hartley, *The Employer's Duty to Supply Information to the Union — A Study of the Interplay of Administrative and Judicial Rationalization,* 58 CORNELL L. REV. 23 (1972).

72 Employers regularly report information to the government in response to regulatory requirements, subpoenas, or informal requests. EMPLOYMENT RECORDS, *supra* note 1, at 90-95.

73 359 F. Supp. 1235 (N.D. Ga. 1973), *aff'd mem.,* 505 F.2d 733 (5th Cir. 1974). Facts asserted in text are based in part on personal communication with Gary P. Peller.

74 *See* Comment, *Employee Privacy Rights: A Proposal,* 47 FORDHAM L. REV. 155 (1978) [hereinafter cited as *Employee Privacy Rights*].

for invasion of privacy or defamation, nor statutory constraints are sufficient to prevent employer abuses of employee informational privacy.

Intrusive data gathering cannot be remedied by a common law action for invasion of privacy, because courts have held that the mere collection of information about an individual does not constitute a privacy invasion.[75] Aside from these holdings, current definitions of explicit and implicit consent would probably vitiate any claim of privacy invasion.[76] There are some specific statutory prohibitions of intrusive collection practices. Several states prohibit submission to a polygraph examination as a condition of employment.[77] In some jurisdictions specific information such as arrest records,[78] certain conviction records,[79] psychiatric histories,[80] and certain medical information[81] cannot be solicited from a job applicant or utilized in an employment decision. Such specific provisions are a commendable attempt to prevent privacy abuses, but their narrow scope and limited availability make them inadequate to the task of preventing privacy abuses.

Legal restrictions on unfair information use by employers are effectively non-existent. An action for privacy invasion simply does not apply to internal use of employee information.[82] The law views employment records as an employer's property, and generally recognizes no employee interest in the records' content.[83] Absent a system of collective bargaining or a violation of employment discrimination law, an employer is privileged, in making employment decisions, to use employee information in any manner he chooses.

There is no general, legally enforceable duty of confidentiality

---

75 PRIVACY LAW, *supra* note 18, at 21.

76 *Employee Privacy Rights*, *supra* note 74, at 163.

77 *See, e.g.*, MASS. GEN. LAWS ANN. ch. 149, § 19B (West 1971). *See* R. E. SMITH, *supra* note 65, at 251; *Employee Privacy Rights*, *supra* note 74, at 188.

78 *See, e.g.*, CAL. LAB. CODE § 432.7 (West Supp. 1980); MASS. GEN. LAWS ANN. ch. 151B, § 4 (West Supp. 1979). *See* PRIVACY LAW, *supra* note 18, at 17-19.

79 *Id.*

80 *See, e.g.*, MD. ANN. CODE art. 100, § 95-A (Supp. 1978).

81 New York's "Flynn Law," N.Y. EXEC. LAW No. 296, forbids discrimination on the basis of a medical problem that does not affect job performance. EMPLOYMENT RECORDS, *supra* note 1, at 49.

82 *See The Publicity Requirement and the Employee's Right to Privacy*, 16 AM. BUS. L.J. 360, 365 (1979).

83 *See* Mironi, *supra* note 69, at 288-89.

on the part of an employer with respect to employment records.[84] A tort action for invasion of privacy in the public disclosure of private facts[85] can be applied to employer dissemination of sensitive employee information. But few such actions have been brought by employees and fewer still have been successful,[86] because of the availability of strong employer defenses. The publication element of a privacy claim requires widespread public dissemination, which is not met by the employer's disclosure of information to a single third party.[87] If the employer's disclosure is in good faith and in furtherance of his business interest, a defense of qualified privilege is available.[88] And if the employer can show implied employee consent to the disclosure, the privacy claim is vitiated. An action for defamation may be available if the employer's disclosure is untrue. However, a defense of qualified privilege will apply to most employer disclosures which are challenged as libel or slander.[89] Several states have "blacklisting" statutes which prohibit employer disclosures with the intent to harm a former employee.[90] Unfortunately, such statutes cover few disclosure abuses and present evidentiary obstacles to recovery.

The Fair Credit Reporting Act (FCRA)[91] provides some specific protections from violations of employee informational privacy.[92] If an applicant is denied employment because of a credit report, he has a right to learn the "nature and substance" of the file, and then may contest the information.[93] The FCRA is considered a weak statute generally,[94] but it has particular lim-

---

84 Mironi, note 69 *supra*.

85 PROSSER, *supra* note 22, at 809-12.

86 *The Publicity Requirement and the Employee's Right to Privacy*, note 82 *supra*.

87 *But see* Beaumont v. Brown, 401 Mich. 80, 257 N.W.2d 522 (1977), which held that disclosure to a single public employer could constitute publication in an action for privacy invasion, because of the subsequent dissemination of information within the organization. *Beaumont* has been criticized, *The Publicity Requirement and the Employee's Right to Privacy*, note 82 *supra*, and has not yet been followed.

88 *See* Mironi, *supra* note 69, at 283-85; Note, *Qualified Privilege to Defame Employees and Credit Applicants*, 12 HARV. C.R.-C.L. L. REV. 143 (1977).

89 Note, *Qualified Privilege to Defame Employees and Credit Applicants*, note 88 *supra*.

90 *See, e.g.*, CAL. LAB. CODE § 1050 (West 1971); CONN. GEN. STAT. ANN. § 31-51 (West 1972). *See* PRIVACY LAW, *supra* note 18, at 18-19.

91 15 U.S.C.A. § 1681 (1974).

92 *See Employee Privacy Rights*, *supra* note 74, at 165-67.

93 15 U.S.C.A. § 1681g(a)(1) (1974).

94 Miller, *supra* note 8, at 2.

itations in the employment context. First, it applies only to employer disseminations through credit reporting agencies, thus reaching only a small portion of the disclosure abuses.[95] Second, the remedy it provides is difficult to obtain and results only in the limitation of prospective harms.[96] Thus, the FCRA is an inadequate constraint on employment record privacy abuses.[97]

4. Social Concern

Violations of employee informational privacy can result in serious personal or economic harms, which the law has failed to remedy or prevent. But a decision to impose a legislative solution to abuses of employee privacy cannot be based on the mere existence of the problem; rather, it requires a social judgment that employment record privacy is a significant concern. Employers have argued that legislation is inappropriate, because privacy abuses are relatively rare and not an important issue among employees.[98] While employee informational privacy may not be as pressing as some other social or even employment relations issues, there is a developing social consensus that it is sufficiently important to warrant a legal response.

There is limited empirical evidence of the frequency of employee informational privacy abuses. One study, conducted by Louis Harris & Associates, Inc. for the Sentry Insurance Co. (Sentry study),[99] measured employees' perceptions of the privacy problem. The results indicate that relatively few employees believe they have been victims of privacy abuses. When questioned about improper disclosure of employment record information by their employers, 76 percent of the employees sampled

---

95 For example, in Peller v. Retail Credit Co., 359 F. Supp. 1235 (N.D. Ga. 1973), the claim under the FCRA was dismissed because the defendants did not come within the definition of "consumer reporting agency."

96 The FCRA forbids invasion of privacy actions based on information gained from access to credit files absent a showing of malice. 15 U.S.C.A. § 1681h(e) (1974).

97 The Privacy Commission recommended several amendments to the FCRA as a means of providing employee privacy protection. PERSONAL PRIVACY, *supra* note 1, at 231-75. This approach has been criticized as misguided because the FCRA's specific approach to the credit reporting industry is not readily adaptable to all employers. *Employee Privacy Rights, supra* note 74, at 176-77.

98 *See* H. Gorlin, Privacy in the Workplace (Conference Board Information Bulletin No. 27, 1977).

99 SENTRY STUDY, note 3 *supra;* Westin, note 3 *supra.*

answered that it was unlikely, 13 percent found it somewhat likely, and 7 percent very likely.[100] Similarly, on the matter of unfair use of personal information by their employers, 83 percent responded that they knew of no occasions, and 14 percent that they did know of occasions.[101] While the sample indicates that only a small percentage of employees have suffered what they perceive to be privacy violations, this percentage amounts to numerous incidents of privacy abuse nationwide.[102]

As evidence of the lack of employee interest in informational privacy, employers have emphasized the low level of interest among organized labor for privacy issues. It is true that few labor unions have negotiated privacy protections,[103] although it is becoming more common.[104] But there are reasons other than lack of interest why unions have not negotiated employment record privacy protections. First, there is often a conflict between the privacy interests of the individual employee and collective interests.[105] Second, more extensive employee privacy protections have generally been overridden by pressing economic issues in collective bargaining.[106] But as indicated by union leaders' testimony at the Workplace Privacy Hearings, employee privacy is an increasingly important concern of or-

---

100 SENTRY, *supra* note 3, at 36-37.

101 *Id.* On the question of denial of a position or a promotion due to unfair, inaccurate, or out-of-date information, 14 percent stated that they had been denied opportunities because of unfair information, 11 percent because of inaccurate information, and 7 percent because of out-of-date information. *Id.*

102 *Workplace Privacy Hearings,* note 2 *supra* (statement of Humphrey Taylor, Louis Harris and Associates, Inc.).

103 Most notably, the Communications Workers of America and the Bell System since the 1950's have had provisions in many of their collective bargaining agreements giving employees access to their personnel records. EMPLOYMENT RECORDS, *supra* note 1, at 65; *Workplace Privacy Hearings,* note 2 *supra* (statement of James W. Webb, Communications Workers of America). Other unions, such as the United Steelworkers of America, have had informal agreements at the local level concerning access to employment records. *Workplace Privacy Hearings,* note 2 *supra* (statement of Leon Lynch, United Steelworkers of America).

104 Recently, the United Auto Workers negotiated agreements with Ford and General Motors giving employees access to their medical records. [1979] EMPL. SAFETY & HEALTH GUIDE (CCH) # 445. Most privacy negotiation activity concerns medical records. *See Workplace Privacy Hearings,* note 2 *supra* (statement of Steven Wodka, Oil, Chemical and Atomic Workers); EMPLOYMENT RECORDS, *supra* note 1, at 65-66.

105 *See* text accompanying notes 172 to 174 *infra.*

106 *Workplace Privacy Hearings,* note 2 *supra* (statement of George J. Poulin, International Association of Machinists and Aerospace Workers). As one union leader phrased it, privacy is important, but it is not a strike issue. *Workplace Privacy Hearings,* note 2 *supra* (statement of John Zaluzsky, AFL-CIO).

ganized labor.[107] Organized labor supports legislative protections of employee privacy, including restrictions on the use of intrusive techniques such as the polygraph[108] and attitude surveys,[109] and guarantees of the confidentiality of employment records,[110] especially medical records.[111]

The failure of employees to exercise their procedural rights under voluntary employer programs or state employment record privacy legislation[112] has also been cited as evidence of lack of employee interest in employment record privacy. However, this phenomenon may simply indicate lack of employee knowledge of the available protections, fear of employer retaliation for exercise of privacy rights, or absence of a need to invoke protections until abuses arise. Employees may simply perceive the availability of protections as sufficient to constrain employer behavior and protect employee privacy.[113]

Direct evidence indicates that employee interest in the protection of informational privacy is high. The Sentry study found

---

107 *Workplace Privacy Hearings, supra* note 2, *passim.*

108 The Privacy Commission recommended a federal law prohibiting the use of lie detectors in gathering employee information, PERSONAL PRIVACY, *supra* note 1, at 238-40, and former President Carter supported such legislation. *President's Message to Congress on Protection of Individual Privacy, reprinted in* [1979] U.S. CODE CONG. & AD. NEWS 522, 545. There is widespread union support for anti-polygraph legislation. *Workplace Privacy Hearings,* note 2 *supra* (statement of Michael Tiner, United Food & Commercial Workers International Union; statement of James W. Webb, Communications Workers of America; statement of Hugh Harley, United Electrical, Radio and Machine Workers of America).

109 Organized labor is acutely concerned with the use of management techniques in organizational drives, and views methods such as attitude surveys as intrusions on workers' privacy. *Workplace Privacy Hearings,* note 2 *supra* (statement of Alan Kistler, AFL-CIO). *See generally Pressures in Today's Workplace: Oversight Hearings before the Subcomm. on Labor-Management Relations of the House Comm. on Education and Labor,* 96th Cong., 1st Sess. (1979).

110 The AFL-CIO's position is "that the worker must have the right to divulge only that information he or she wishes to communicate, to see what information is in the hands of the employer, and to have items removed from the file and to comment on the content of the file." *Workplace Privacy Hearings,* note 2 *supra* (statement of Alan Kistler, AFL-CIO). At the recent UE convention the union adopted an employees bill of rights including "full access for employees to their employment files." *Id.* (statement of Hugh Harley, United Electrical, Radio and Machine Workers of America). *See also id.* (statement of Leon Lynch, United Steelworkers of America).

111 *See id.* (statement of Nolan W. Hancock, Oil, Chemical and Atomic Workers International Union; statement of Leon Lynch, United Steelworkers of America).

112 There was consensus in testimony at the Workplace Privacy Hearings that employee response to state employment records legislation and voluntary employer programs has been low. *See id.* (statement of Robert Ellis Smith, Editor, *The Privacy Journal*).

113 *See id.* (statement of Robert F. Phillips, Aetna Life and Casualty Co.).

that 93 percent of the employees surveyed thought that their employers should have specific policies for protecting employment record privacy.[114] Employees generally support the legislative enactment of such protections: 70 percent of an employee sample responded that a law giving employees access to their employment records should be enacted;[115] 62 percent of a general sample stated that it is important that Congress enact legislation regulating what information employers can collect about employees.[116] In short, there is substantial support for legal protections of employment record privacy.[117]

The responses of the legal system attest to the strength of the social concern for informational privacy. Most obviously, enactment of state employment record privacy statutes, often in the face of strong employer opposition,[118] indicates legislative recognition of employee privacy abuses.[119] Administratively, the effort by OSHA to enact regulations giving employee access to medical records[120] and the controversy it generated shows the importance of employment record privacy, at least with respect to medical records.[121] Judicially, the Supreme Court's

---

114 SENTRY STUDY, *supra* note 3, at 40.

115 *Id.* at 41. Support among other groups for access legislation was as follows: general public — 65 percent; employers — 33 percent; congressmen — 44 percent. *Id.*

116 *Id.* at 97. Results among other groups were as follows: employers — 30 percent; congressmen — 51 percent. *Id.*

117 "There is a clear public and employee mandate calling on Congress to pass 'privacy legislation' covering private employment . . ." *Workplace Privacy Hearings,* note 2 *supra* (statement of Alan F. Westin, Columbia University).

118 *See, e.g., Workplace Privacy Hearings,* note 2 *supra* (statement of Margaret A. Regan, New York Chamber of Commerce and Industry).

119 There is also evidence of congressional support for employee privacy legislation. *See* notes 115 & 116 *supra.*

120 In July, 1978, OSHA proposed a rule requiring employers to retain medical records and exposure records for a period of five years after employment, and to give present and former employees, their representatives, and OSHA access to such records. 43 Fed. Reg. 31371 (1978). The proposal generated strong labor support and equally strident employer opposition. The comment period on the proposal was extended, and final regulations were not promulgated until May, 1980. 45 Fed. Reg. 35294 (1980); 29 C.F.R. § 1913 et seq. (1980). This OSHA proposal was a background concern of the Workplace Privacy Hearings, since any broader right of access to information would encompass medical records. *See OSHA Records and Privacy,* note 59 *supra.*

121 The Privacy Commission recognized the special importance of the medical record issue, and recommended that the Department of Labor investigate legal controls on the use and dissemination of medical records by employers. PERSONAL PRIVACY, *supra* note 1, at 273-74. Former President Carter proposed legislation dealing generally with medical record privacy that would indirectly affect employers. PRIVACY OF MEDICAL INFORMATION ACT, H.R. Doc. No. 96-84, 96th Cong., 1st Sess. (1978).

decision in *Detroit Edison Co. v. NLRB*[122] is significant because the Court, which formerly had been unreceptive to claims of informational privacy,[123] took judicial notice of an employee's interest in the confidentiality of employment records,[124] and then gave legal effect to that interest.[125]

II. Protecting Employment Record Privacy

Formulating a policy for the protection of employment record privacy requires two steps. First, goals for privacy protection within the employment relationship must be established. Second, a strategy for implementing those practices among employers must be adopted. The Privacy Commission developed a policy of employee privacy protection through procedural rights of participation in recordkeeping, to be implemented primarily through voluntary employer programs. This section analyzes the limitations of such an approach, and considers alternatives.

A. *Privacy Protection Goals*

1. The Procedural Approach in the Employment Relationship

The Privacy Commission approached the employment record privacy problem with unease, at one point even confessing a lack of competence in employment relations.[126] Employment recordkeeping presented difficulties to the Privacy Commission

---

122 440 U.S. 301 (1979).

123 Miller, *supra* note 9, at 5. *See* United States v. Miller, 425 U.S. 435 (1976), in which the Supreme Court held that an individual did not have a legitimate expectation of confidentiality with respect to bank records.

124 In taking judicial notice of "the sensitivity of any human being to disclosure of information that may be taken to bear on his or her basic competence," 440 U.S. at 318, the Court cited the legal effect this interest has been given under state and federal law protecting public employees' informational privacy. 440 U.S. at 318, n.16.

125 The Supreme Court held that the employer's duty to provide information to the union under section 8(a)(5) did not encompass the uncontested disclosure of test scores indicating an employee's competence. Prior to *Detroit Edison*, employers' assertions of employee privacy as a defense to unfair labor practice claims for failing to provide information to unions had been rejected by the courts. *See* Bartosic & Hartley, *supra* note 71, at 29.

126 Employment Records, *supra* note 1, at 70.

because it is quite different from the other recordkeeping areas that were studied. First, computerization and distrust of recordkeeping institutions — the sources of the privacy revolution — are not significant factors in employment recordkeeping. Second, the diversity in employment record content and the uncertain use of employment information in decision-making pose obstacles to the development of general employment record privacy goals. Third, the employment relationship is itself problematic, because of its importance, complexity, and lack of a general structure of rights and duties.

The central problem in formulating a policy for the protection of employment record privacy is balancing the employee's interest in informational privacy against the employer's interest in efficiency through informed decision-making. An employee's interest in controlling the information may be a clear privacy interest, as with information about sexual preference, or it may simply be an interest in avoiding adverse decisions, as with information about poor work performance. An employer's interest in utilizing the information may be legitimate, as with information about prior work experience, or it may be prejudicial, as with information about political beliefs. Thus, a policy determination of whether an employer should be permitted to utilize specific information requires an assessment of the sensitivity of the information to the employee and the relevance of the information to the employer, and a balancing between the competing interests. Such balancing is difficult to institute, both because it depends upon the facts of specific situations and because it requires choosing between competing social values: privacy and efficiency. The Privacy Commission concluded that in a few situations information could be of such high sensitivity and such low relevance as to warrant blanket prohibition. Arrest records were identified as such information, and it was recommended that their use be proscribed.[127] Similarly, the Privacy Commission recommended that the use of truth verification devices be prohibited,[128] because the information they yield is of such high sensitivity and low relevance.

The general policy approach of the Privacy Commission was

---

127 PERSONAL PRIVACY, *supra* note 1, at 242-43.
128 *Id.* at 238-40.

based on procedural protections of employment record privacy, rather than substantive determinations between competing employee and employer informational interests. Mediation through procedural protections was the method for protecting informational privacy developed by the Advisory Committee, and adopted by the Privacy Commission in all recordkeeping areas. In the employment context the procedural approach consists of a set of employee rights of participation in employment recordkeeping. The Privacy Commission's recommendations may be categorized in terms of five basic employee procedural rights:

> (1) *Notice* — The employee should be informed of the type of information about the employee kept by the employer, and how such information is used.[129]
> (2) *Authorization* — Collection of information about the employee from third party sources should require specific authorization by the employee.[130]
> (3) *Access* — The employee should be able to see and copy information about him kept by the employer.[131]
> (4) *Correction* — The employee should be able to contest the accuracy, timeliness or completeness of information about him kept by the employer, and be able to either correct or add a statement of dispute to the record.[132]
> (5) *Confidentiality* — Disclosure of information about an employee to third parties by the employer should generally require employee authorization.[133]

These procedural rights are relatively specific, but implementing them in the employment context requires more specific determinations of their scope and form. Indeed, the choice of implementation strategy is primarily a question of how and by whom these procedural rights will be defined in practice.

The question posed by the procedural approach of employee rights of recordkeeping participation is whether it is an appropriate and effective method for protecting employee informational privacy. These procedural rights appear to be directed at a particular problem: the loss of employee control over information after disclosure. By giving an employee rights of access and correction (so he can see and challenge what is in his

---

129 *Id.* at 236-37.
130 *Id.* at 251-53.
131 *Id.* at 253-61.
132 *Id.* at 261-65.
133 *Id.* at 270-73.

record), and rights of authorization and confidentiality (so he can limit information transfer between the employer and third parties), harm to the employee due to unrestricted circulation of sensitive information can be limited.

These procedural rights, however, seem less appropriate for mediating between the employee's interest in non-disclosure and the employer's interest in information use. Underlying the procedural participation approach, especially the right of notice, is a notion that informing an employee of an employer's recordkeeping practices will lead to fewer privacy violations. This could occur either through the facilitation of informed employee consent to disclosure or bargaining between individual employee and employer concerning disclosure. But the Privacy Commission realized that in some relationships an individual does not have the option of non-disclosure.[134] In the employment relationship, where the consequence of nondisclosure may be unemployment, informed consent seems unlikely. Bargaining concerning information disclosure seems even less plausible; indeed, the untenability of bargaining between the employer and the individual employee is the basic tenet of labor law policy. Nevertheless, procedural rights of participation can lead to protection of employment record privacy indirectly, by exposing recordkeeping practices to examination by employers, employees, and the general public. Employers may limit the intrusiveness of their recordkeeping practices because they discover that some intrusions upon employee privacy lack an efficiency justification, or simply because it is good public relations. Thus, it seems that the privacy protection effect of the procedural approach will be limited to mitigation of privacy abuses in unrestricted circulation of sensitive information and less intrusive employment recordkeeping practices resulting from social exposure and examination.

A substantive approach to privacy policy would require a determination of what type of information about an employee could be collected by the employer, and how it could be used in decision-making. The Privacy Commission rejected a general substantive approach, apparently because they viewed the diversity in employment record content and use as presenting

134 *Id.* at 19.

intractable problems to general policy development.[135] More importantly, a substantive approach to protecting employee informational privacy carries a risk of creating substantial changes in employment decision-making. Currently, with the exception of employment discrimination law, there is no social policy requiring the review of employer decision-making. Substantive protection of employee informational privacy would require a structure for reviewing the fairness of employment decisions with respect to sensitive information use. Since the scope of informational privacy and fairness in information use are poorly defined, it would be difficult in practice to distinguish protections of fairness in the use of sensitive information from general fairness protections. In other words, a substantive system for protecting employee informational privacy could become a means of controlling subtler forms of employment discrimination,[136] or even imposing a fairness or relevance requirement in all employment decision-making. While rationalizing the employment decision-making process is a laudable policy goal, it is a mammoth undertaking, and should not be attempted through a guise of privacy protection.

The procedural approach advocated by the Privacy Commission, while inherently limited in its privacy protection effect to lessening unrestricted circulation of sensitive information and exposing employment recordkeeping practices to public scrutiny, is the best available policy for protecting informational privacy. More extensive privacy protections would require and result in changes in employment decision-making which transcend the privacy concern.

2. Cost

There are three types of costs to employment record privacy protections: (1) the administrative costs of changing employment recordkeeping practices, (2) the informational costs of removing relevant information from employment decisions, and (3) the regulatory costs of implementing a social policy of employee privacy protection.

---

135 *Id.* at 231-32.

136 *Workplace Privacy Hearings*, note 2 *supra* (statement of Ronald L. Plesser, former General Counsel, Privacy Protection Study Commission).

All the available evidence suggests that the administrative costs to employers of instituting privacy protections are not burdensome. An empirical study of the cost of implementing privacy protections in computer systems, including rights of access, correction, and confidentiality, found that for an employment record system the cost of privacy protection would be only $4 per subject per annum.[137] The consensus among employers who have adopted voluntary privacy protection programs is that administrative cost is low.[138] One of the administrative costs of privacy protection to employers is policy development; since a number of innovative employers have developed adaptable policies, that cost is now minimized.[139]

The informational cost to employers of privacy protection is more difficult to assess. Some costs are undoubtedly borne by employers in information loss, but most of the information removed because of privacy protection is probably irrelevant or low in informational value.[140] Large employers who have implemented privacy protection programs seem relatively unconcerned with this cost factor. Some smaller employers who depend upon intrusive practices to screen out high security risk employees, however, are quite concerned with the informational cost of privacy protection.[141]

The support among employers, unions, and the federal government for some form of employment record privacy protection may indicate a social consensus that the administrative and informational costs of privacy protection are outweighed by the

---

137 Goldstein & Nolan, *Personal Privacy Versus the Corporate Computer,* HARV. BUS. REV., Mar.-April, 1975, at 62, 66. Many early estimates of the cost of privacy protection for employment records were quite high because they included the cost of eliminating the Social Security number (SSN) as an identifier. *See* Gorlin, *supra* note 98, at 8. However, the Privacy Commission seemed to minimize the problem of SSN use, PERSONAL PRIVACY, *supra* note 1, at 605-18, and it has subsequently received less attention.

138 *Workplace Privacy Hearings,* note 2 *supra* (statement of T. Michael Jackson, American Society for Personnel Administration; statement of Naomi O. Seligman, McCaffrey, Seligman & Von Simson, Inc.).

139 *Id.* (statement of Edward S. Cabot, Equitable Life Assurance Society of the United States). Employee privacy program development is now common enough that there are instructional manuals available for employers. Westin, *supra* note 3, at 30.

140 D. W. EWING, FREEDOM INSIDE THE ORGANIZATION: BRINGING CIVIL LIBERTIES TO THE WORKPLACE 137-38 (1977).

141 *Workplace Privacy Hearings,* note 2 *supra* (statement of Thomas W. Norton, Fidelifacts/Metropolitan; statement of Harry C. Hunter, National Association of Convenience Stores).

social benefits of employee privacy. There is considerable disagreement, however, over the appropriate implementation strategy for employee privacy protection, which is largely a matter of concern for regulatory costs.

B. *Implementation Strategies*

A decision to protect employee informational privacy through procedural rights of participation in recordkeeping requires the choice of a legal mechanism to implement those rights in employment relationships. Four implementation strategies are available: (1) voluntary compliance, (2) collective bargaining, (3) state legislation, and (4) federal legislation. In this section the advantages and disadvantages of each strategy are examined.

1. Voluntary Employer Compliance

The Privacy Commission recommended voluntary employer compliance as the primary mechanism for implementing its recommended procedural rights in the employment relationship.[142] That recommendation, however, was qualified. First, the voluntary approach was not justified by any lack of severity in employee informational privacy problems. Rather, the Privacy Commission cited the diversity in content and use of employment records,[143] its self-confessed lack of competence in employment relations,[144] and doubts about the efficacy of mandatory enforcement mechanisms.[145] Second, the voluntary approach was experimental: the Commission did not preclude future legislation if voluntary compliance should fail.[146]

President Carter ratified the voluntary approach to employee privacy protection, and directed the Secretary of Labor to cooperate with employers in the establishment of voluntary privacy protection programs.[147] In response, the Secretary initiated

142 PERSONAL PRIVACY, *supra* note 1, at 34.
143 *Id.*
144 EMPLOYMENT RECORDS, *supra* note 1, at 70.
145 *Id.* at 33.
146 PERSONAL PRIVACY, *supra* note 1, at 232-33; Linowes, *supra* note 29, at 42.
147 *President's Message to Congress on Protection of Individual Privacy*, *supra* note 108, at 545.

the Workplace Privacy Hearings to monitor the effectiveness of voluntary compliance and facilitate its implementation.[148] Business presented a united front at the hearings, emphasizing the successful employer programs, stressing the need for more implementation time, and generally praising the voluntary approach.[149] The Workplace Privacy Hearings illustrate the conscious and concerted effort on the part of the business community to anticipate and either prevent or shape employment record privacy legislation.[150] This strategy appears to be based in employers' disapproval of the recent growth in employment regulation,[151] and fear of extensive regulation of recordkeeping practices.[152] But this response is an overreaction, since realistically, any federal legislation protecting employment record privacy is likely to require only the procedural rights already provided by many employers, and not result in extensive regulation of the employment relationship.

The Workplace Privacy Hearings also showed that many employers have made a serious effort to respond to the Privacy Commission's recommendations. The leaders in the development and implementation of voluntary privacy protection programs have been large corporate employers in industries such as data processing, insurance, communications, and finance, for whom privacy has long been a concern.[153] Noteworthy in this

---

148 *Workplace Privacy Hearings*, note 2 *supra* (statement of Ray Marshall, Secretary of Labor).

149 *See id.* (statement of Charles F. Bacon, Business Roundtable; statement of Brenda McChriston, National Association of Manufacturers; statement of Robert M. Hawk, Chamber of Commerce of the United States of America).

150 *See* Schien, note 7 *supra;* WALL ST. J., Oct. 23, 1979, at 1, col. 2; Westin, *supra* note 3, at 30.

151 A frequently raised argument against employment record legislation is that employers are currently overregulated. *See, e.g., Workplace Privacy Hearings*, note 2 *supra* (statement of Robert M. Hawk, Chamber of Commerce of the United States of America; statement of Margaret A. Regan, New York Chamber of Commerce and Industry).

152 The primary source of this anxiety is H.R. 1984, 95th Cong., 1st Sess. (1977), a proposal that would establish a Federal Privacy Board which would monitor every data system in the country, including employment record systems. The prospect of an omnibus, pervasive, and intrusive system of informational privacy regulation has drawn sharp business attention. *See* Gorlin, *supra* note 98, at 4-5; *Let Industry Beware, supra* note 13, at 76-80.

153 Among the most innovative corporations are AT&T, Equitable Life Assurance, Aetna Life and Casualty, General Electric, Cummins Engine, Eastman Kodak, Bank of America, Control Data Corporation, Prudential Insurance. *Workplace Privacy Hearings*, note 2 *supra;* Westin, *supra* note 3, at 30.

regard is IBM, which was one of the first employers to adopt privacy protection policies and practices, and whose management views privacy as a special concern because of the potential privacy threat of data processing.[154] Beyond an interest in good public relations, many large corporations seem genuinely concerned with protecting employee privacy.[155] Progress in privacy protection among smaller employers is much less apparent, but smaller employers are at least becoming aware of privacy issues.[156]

The primary advantage of voluntary compliance as an implementation strategy is that it avoids the regulatory costs of a governmental structure for insuring employee privacy protections. But the current voluntary approach is not cost-free to the government, since a system of informal monitoring of employer progress now exists, as evidenced by the Workplace Privacy Hearings. Indeed, one can question the voluntariness of a system that depends upon the threat of federal legislation for its effectiveness. Because of its avoidance of a regulatory structure, the voluntary approach is politically attractive, especially among businessmen, in a period of anti-regulation sentiment.[157] Another major advantage of voluntary implementation is that it permits flexibility in the creation of employer programs to meet the diverse needs of specific employment situations. Some pro-

---

154 Ewing & Lankenner, note 37 *supra;* EWING, *supra* note 140, at 133-38.

155 This may be based on the personal interests of managers, who can themselves be victims of corporate intrusiveness, and therefore can identify with lower level employees' privacy concerns. Generally, there is broad support among managers for protections of employee privacy. One study found that of a survey of managers, 94 percent agreed with limitations on disclosure of information to third parties without employee consent, 87 percent agreed with employee access to most employment records. Ewing, *What Business Thinks About Employee Rights,* HARV. BUS. REV., Sept.-Oct, 1977, at 81, 86.

156 It is unclear whether privacy violations are a greater or lesser problem with small as compared to large employers. Intrusive practices, such as polygraph use, appear to be more common among smaller employers. But reliance on records and computerization are more common with larger employers. The Privacy Commission's concern and recommendations were primarily directed at larger employers. Small employers' interest in employee privacy is probably based mostly on fear of federal regulation. *See Workplace Privacy Hearings,* note 2 *supra* (statement of Robert M. Hawk, Chamber of Commerce of the United States of America).

157 It is noteworthy that one study found that even in an employee sample that overwhelmingly favored mandatory legal protections of employment record privacy, two-thirds opposed the creation of a federal regulatory structure to implement such protections. Westin, *supra* note 40, at 29.

gressive employers have adopted particular fair information practices that not only meet their specific needs, but exceed the protections recommended by the Privacy Commission.[158]

The disadvantage of the voluntary approach is that it fails to provide sufficient protection of employee informational privacy. Many employers have not adopted voluntary privacy protection programs, and some simply will not, absent legal compulsion. A study conducted by David F. Linowes, former Chairman of the Privacy Commission, found that of the employers surveyed, 64 percent do not have privacy protection programs.[159] The Sentry study found that 36 percent of the employers surveyed would not institute privacy protections until required to do so by law.[160] Both these surveys were directed at the largest corporate employers in the country; with smaller employers, voluntary compliance is surely even less successful. In addition, there is a danger that even employers who adopt privacy protections will compromise employee interests in the formulation or operation of the programs. For example, employers may narrowly define an employee's right of access to employment records by creating exceptions for performance evaluations or medical records.[161] Also, unless there is some assurance that an employee will not be discriminated against by the employer for exercising privacy rights, it is unlikely that voluntary programs will succeed in protecting employee informational privacy.

Thus, while voluntary compliance has provided an opportunity for further examination of the employment record privacy problem by employers, and has been surprisingly successful with some large corporate employers, even advocates of the voluntary approach have suggested that it may be reaching the limits of its effectiveness.[162]

---

158 For example, IBM has minimized the information it requests from job applicants, eliminated checks on credit references, instituted rules for periodic elimination of outdated information from files, and will not release information about a prior employee unless authorized upon termination. Ewing & Lankenner, note 37 *supra;* EWING, *supra* note 140, at 133-38; EMPLOYMENT RECORDS, *supra* note 1, at 89.

159 *Workplace Privacy Hearings,* note 2 *supra* (statement of Prof. David F. Linowes, former Chairman, Privacy Protection Study Commission).

160 SENTRY STUDY, *supra* note 3, at 95.

161 *See* Gorlin, *supra* note 98, at 6-7.

162 *Workplace Privacy Hearings,* note 2 *supra* (statement of Prof. David F. Linowes, former Chairman, Privacy Protection Study Commission; statement of Alan F. Westin, Columbia University).

2. Collective Bargaining

Collective bargaining was viewed by the Privacy Commission as an appropriate means, in addition to voluntary employer programs, for implementing privacy protections in the employment relationship.[163] Privacy protection, especially with respect to medical records, has recently become a more frequent subject of collective bargaining.[164] However, as indicated by the testimony of union leaders at the Workplace Privacy Hearings, negotiated employee privacy rights are still relatively uncommon.

The primary advantage of collective bargaining as an implementation strategy is that it utilizes the currently existing regulatory structure of labor law. Privacy protections can be negotiated in collective bargaining, specified in agreements, and enforced through grievance procedures.[165] Since collective bargaining and grievance procedures already constrain managerial discretion in employment decision-making, there is less concern that the imposition of privacy protections will change the employment decision-making process. Another advantage of collective bargaining is that it permits tailoring of privacy protections to meet employer and employee needs in the particular employment setting.

There are limitations to the collective bargaining mechanism that ultimately make it an inadequate means for implementing employee privacy protections. Most obviously, the fact that over three-quarters of the employees in the United States are unorganized means that most employees cannot directly benefit from negotiated privacy protections.[166] Even unionized workers may not be able to achieve sufficient privacy protection unless they possess considerable bargaining strength.[167] This is illustrated by the Oil, Chemical and Atomic Workers, who have achieved limited success in bargaining for access to medical records.[168] Thus, unions can provide, at best, piecemeal privacy

---

163 PERSONAL PRIVACY, *supra* note 1, at 255.

164 *See* notes 107 to 111 *supra*.

165 EMPLOYMENT RECORDS, *supra* note 1, at 8. Such enforcement includes protection from recriminations by employers for the employee's exercise of privacy rights.

166 PERSONAL PRIVACY, *supra* note 1, at 225.

167 *Workplace Privacy Hearings*, note 2 *supra* (statement of James W. Webb, Communications Workers of America).

168 *Id.* (statements of Nolan W. Hancock and Steven Wodka, Oil, Chemical and Atomic Workers International).

protection.[169] Another limitation is that employment record privacy is an issue for applicants and former employees, as much, if not more than, for current employees. This means that unions, as representatives of current employees, may have limited incentive for bargaining for employment record privacy protections; and legally, such protections may be outside the scope of mandatory bargaining.[170]

Most importantly, there is an inherent tension between employee privacy and collective representation. Under section 8(a)(5) of the National Labor Relations Act,[171] an employer has a duty to supply the union with information relevant to collective bargaining and grievance processing,[172] which results in union access to employment record information. This may provide a means for an employee to get access to his records through the union, and lessens the need for employee access rights in organized workplaces. The employee's and union's interests, however, may be adverse: the employee may not want the employer to disclose sensitive information that the union deems relevant. In *Detroit Edison Co. v. NLRB*[173] the Supreme Court recognized these conflicting interests and held that confidentiality of employee information is valid grounds for an employer's refusal to disclose to the union.[174] Still, the most efficient and effective way for unions to get access to employee information is through the employer under section 8(a)(5). Thus, it may be against the union's interest to negotiate employee privacy protections that limit employer access to or disclosure of employee information.

Collective bargaining has not been shown to be an effective mechanism for securing employee informational privacy rights, and may be an inherently inappropriate implementation strat-

---

169 *Id.* (statement of George J. Poulin, International Association of Machinists and Aerospace Workers).

170 *See* Allied Chemical & Alkali Workers v. Pittsburgh Plate Glass Co., 404 U.S. 157 (1971).

171 29 U.S.C. § 158(a)(3) (1970).

172 *See* Bartosic & Hartley, note 71 *supra*.

173 440 U.S. 301 (1979).

174 *See* note 125 *supra*. *Detroit Edison* dealt with standardized test results, but would appear to have broader scope. Currently, the NLRB is reviewing the same issue with respect to medical records in cases involving Borden Chemical, Colgate Palmolive Co., Minnesota Mining and Manufacturing Co., and the International Oil, Chemical and Atomic Workers. *NLRB Mulls Access to Health Records,* LEGAL TIMES OF WASH., Jan. 21, 1980, at 4. A recent NLRB Administrative Law Judge opinion, Washington Gas Light Co., JD-735-80 (Dec. 17, 1980), found *Detroit Edison* applicable generally to personnel files.

egy. Organized labor's interest in employee privacy rights would be best expressed in political support for employment record privacy legislation and not in negotiation of privacy protections.

3. State Legislation

Eight states have enacted legislation providing mandatory employee privacy protection.[175] Employer opposition to state legislation, however, has been strong,[176] and the statutes that have been enacted are characterized by a weakness denoting political compromise.[177] The state approach is attractive because it seems less burdensome and less intrusive than federal regulation. But this probably is an illusion. The problems presented by employment record privacy do not lend themselves to simple statutory solutions; for this reason, three states have adopted some form of state administrative approach for protecting employee privacy.[178]

The primary disadvantage presented by state implementation is that until all states enact legislation, many employees will remain without informational privacy protection. Another disadvantage in state implementation is that the lack of uniformity among the different state statutes would impose a significant burden upon employers with operations in more than one state. Although inconsistent state standards do not appear to be a current problem,[179] employers may be forced to become advocates of pre-emptive federal legislation if states continue to enact employment record statutes.[180] A state approach to em-

---

175 CAL. LAB. CODE § 1198.5 (West Supp. 1980); Pub. Act No. 79-264, 1979 Conn. Legis. Serv. 554, Pub. Act No. 80-158, 1980 Conn. Pub. Acts 159; ME. REV. STAT. ANN. tit. 26, § 631 (Supp. 1979); MICH. COMP. LAWS ANN. §§ 423.501-423.512 (Supp. 1979); OHIO REV. CODE ANN. § 4113.23 (Page Supp. 1979); OR. REV. STAT. § 652.750 (1977); PA. STAT. ANN. tit. 43, §§ 1321-1324 (Purdon Supp. 1979); Ch. 339, Laws of 1979, 1979-81 Wis. Legis. Serv. 2132 (West).

176 *See* note 118 *supra*.

177 Even the Michigan statute, the strongest and most comprehensive of the acts, was compromised to meet corporate opposition. Letter from Perry Bullard to David Williams (Aug., 1978).

178 CAL. LAB. CODE § 1198.5 (West Supp. 1980) gives interpretive authority to the Labor Commissioner; OHIO REV. CODE ANN. § 4113.23 (Page Supp. 1979) and PA. STAT. ANN. tit. 43, § 1324 (Purdon Supp. 1979) give enforcement authority to, respectively, the Industrial Commissioner and the Bureau of Labor Standards.

179 *Workplace Privacy Hearings*, note 2 *supra* (statement of Naomi O. Seligman, McCaffrey, Seligman & Von Simson, Inc.; statement of Alan F. Westin, Columbia University).

180 Gorlin, *supra* note 98, at 6.

ployment record privacy seems inconsistent with current federal primacy in both employment and privacy regulation.

A review of the current state employment record privacy statutes shows that state implementation is ineffective. Generally, the only right the eight state statutes create is employee access to employment records. Michigan's "Bullard-Plawecki Employee Right to Know Act,"[181] the most comprehensive of the state acts, alone includes restrictions on employer data collection practices.[182] The Michigan and Connecticut statutes also regulate employer disclosure of information.[183] The statutes do not require an employer to notify an employee of the existence and nature of the employment file. None of the acts provide protections for job applicants, and only three of the statutes provide protections for former employees.[184] The statutes qualify the employee right of access by defining the personnel file to exclude certain types of information: references,[185] security investigation reports,[186] future personnel plans,[187] and documents relating to grievance or legal proceedings.[188] The treatment of medical records under the statutes illustrates the danger of inconsistency in state implementation: the Pennsylvania statute explicitly excludes medical records from coverage;[189] the

---

181 MICH. COMP. LAWS ANN. §§ 423.501-423.512 (Supp. 1979).

182 MICH. COMP. LAWS ANN. § 423.508 (Supp. 1979) prohibits unauthorized collection of information about an employee's off-premises associational activities.

183 MICH. COMP. LAWS ANN. § 423.506 (Supp. 1979) prohibits disclosure of disciplinary information to third parties other than the union or the government without notice to the employee. Pub. Act No. 80-158, § 5, 1980 Conn. Pub. Acts 159, prohibits employer disclosure of personnel or medical records to a third party without employee authorization, but has exceptions for specific types of information and certain third parties.

184 MICH. COMP. LAWS ANN. § 423.501 (Supp. 1979); ME. REV. STAT. ANN. tit. 26, § 631 (Supp. 1979); OR. REV. STAT. § 652.750(3) (1977).

185 Pub. Act No. 79-264, § 1(B), 1979 Conn. Legis. Serv. 554; MICH. COMP. LAWS ANN. § 423.501(2)(c)(i) (Supp. 1979); OR. REV. STAT. § 652.750(1)(b); PA. STAT. ANN. tit. 43, § 1321 (Purdon Supp. 1979); Ch. 339, § (6)(b), Laws of 1979, 1979-81 Wis. Legis. Serv. 2133 (West).

186 CAL. LAB. CODE § 1198.5 (West Supp. 1980); Pub. Act No. 79-264, § 1(3), 1979 Conn. Legis. Serv. 554; MICH. COMP. LAWS ANN. § 423.501(2)(c)(v) (Supp. 1979); OR. REV. STAT. § 652.750(1)(b); PA. STAT. ANN. tit. 43, § 1321 (Purdon Supp. 1979); Ch. 339, § (6)(a), Laws of 1979, 1979-81 Wis. Legis. Serv. 2133 (West).

187 Pub. Act No. 80-158, § 1(3), 1980 Conn. Pub. Acts 159; MICH. COMP. LAWS ANN. § 423.501(2)(c)(ii) (Supp. 1979); PA. STAT. ANN. tit. 43, §·1321 (Purdon Supp. 1979); Ch. 339, § (6)(d), Laws of 1979, 1979-81 Wis. Legis. Serv. 2133 (West).

188 Pub. Act No. 79-264, § 1(3), 1979 Conn. Legis. Serv. 554; MICH. COMP. LAWS ANN. § 423.501(2)(c)(vi) (Supp. 1979); PA. STAT. ANN. tit. 43, § 1321 (Purdon Supp. 1979); Ch. 339, § (6)(g), Laws of 1979, 1979-81 Wis. Legis. Serv. 2133 (West).

189 PA. STAT. ANN. tit. 43, § 1321 (Purdon Supp. 1979).

California and Oregon statutes do not mention medical records, but have been interpreted as excluding medical records from coverage;[190] the Maine act covers "nonprivileged medical records";[191] the Michigan act covers only medical records otherwise unavailable to the employee;[192] the Connecticut and Wisconsin acts cover medical records, but under certain conditions only permit access through a physician;[193] the Ohio statute only covers medical records.[194] The weakest aspects of the state statutes are the remedies they provide. Only Michigan gives the employee a civil remedy of injunction or damages for employer violations of the act.[195] Pennsylvania provides the employee with an administrative remedy.[196] California, Maine, and Ohio enforce their acts with civil fines or misdemeanor prosecutions.[197] Connecticut, Oregon, and Wisconsin provide no explicit remedies under their statutes. Most importantly, none of the state acts protect employees from recriminations by their employers for the exercise of their privacy rights. Therefore, it is not surprising that few employees have exercised their rights under these statutes.[198] Because of their narrow scope and weak protections, state employment record privacy statutes have had limited impact on the employment record privacy problem.

4. Federal Legislation

Despite its recommendation for voluntary compliance, the Privacy Commission left open the possibility of federal legis-

---

190 Both statutes apply to records "which are used or have been used to determine the employee's qualifications for employment, promotion, additional compensation or termination or other disciplinary action." CAL. LAB. CODE § 1198.5 (West Supp. 1980); OR. REV. STAT. § 652.750(2) (1977). This language is interpreted as not encompassing medical records. Gorlin, *supra* note 98, at 4.

191 ME. REV. STAT. ANN. tit. 26, § 631 (Supp. 1979).

192 MICH. COMP. LAWS ANN. § 423.501(2)(c)(iii) (Supp. 1979).

193 Pub. Act No. 80-158, § 3, 1980 Conn. Pub. Acts 159; Ch. 339, § (5), Laws of 1979 1979-81 Wis. Legis. Serv. 2133 (West).

194 OHIO REV. CODE ANN. § 4113.23 (Page Supp. 1979).

195 MICH. COMP. LAWS ANN. § 423.511 (Supp. 1979).

196 PA. STAT. ANN. tit. 43, § 1324 (Purdon Supp. 1979).

197 CAL. LAB. CODE § 1199 (West Supp. 1980); ME. REV. STAT. ANN. tit. 26, § 631 (Supp. 1979); OHIO REV. CODE ANN. § 4113.23(D) (Page Supp. 1979).

198 *Workplace Privacy Hearings*, note 2 *supra* (statement of Robert Ellis Smith, Editor, *The Privacy Journal;* statement of Ronald L. Plesser, former General Counsel, Privacy Protection Study Commission).

lation protecting employment record privacy if the voluntary approach should fail.[199] The Workplace Privacy Hearings, even though conducted in the context of executive support for voluntary implementation, implicitly considered the prospect of federal legislation.[200] While there are currently no specific proposals for federal employment record privacy legislation under consideration,[201] there is political support for such legislation and substantial sentiment that federal legislation may be in the offing.[202]

Federal legislation presents some clear advantages as a strategy for implementing employee information privacy protections. First, only a federal approach can provide a uniform minimum level of privacy protection for all employees. The major shortcoming of voluntary compliance, collective bargaining, and state legislation is that both individually and collectively they fail to provide adequate privacy protection for a sufficient number of employees. Second, pre-emptive federal legislation would eliminate the compliance problem for multistate employers presented by potentially inconsistent state statutory requirements. Indeed, the proliferation of state employment record privacy legislation may prove to be the primary force leading to enactment of federal legislation.[203] Third, the federal approach is consistent with federal primacy in employment and privacy regulation. The federal experience in privacy regulation, especially with respect to credit agencies[204] and educational institutions,[205] may be helpful in developing regulation in the employment context. More importantly, a federal approach to employment record privacy can involve concurrent examination of the record-

---

199 *See* note 146 *supra*.

200 Seth Zinman, Associate Solicitor for Legislation, Department of Labor, stressed that the hearings were not directed toward the development of federal legislation, *Labor Tackles Workplace Privacy*, LEGAL TIMES OF WASH., Jan. 7, 1980, at 2; but the prospect of federal legislation was raised frequently in testimony.

201 *Workplace Privacy Hearings*, note 2 *supra* (statement of Fred Feinstein, Chief Counsel, House Subcommittee on Labor-Management Relations). Since the omnibus proposal, H.R. 1984, note 152 *supra*, no federal legislation directed toward employer recordkeeping practices has been proposed.

202 *See* note 7 *supra*.

203 *See* Gorlin, *supra* note 98, at 6.

204 *See* the Fair Credit Reporting Act, 15 U.S.C.A. § 1681 (1974); note 97 *supra;* PERSONAL PRIVACY, *supra* note 1, at 41-100.

205 *See* the Family Educational Rights and Privacy Act of 1974, 20 U.S.C.A. § 1232g (1978); PERSONAL PRIVACY, *supra* note 1, at 393-444.

keeping requirements of federal employment law and the protection of privacy in governmental collection of employee information. In short, only federal legislation can provide comprehensive and uniform protection of employee informational privacy while addressing the federal government's role in the employment record privacy problem.

The disadvantages of the federal implementation approach are political and administrative. Although there is labor and public support for federal legislation protecting employee privacy rights, the business community is adamantly opposed to federal regulation of employment recordkeeping practices. However, employer opposition is not directed at the principle of employee privacy protection, for as the experience of voluntary compliance shows, there is genuine management support for privacy protections. Rather, employers oppose a burdensome system of regulation which intrudes upon employment decision-making.[206]

If the problems of regulatory burden and potential intrusion upon the employment relationship can be overcome in the design of a system of federal employment record privacy protection, then the obvious advantages of the federal approach should lead to the adoption of some form of federal employment record privacy protection.

III. A Federal Approach

While the general prospect of federal legislation protecting employment record privacy has been widely discussed, limited attention has been directed to questions of the specific scope and form of a system of federal legal protections.[207] This section outlines a practical federal approach to employee privacy protection that can serve as a basis for future legislation.

---

206 *See Workplace Privacy Hearings*, note 2 *supra* (statement of Robert M. Hawk, Chamber of Commerce of the United States of America; statenent of T. Michael Jackson, American Society for Personnel Administration).

207 The most comprehensive treatment of the subject is *Employee Privacy Rights*, note 74 *supra*.

A. *Principles*

An ideal system of federal legal protection of employment record privacy would prevent and remedy abuses of employee informational privacy without incurring excessive regulatory costs. To be politically feasible, a federal mechanism must convince employers that it will neither prove administratively burdensome nor intrude upon the employment decision-making process. There are several basic principles that a well designed federal approach should embrace.

Federal employment record privacy protection must be limited in its effect upon the employment relationship to the mitigation of employee informational privacy abuses. Attempts to use privacy protection to reach subtler forms of employment discrimination or to impose a comprehensive system of rational decision-making upon the employment relationship must be expressly disavowed. These more extensive aims should be avoided because they would generate strong adverse reaction from the business community and they would involve an unmanageable regulatory undertaking.[208] The goal of federal legislation should be simply to institute the Privacy Commission's procedural protections in all employment relationships.

Federal protection of employee informational privacy must be based upon enforcement of individual employee privacy rights rather than regulation of employment recordkeeping practices.[209] Regular federal review of employers' recordkeeping systems, as was suggested under earlier omnibus proposals,[210] must be avoided as administratively burdensome and intrusive, if not practically impossible. Rather, legal enforcement of the Privacy Commission's recommended rights of notice, authorization, access, correction, and confidentiality should be the core of a federal approach.

---

208 Alan F. Westin has stated that regulating the relevance of information use in employment decision-making would be the equivalent of drafting a labor code for the United States. *Workplace Privacy Hearings*, note 2 *supra* (statement of Alan F. Westin, Columbia University).

209 The preferability of a rights rather than a regulatory approach is agreed upon by such diverse groups as the ACLU and the Chamber of Commerce. *Id.* (statement of John H. F. Shattuck, ACLU; statement of the Board of Directors of the Chamber of Commerce of the United States).

210 *See* note 152 *supra*.

Federal legislation protecting employee informational privacy must address the federal government's role in the employment record privacy problem.[211] As discussed earlier,[212] the recordkeeping requirements of federal employment law have increased the volume and sensitivity of employment record information. This phenomenon presents a threat of informational privacy abuses by the government and places a paperwork burden upon the employer. While the Privacy Act mitigates informational privacy abuses by the government,[213] the Privacy Commission found that data collection by the federal government still poses a threat to individual privacy.[214] The Privacy Commission directed particular attention to governmental data gathering through third party recordkeepers, such as employers, and recommended that the practice be constrained.[215] It therefore is appropriate that federal legislation protecting employee informational privacy also include mechanisms for examining and constraining the data gathering practices of the federal government with respect to federal employment regulation.

The burden placed on employers by the reporting requirements of federal employment law warrants special concern. The Commission on Federal Paperwork (CFP)[216] studied the general matter of regulatory burden upon businesses and found that it is a genuine problem resulting from the lack of coordination among federal programs.[217] In the employment law area, the paperwork burden is a sufficient problem to actually discourage some employers from hiring.[218] It would seem both efficient and politically expedient for federal legislation protecting employment record privacy to also provide mechanisms for reducing the regulatory burden of federal employment law.

---

211 Murg & Maledon, *supra* note 7, at 176.

212 *See* text accompanying notes 60 to 74 *supra*.

213 *See* PERSONAL PRIVACY, *supra* note 1, at 497-536.

214 *Id.* at 345-91.

215 *Id.* at 388-89.

216 The CFP was a temporary body created by Congress in 1974 with a mission of minimizing the burden of governmental reporting requirements upon the public, especially small business. *See* Note, *Red Tape and National Information Policy: The Commission on Federal Paperwork,* 26 AM. U.L. REV. 1208 (1977) [hereinafter cited as *Red Tape*].

217 *Id.* The Privacy Commission reached a similar conclusion. PERSONAL PRIVACY, *supra* note 1, at 380.

218 *Red Tape, supra* note 216, at 1210 n.9.

B. *Administrative Structure*

Perhaps the most difficult question presented by a federal approach to employment record privacy protection is what type of legal and administrative structure should be used to enforce employee privacy rights. The choice is among using an existing administrative agency, relying upon the judicial system, or creating a new regulatory body.

The Privacy Commission recommended generally that privacy protections be implemented through existing regulatory structures.[219] Underlying such an implementation strategy is a concern over the regulatory cost of a new agency, a conviction that privacy problems can be adequately addressed within current governmental structures, and an uncertainty that privacy issues are sufficiently important to warrant separate regulatory treatment. The National Labor Relations Board (NLRB) has been mentioned in discussions of federal employee privacy protection as an appropriate body for the enforcement of employee privacy rights.[220] But the NLRB, with its particular emphasis on organizational rights and collective interests, would be an inappropriate instrument for enforcing employee privacy protections. The Equal Employment Opportunity Commission (EEOC) and the Occupational Safety and Health Administration (OSHA) would be better choices from among existing federal agencies because they are both involved in the enforcement of individual rights and have experience dealing with employment recordkeeping practices. But neither EEOC nor OSHA seem well suited to employee privacy protection, if for no other reason, because their own information demands intrude upon employee privacy. And since each agency is dedicated to a particular regulatory mission, if privacy protection were added to its functions, it would receive short shrift in the allocation of resources. OSHA and EEOC are also unpopular among employers, so choosing either agency as a privacy enforcer would be a political mistake. Thus, it does not seem that existing regulatory bodies can be utilized effectively to enforce employee privacy rights.

---

219 PERSONAL PRIVACY, *supra* note 1, at 30-31.

220 *Workplace Privacy Hearings*, note 2 *supra* (statement of Alan F. Westin, Columbia University; statement of Robert Ellis Smith, Editor, *The Privacy Journal*).

Federal legislation could rely upon the federal courts as the sole means for enforcing employee privacy rights. However, judicial enforcement presents obstacles of time and money to employees seeking relief from privacy violations. In Michigan, which relies upon judicial enforcement of its employment record privacy act, these obstacles appear to have hindered employee exercise of privacy rights.[221] The complexity of the employment record privacy problem and the difficulty of applying legal protections to diverse employment recordkeeping practices support an administrative approach.

Thus, the question is posed whether the protection of employee informational privacy is a sufficiently important policy goal to warrant the creation of a separate governmental body. Employee privacy is an important social concern, but it does seem less important than other employee interests, such as organizational rights, freedom from race and sex discrimination, and protection of health and safety, which currently warrant separate regulatory structures. Also, the probable limited effectiveness of the procedural approach to employee privacy protection argues against incurring the regulatory costs of a separate agency. Therefore, it would be difficult to justify the creation of a separate regulatory body having as its sole purpose the enforcement of employee privacy rights. For this reason, it is crucial that any attempt to institute federal protection of employment record privacy must be combined with a reorganization and reexamination of the government's data gathering activities under federal employment law. Through reorganization, the costs of a new regulatory body would be minimized, since it would perform the data gathering activities currently conducted by several agencies. Reexamination of the government's recordkeeping activities can produce the additional social benefits of further employee privacy protection and reduction of administrative burden upon employers.

A separate federal regulatory body within the Department of Labor, which could be termed the Federal Employment Records Agency (FERA), should be the administrative structure adopted to enforce employee privacy rights.

---

221 Letter from the American Civil Liberties Union of Michigan to David Williams (July 2, 1979).

C. *Rights and Remedies*

The most important aspect of federal legislation protecting employment record privacy would be the specific rights and remedies it would provide for employees. Federal legislation should provide every employee with the rights of notice, authorization, access, correction, and confidentiality. The statutory definition of these rights should be quite general, thus permitting latitude to employers in implementing specific practices, and discretion to FERA and the courts in interpreting the extent of the rights. Federal employment record privacy legislation would thus parallel the "enforced self-regulation" approach of the Family Educational Rights and Privacy Act (FERPA),[222] which gives students and their parents rights of access, correction, and confidentiality, and permits flexibility to educational institutions and the Department of Health and Human Services in application of these rights.[223]

FERA should be given authority to issue interpretations of the five employee privacy rights. This will be particularly important with the rights of access and confidentiality, for which some exceptions seem necessary. Voluntary employer programs and state statutes have generally defined some information to be outside the scope of employee access rights. For example, it would seem justifiable to exclude security investigation reports and promotional plans from employee access. With respect to confidentiality, some routine "directory information"[224] should be allowed to be disclosed by employers without employee consent. A more difficult question FERA would face is whether disclosures to certain parties, particularly collective bargaining agents and the government, are outside the confidentiality right. Giving FERA discretion to interpret these rights

---

222 20 U.S.C.A. § 1232g (1978).

223 The Privacy Commission indicated support for the "enforced self-regulation" approach, even though it found that FERPA has achieved limited success. The major shortcomings of FERPA stem from the lack of participation by educational institutions in the formulation of its standards. FERPA's application to private institutions is analogous to the application of federal privacy protections to private employers, so its experience should be drawn upon in the formulation of employment record legislation. *See* PERSONAL PRIVACY, *supra* note 1, at 393-444.

224 The Privacy Commission considered the fact of employment, dates of employment, title or position, wage or salary, and location of job site as "directory information" which could be disclosed without notice to the employee. *Id.* at 272.

permits it to assess the ongoing experience of employers, and then make policy determinations of the appropriate scope of employee privacy protections.

In addition to the five procedural rights, it is crucial that employees be given a general right to exercise their privacy rights, in the form of a protection against employer recriminations.[225] The failure to include such a protection may be a major reason for low employee response to state legislation and voluntary compliance programs. Indeed, concern for the lack of an enforcement mechanism was the primary reason the Privacy Commission did not recommend giving legal form to its employee procedural rights.[226] Therefore, statutory protection against discrimination for employees exercising their rights[227] must be provided in federal employment record privacy legislation.

The remedies provided for violations of employee privacy rights should be primarily injunctive, with nominal damages for most violations, and possibly punitive damages for willful disregard of employee privacy rights. Since breaches of confidentiality can lead to direct harms, the statute should provide actual damages (including back pay and reinstatement) for such violations.

D. *Administrative Operation*

In fulfilling its regulatory purposes of protecting employee informational privacy from employer and governmental abuse, and reducing the paperwork burden of federal employment law, FERA should engage in two types of activities: enforcing employee privacy rights, and supervising governmental data gathering.

FERA's enforcement of employee privacy rights would consist of regulatory interpretation of the scope of the statutory rights, and the processing of employee complaints. Under the statute, FERA should have authority to issue interpretive guide-

---

225 *Employee Privacy Rights*, *supra* note 74, at 200; *Workplace Privacy Hearings*, note 2 *supra* (statement of John Zaluzsky, AFL-CIO).

226 PERSONAL PRIVACY, *supra* note 1, at 231-33.

227 The provision should be analogous to section 8(a)(3) of the National Labor Relations Act. *See Employee Privacy Rights*, *supra* note 74, at 199-200.

lines to aid employers in complying with statutory requirements. A system for the processing of employee complaints should be modeled after the EEOC.[228] FERA would receive an employee's complaint, investigate the situation, try to negotiate a settlement, and then determine whether to issue a "right to sue letter." After FERA processed an employee's claim, the employee would have to determine whether to pursue his judicial remedies. Like the EEOC, FERA should also be given power to prosecute patterns or practices of employee privacy violation.[229]

FERA's resources would not be focused upon the processing of employee complaints or upon litigation. Rather, its regulatory activities would be much more informal. Larger employers who develop privacy programs could submit them to FERA for review, thus effectively receiving immunity from future employee suits. Few employees or employers would want to litigate privacy issues, so that most complaints would result in FERA negotiated settlements. FERA's primary contribution to employee privacy protection would be as an informational resource for employers, providing model privacy protection programs and information about the experiences of other employers. FERA should not be envisioned as an active and permanent agency for adjudicating employee privacy claims. Rather, it should be viewed as a method for providing the means and the incentive for self-regulation by employers.[230] For this reason, it would be useful if federal legislation provided a long lag time — perhaps two years — between the establishment of FERA and the time that employee privacy rights become operative. This would provide time for the education of employers and for changes in recordkeeping systems. In summary, it seems possible for a regulatory body such as FERA to effect the implementation of employee privacy protections without an expensive, burdensome, and intrusive administrative process.

FERA's primary activity, at least in terms of resource investment, would be the standardization and centralization of

---

228 *See* 42 U.S.C.A. § 2000e, at 705-09 (1974); B. L. SCHLEI & P. GROSSMAN, EMPLOYMENT DISCRIMINATION LAW (1976).

229 *See* note 228 *supra*.

230 The success of legislation, particularly in the employment area, depends upon changes in social values as much as specific regulatory mechanisms. Mironi, *supra* note 69, at 287.

the data gathering functions under federal employment law.[231] FERA would mediate between employers and federal agencies seeking employment information, including EEOC, OSHA, the Department of Labor, the NLRB, and perhaps the Internal Revenue Service and the Social Security Administration. This mediation would include the elimination of redundant information requests and the standardization of employee reporting forms. But FERA's activities would go beyond mere standardization, and would include the processing of all employee information requests.[232] Federal agencies would request employee information from FERA, which would judge regulatory need, employee privacy, and administrative burden in determining whether to process the request. Centralization of governmental records of employee information under control of FERA should be considered as a method for increasing efficiency in the processing of information requests.[233] FERA would also be responsible for policing internal federal use of employee information to insure privacy protection.[234] FERA's activity might result in other federal agencies obtaining less information about employees than they desire. But it seems that an independent body with functions of privacy protection and the elimination of administrative burden would be less biased in determining informational relevance than an agency with a single substantive regulatory mission. Any transaction costs incurred in using an intermediary such as FERA in the processing of information requests should be outweighed by the efficiency of centralized and standardized information processing, and the benefits of employee privacy and reduced paperwork for employers. But the primary benefit of using FERA for employment data gathering would be political. FERA's regulatory goals would include

---

231 These functions would be consistent with the recommendations of the CFP, which advised general centralization and standardization of federal information gathering functions. *See Red Tape*, note 216 *supra*. Standardization and centralization in a limited regulatory area such as employment law would seem more practical and politically feasible than CFP's general approach.

232 Agencies such as OSHA and EEOC would still be able to conduct investigations and inspections; FERA would only supervise regular data-gathering functions.

233 This aspect of FERA's function is left open both because it would involve great physical and administrative reorganization, and because centralization of governmental data files may itself pose privacy problems. *See* Note, *Privacy and Efficient Government: Proposals for a National Data Center*, 82 HARV. L. REV. 400 (1968).

234 This is the primary function of the Privacy Act, and thus would involve some modifications in its provisions.

aiding employers in reducing paperwork, and its information requests would be separated from the substantive enforcement of federal employment law.

CONCLUSION

A federal system of employee rights of notice, authorization, access, correction, and confidentiality with respect to employment records can provide an effective, efficient, and politically feasible solution to the employment record privacy problem. Employee privacy protections should be administered in a regulatory structure that addresses the governmental role in employee privacy abuses and provides employers with paperwork relief from federal employment regulation. Since, as the Chairman of IBM has noted, "Privacy is not a passing fad,"[235] federal legislation protecting employee informational privacy seems inevitable.

235 Ewing & Lankenner, *supra* note 37, at 82 (statement of Frank T. Cary).

BOOK REVIEW

SOVIET LEADERSHIP IN TRANSITION. By *Jerry F. Hough.* Washington, D.C.: The Brookings Institution, 1980. Pp. 175

*Review by Harold J. Berman**

This book combines sharp criticism of certain postulates of current U.S. foreign policy vis-à-vis the Soviet Union with an analysis of generational changes in the Soviet middle and upper strata from which the top leadership comes. The author, himself a member of the younger generation of American students of Soviet affairs, charges an older generation of American foreign policy makers with failure to understand the dynamic aspects of the Soviet political system. He states that Zbigniew Brzezinski, Samuel Huntington, and others who have been regarded as experts on Soviet intentions see only a dead or dying totalitarianism, an ossified bureaucracy, incapable of significant change. In reality, however, the Soviet Union has undergone, is undergoing, and is likely to undergo in the future, very substantial changes. Moreover, the United States can, if it recognizes this, exert some influence on the course of those changes, so that they will be more compatible with our own interests.

Prevailing American perceptions of the Soviet Union are based, in Professor Hough's view, on highly abstract conceptions and have little relationship to the actual life experience of the people who run Soviet society. "[W]e have almost no understanding of the bureaucrats," he writes, "the people belonging to the Soviet middle class or upper middle class . . . [the] changes [that] are taking place among them . . . their thinking . . . [the] policy dilemmas [that] they face . . . ." (p. 7). Yet it is these people whose thoughts and values form the soil, so to speak, in which Soviet policy grows.

But how are we to find out what Soviet "bureaucrats" are really like? Hardly any foreigners ever even meet members of the Central Committee of the Communist Party, or of its staff,

---

* Ames Professor of Law, Harvard Law School. A.B., Dartmouth, 1938; M.A., Yale, 1942; L.L.B., Yale, 1947.

other than the very top leaders. The Party *apparat* — which runs the country — operates almost entirely behind the scenes. Apparently it does not want us — or even its own people — to know at first hand the changes that are taking place within it, in its thinking, or in the policy dilemmas it faces.

Professor Hough's principal method of penetrating to these secrets is to analyze some of the main events through which successive generations of the Soviet bureaucracy have lived during the past sixty years, and to give some statistics concerning the education, the activities, and the careers of people of various age groups. The results are interesting and important, although they constitute a less than adequate foundation for a substantial critique of U.S. foreign policy, since at the end we still do not know much about the people who shape Soviet policy.

The author's "generations" are, in fact, decades; that is, he takes Brezhnev as a member of the "generation," now in its seventies, that experienced, as adolescents and young adults, the first ten years of the Bolshevik Revolution and that rose to prominence in the late 1930's and in World War II. Hough sees a difference in the "life experience" of the next "generation," born between 1910 and 1920, many of whom began their careers after World War II, and a still greater difference in that of the men born between 1920 and 1930. More generally, he stresses that the generation of Soviet bureaucrats who are now leaving the scene, or who will leave it soon, are much less well-educated than the generation of those who are about to succeed them (today, he emphasizes, half the college graduates between the ages of thirty and sixty are members of the Party); that the present generation of bureaucrats was almost wholly deprived of contact with foreign countries until it was middle-aged; and that its leadership careers began in the worst period of the Stalin era. In addition, members of the generation now in its fifties were subjected, early in their careers, to the genuine shock of Khrushchev's revelations about Stalin's terror.

Professor Hough also stresses socio-economic changes; for example, in 1926 only ten million Soviet people, comprising seven percent of the total population, lived in cities of over 100,000, whereas in 1979 the figure was almost 100 million people, comprising almost forty percent of the population. The

present generation of Soviet leaders grew up in a developing country; their successors will have grown up in a highly industrialized economy.

These and many other statistical data are used to refute the prevailing stereotypes of an unchanging people ("It was the same under the tsars — they're the same people") and an unchanging system (whether that of a petrified bureaucracy concerned only with maintaining itself in power or that of a committed conspiracy to spread Communism throughout the world). Two general conclusions emerge. The first is that the people and the system are changing and will continue to change and therefore we should be prepared for change and should give inducements for such changes as would be beneficial to us. The second, which is implied rather than expressed, is that a new generation of Soviet leaders is apt to be more sophisticated, more cosmopolitan, more tractable, more open to negotiation, yet at the same time perhaps less cautious, than their predecessors.

More specifically, Professor Hough states that "it is vital that we offer "[the Soviet Union] some hope that it can pursue a policy of economic reform, liberalization, and reduction of military expenditure without national humiliation" (p. 166). He also states his belief that a new and younger Soviet leadership "may be more willing to take risks involving military action" (p. 167) — which means, once again, that we should act so as to influence Soviet cost-benefit calculations in our favor.

Although Professor Hough's main emphasis is on cost-benefit calculations in a changing world, and on a flexible American policy that would seek to reward Soviet actions that we favor, he also recognizes, at several points, the importance of certain basic principles in international relations — principles that seem to go beyond immediate mutual advantage. Thus he states that we must not demand that the Soviet Union "abstain from action in the third world that we take for granted for ourselves . . . [or] issue ultimatums on types of behavior that we are not willing to forswear ourselves . . . ." (p. 168). (Presumably the Soviet leaders, too, should adhere to universal principles for whose violation they attack us.) More striking, in view of the author's main emphasis on what we can do by way of accommodation to induce a reduction of Soviet military expenditures and of Soviet military power generally, is his analysis of the Soviet

invasion of Afghanistan: he writes that the Soviet leaders may have thought that "because Afghanistan had been in the Soviet sphere of influence for decades, it was a special case," but that if, on the contrary, they acted on the principle that Marxist-Leninist revolutions in the third world are irreversible and that the Soviet Union "should or must intervene militarily if they are near collapse," "the situation becomes dangerous." "[T]he invasion of Afghanistan," he concludes, "has set a precedent that, if followed, will ultimately lead to World War III" (p. 166).

This dramatic and solitary reference to a "precedent" leading to World War III brings the reader — or at least brought this reader — up short. Is there a mailed fist inside the author's velvet glove? Is he saying that the Soviets are changing and therefore we can make peace with them, but if they do not change in certain ways, then a major war is inevitable? More specifically, is his emphasis on the value of negotiation, with give and take on both sides, premised on an assumption not merely of change but also of Soviet modification of long-range goals? If so, then it becomes important to identify those goals.

What is missing is what the statistics cannot give us, namely, information about the actual beliefs of successive generations of Soviet leaders. What do the members of the Central Committee, including the staff, and the thousands of heads of important Party Organizations throughout the country that are responsible to the Central Committee staff, really believe? What do they really believe *in*? As Professor Hough emphasizes, the models of "totalitarianism" and "bureaucracy," from which much of the prevailing American thought about Soviet policy derives, are too static and too abstract to be satisfactory. But the statistical information that he presents only takes us a short distance further. It helps us to know that the old models are too static and too abstract, but it does not give us anything to replace them with except a model of changing circumstances.

Does the coming generation of Soviet leaders believe, or not believe, that "Marxist-Leninist revolutions in the third world" (Ethiopia? Cuba?) are irreversible and should or must be preserved by Soviet military intervention if they are about to collapse? Does it believe, or not believe, that substantially greater decentralization of economic decision-making and the intro-

duction of certain principles of market socialism, as in Hungary, would be compatible with the long-term political and military goals of the Soviet planned economy? How seriously does it take the present leadership's campaign to make socialist legality and protection of the Soviet legal order matters of prime concern both of Soviet officialdom and of the population as a whole? Does it pay more than lip service to the Leninist vision of the future? What, if anything, has taken the place of that vision?

It is possible, perhaps even likely, that the answers to these and similar questions would show that the kinds of changes explored by Professor Hough really do matter in the making of Soviet foreign policy; but without those answers one can only speculate concerning the responses that a new Soviet leadership might make to American initiatives.

RECENT PUBLICATIONS

AFFIRMATIVE DISCRIMINATION: ETHNIC INEQUALITY AND PUBLIC POLICY. By *Nathan Glazer*. New York: Basic Books, 1978. Pp. xvii, 221, notes, index.

In *Affirmative Discrimination,* Nathan Glazer attacks the wisdom of shifting attention away from individual and towards group rights in developing policies to overcome the legacy of racial discrimination. Since 1975, when the book first appeared, the issue has expanded considerably. In the *Bakke* and *Weber* cases, the Supreme Court has attempted to determine the proper role of ethnicity in college admissions and obtaining employment. Yet, the controversy continues. Glazer's provocative book is a useful reference for anyone interested in the subtleties of affirmative action policy.

The first chapter traces the emergence of an American ethnic pattern. From the very inception of the Republic, Glazer maintains, no separate ethnic group was to be allowed to establish itself independently within the United States. The enslavement of the blacks, the near extermination of the American Indians, and the official persecution of Asian-Americans are dismissed as aberrations from an otherwise color-blind American national character. The dismantling of an entrenched system of prejudice and discrimination in law and custom, however, beginning with the Supreme Court decision in 1954 barring segregation in public schools, to the Civil Rights Act of 1964 and the Voting Rights Act of 1965, represented the culmination of the fight for legal equality.

Yet, as civil rights advocates shifted their attention to advancement in economic and political strength, an internal contradiction arose within the affirmative action policies themselves. In 1964, legislation was passed declaring that race, color, national origin, and religion should not be used to determine qualifications for voting, employment, or educational opportunity. Yet, as Glazer points out, "Having placed into law the dissenting opinion of *Plessy v. Ferguson* that our Constitution is color-blind, we entered into a period of color and group-consciousness with a vengeance" (p. 31).

In the first part of the book, Glazer assails the results of affirmative action in employment. He points out that during the

early 1970's, increasing numbers of employers were subjected to stringent bureaucratic controls so that each job offer, each promotion, each dismissal had to be considered for its effects on group ratios in employment. Inevitably, Glazer asserts, this meant that the ethnic affiliation of an individual began to determine the individual's future employment prospects. Glazer deeply deplores the change in emphasis away from equal opportunity towards statistical parity.

Glazer carefully distinguishes the civil rights movement from the present system of affirmative action. While acknowledging the need for universal enfranchisement and compulsory education, he points out that hiring is based on qualifications, which necessarily vary with the individual. Glazer's most trenchant statement is that actual acts of discrimination against individuals have been ignored by the Equal Employment Opportunity Commission in favor of statistical pattern-seeking. In emphasizing statistics, rather than personal discrimination, the Commission assumes that racism is the operative cause in any case of differential treatment, rather than concern about qualifications. "Opportunity, it seems, is being redefined as result," Glazer concludes, "[while] . . . equal opportunity, not statistical distribution, is the proper objective of public policy" (p. 48).

The second part of the book addresses the social and political consequences of public policies which emphasize rigid lines of division between ethnic groups and make the ethnic characteristics of individuals primarily responsible for their personal fate. The comparison of experience, the rationale goes, causes white ethnic group to perceive that they were treated unfairly relative to the groups benefited by affirmative action policies (p. 195).

In the last section of the book, Glazer considers the future of affirmative action. He argues that the legacy of discrimination can be overcome by simply attacking discrimination as such. He further asserts that the statistical-pattern approach was instituted *after,* not before, the economic position of blacks improved dramatically in the 1960's. Glazer fears the political consequences of the increasing resentment and hostility between groups that is fueled by special benefits for some. Glazer's final argument is that affirmative action abandons a fundamental principle of any liberal society: that the rights of the individual are sacrosanct.

Affirmative Discrimination forces the reader to take a hard look at the assumptions upon which affirmative action policies are based. Before we automatically applaud well-intentioned, but specious attempts to redress past wrongs, we must reckon with the consequences to society. The author argues convincingly that the goal of integration will not be served by codifying fixed categories of ethnic identification and using these categories to determine the fate of the individual.

Matthew J. Goldman

THE PERMANENT CAMPAIGN: INSIDE THE WORLD OF ELITE POLITICAL OPERATIVES. By *Sidney Blumenthal.* Boston: Beacon Press, 1980. Pp. 264. $12.95

Sidney Blumenthal's book, *The Permanent Campaign: Inside the World of Elite Political Operatives,* examines the impact which media experts, direct mail specialists, and other consultants have had on recent American political campaigns. While the topic deserves careful attention, this collection of profiles leads the reader to suspect that the author just barely beat an editor's deadline. Insights and cliches coexist uneasily in this frantic blend of history, biography, political analysis, and psychoanalysis

The Permanent Campaign purports to be a "description of the new process of American politics" (p. 10), but it never develops coherent themes nor transcends conventional wisdom concerning "political operatives." Chapters lose momentum while the author pursues fascinating but frustrating tangents, and the overall structure of the book is puzzling. The final section, which might have tied the book together, is particularly disappointing. Instead of synthesizing his ideas or suggesting a plan for reform, Blumenthal concludes with a tribute to Edward Kennedy which seems unrelated to the rest of the book.

Blumenthal also runs into problems by portraying his subjects as a special breed of men. He wants us to feel awed by these techno-political geniuses, yet he also wants us to feel vaguely disturbed by their manipulative skills. In order to accomplish this goal, he tends to inflate the achievements of his subjects. For instance, he would have us believe that Carter pollster Pat

Caddell "has forged the era of the antipolitician, the outsider candidate who can inflict piercing wounds on an incumbent, but whose platform is thematic rather than programmatic" (p. 29). One should be skeptical whenever someone credits a Harvard graduate in his twenties with "forging an era."

We should also remain skeptical about the claims made by, and on behalf of, the top political consultants. On the average, about fifty percent of their clients will win, hence it is not remarkable that many consultants have chalked up impressive win-loss records. When a consultant wins a few races and builds up a reputation, he can charge higher prices and attract better financed candidates. Unsurprisingly, his new candidates tend to win elections and add to his reputation. This self-fulfilling prophecy is in large part responsible for the mystique of David Garth and his prominent colleagues; Garth's manhandling of the Anderson campaign demonstrates that these "wizards" have not yet reduced campaigning to a science. Indeed, as media experts multiply and cancel out each other, political organization becomes increasingly important, particularly in caucuses and primaries.

Despite the flaws of *The Permanent Campaign*, the book does have occasional flashes of brilliance. Blumenthal's analysis of the Carter-Caddell relationship provides insight into a baffling administration; every potential king or kingmaker should read the chapter on Don Rose's role in the Jane Byrne campaign. The political junkie will also find a healthy fix of trivia. The more one considers *The Permanent Campaign*, however, the more one doubts Blumenthal's claim that consultants have created a "new process of American politics" (p. 10). Political hacks are much the same as they have always been: they merely have new toys.

Michael J. Astrue

BOOKS RECEIVED

ABUSE ON WALL STREET: CONFLICTS OF INTEREST IN THE SECURITIES MARKETS. By *Twentieth Century Fund.* Westport, Conn.: Quorum Books, 1980. Pp. xv, 620, index.

BRITAIN'S ECONOMIC PERFORMANCE. Edited by *Richard E. Caves* and *Lawrence B. Krause.* Washington, D.C.: The Brookings Institute, 1980. Pp. viii, 388, index. $18.95 cloth, $7.95 paper.

THE COLLAPSE OF WELFARE REFORM: POLITICAL INSTITUTIONS, POLICY AND THE POOR IN CANADA AND THE UNITED STATES. By *Christopher Leman.* Cambridge, Ma.: The MIT Press, 1980. Pp. xvii, 292, notes, index, $19.95.

DECISION TO PROSECUTE: ORGANIZATION AND PUBLIC POLICY IN THE ANTITRUST DIVISION. By *Suzanne Weaver.* Cambridge, Ma.: The MIT Press, 1980. Pp. viii, 196, index, notes. $5.95.

DEEPSEA MINING. By *Judith T. Kildow.* Cambridge, Ma.: The MIT Press, 1980. Pp. x, 251. $17.50.

ECONOMIC INTERACTION IN THE PACIFIC BASIN. By *Lawrence B. Krause* and *Suzo Sekiguchi,* eds. Washington, D.C.: The Brookings Institution, 1980. Pps. xii, 269, index. $14.95 cloth, $5.95, paper.

FINANCING GOVERNMENT IN A FEDERAL SYSTEM. By *George F. Break,* Washington, D.C.: The Brookings Institution, 1980. Pp. viii, 276, index.

FOUNDATIONS OF POLITICAL ANALYSIS: AN INTRODUCTION TO THE THEORY OF COLLECTIVE CHOICE. By *Robert Abrams.* New York: Columbia University Press, 1980. Pp. x, 357, references, indices. $16.50.

HER HOME: A WOMAN'S GUIDE TO BUYING REAL ESTATE. By *Ruth Rejnis.* New York: Anchor Press/Doubleday, 1980. Pp. 183, index, glossary. $8.95.

HISTORY OF THE IDEA OF PROGRESS. By *Robert Nisbet.* New York: Basic Books, 1980. Pp. xi, 370, index. $16.95.

INDIVIDUALISM AND THE PHILOSOPHY OF THE SOCIAL SCIENCES. By *Murray N. Rothbard.* San Francisco: Cato Institute, 1979. Pp. x, 67. $2.00, paper.

LABOR-LAW IN CONTRACTORS' LANGUAGE. By *McNeill Stokes.* New York: MrGraw-Hill, 1980. Pp. xviii, 434, table of case, index. $22.50.

THE LAW PRACTICE OF ALEXANDER HAMILTON: DOCUMENTS AND COMMENTARY. By *Julius Goebel, Jr.* and *Joseph H. Smith,* eds. New York: Columbia University Press. Vol. III, xxviii, 844, index, $50.00. Vol. IV, xviii, 668, index.

LAW AND SCIENCE: A SELECTED BIBLIOGRAPHY. By *Morris L. Cohen, Naomi Ronen,* and *Jan Stepan,* compilers. Cambridge, Mass.: The MIT Press, 1980. Pp. xvii, 155, index. $15.00.

LAW OF SENTENCING. By *Arthur W. Campbell.* Rochester, N.Y.: Lawyers Cooperative Publishing Co., 1978. Pp. xiii, 560, appendices, table of casis, index.

LEGISLATION AND SOCIETY IN AUSTRALIA. By *Roman Tomasic,* ed. Sydney, NSIU, Australia: The Law Foundation of New South Wales, 1980. Pp. 420, notes, appendix.

MASS MEDIA AND AMERICAN POLITICS. By *Don A. Graber.* Washington, D.C.: Congressional Quarterly Press, 1980. Pp. xiii, 304, index, paper.

NEW ESTATE PLANNING HANDBOOK WITH FORMS AND TABLES. By *Howard M. Zaritsky* and *Martha Altschuler Zaritsky.* Englewood Cliffs, N.J.: Prentice Hall, Inc., 1980. Pp. xv, 438, index.

THE POLITICAL ECONOMY OF REGULATION: CREATING, DESIGNING AND REMOVING REGULATORY FORMS. By *Barry M. Mitnick,* New York: Columbia University Press, 1980. Pp. xv, 506, references, index. $20.00.

THE POLITICS OF PRIVACY, COMPUTERS, AND CRIMINAL JUSTICE RECORDS. By *Donald A. Marchard.* Arlington, Va.: Information Resources Press, 1980. Pp. xvi, 433, bibliography, appendices, index. $34.95.

THE POLITICS OF SCHOOL DESEGREGATION: THE METROPOLITAN REMEDY IN DELAWARE. By *Jeffrey A. Raffel.* Philadelphia, Pa.: Temple University Press, 1980. Pp. viii, 279, notes, glossary, index. $19.50.

THE PRESENT DANGER. By *Norman Podhoretz.* New York: Simon and Schuster, 1980. Pp. 109, notes. $7.95.

THE PUBLIC'S BUSINESS: The Politics and Practices of Government Corporations. By *Annmarie Hauck Walsh.* Cambridge, Mass.: The MIT Press, 1980. Pp. 436, appendix, notes, bibliography, index. $9.95.

REGIONAL GROWTH: FLOWS OF FEDERAL FUNDS 1952-76. By *The Commission on Intergovernmental Relations.* Washington, D.C., 1980. Pp. ix, 141.

REGULATING SENSITIVE LANDS. By *Jon A. Kusler,* Cambridge, Mass.: Ballinger Book Company, 1980. Pp. viii, 248, appendix, bibliography, index. $25.00.

THE REWARDS OF PUBLIC SERVICE: COMPENSATING TOP FEDERAL OFFICIALS. By *Robert W. Hartman* and *Arnold R. Weber,* eds. Washington, D.C.: The Brookings Institution, 1980. Pp. xi, 238, index. $11.95 cloth, $4.95 paper.

ROLLING THUNDER: UNDERSTANDING POLICY AND PROGRAM FAILURE. By *James Clay Thompson.* Chapel Hill, N.C.: The University of North Carolina Press, 1980. Pp. xv, 199, notes, bibliography, index. $14.00 cloth, $6.50 paper.

RUFUS CHOATE: THE LAW AND CIVIL VIRTUE. By *Jean V. Alexander.* Philadelphia, Pa.: Temple University Press, 1980. Pp. 328, notes, bibliography, index.

SELF-STUDY PROCESSES: A GUIDE FOR POST-SECONDARY INSTITUTIONS. By *H. R. Kells.* Washington, D.C.: American Council on Education, 1980. Pp. xiii, 131, appendices.

THE STATE AND HUMAN SERVICES: ORGANIZATIONAL CHANGE IN A POLITICAL CONTEXT. By *Lawrence E. Lynn, Jr.* Cambridge, Ma.: The MIT Press, 1980. Pp. xii, 217, notes, bibliography, index. $15.00.